Panauti Journal
Volume 3, 2025
DOI: https://doi.org/10.3126/panauti.v3i01.83979 ISSN: 2990-7888 (Print) 2990-7896 (Online)

Agricultural Transformation in Nepal: An Econometric Analysis

Beg Prasad Garbuja[1*]

1. PhD Scholar, Central Department of Rural Development, T.U. Campus, Kirtipur, Kathmandu

Abstract

Agricultural transformation in Nepal, reflecting changes in productivity, structure, and economic contribution, is influenced by various socioeconomic and demographic factors. This study examines how these factors influence agricultural value added (% of GDP) in the country. For this purpose, data from two sources—the World Bank and the World Population Prospects 2024—were used, in which an econometric analysis approach was applied to assess the level of agricultural transformation in Nepal. While the overall relationship was statistically significant, the individual effects of the independent variables were too trivial to establish meaningful interactions. To address this issue, econometrics of three additional regression techniques—Ridge, Lasso, and Elastic Net—were employed to better identify the true contributions of the independent variables to the dependent variable. Key variables such as sex ratio, population growth rate, agriculture-related imports and exports, urban population, remittances, and per capita gross national income emerged as the most relevant factors in explaining variations in agriculture value added. The findings imply that targeted policies addressing demographic, economic, and trade-related factors are essential to effectively enhance agricultural value added and support Nepal's agricultural transformation.

Keywords: agricultural transformation, Pooled OLS, Ridge, Lasso, Elastic Net Regressions

Manuscript Received	Final Revision	Accepted
19 June, 2025	27 August, 2025	28 August, 2025

*B.P. Garbuja (garbeg436@gmail.com)

1. Introduction

Agricultural transformation represents the shift from traditional, subsistence-oriented farming methods to contemporary, commercially-oriented agricultural systems characterized by increased productivity (Pasa et al., 2024). This fundamental transition involves significant changes in food production systems, which are shaped by multiple interrelated factors. While technological progress, efficient resource allocation, and improved management techniques play crucial roles, demographic characteristics and social dynamics equally influence agricultural outcomes (Pasa, 2017). Population trends, labor availability, and societal structures interact with production systems, creating complex relationships that determine the pace and nature of agricultural development. The modernization process consequently reflects not just economic and technological changes, but also broader societal transformations that affect how food is produced and distributed. Variations in population size, structure, and behavior exert pressure on agricultural production through shifts in demand and labor supply. Conversely, changes within the agricultural sector can impact rural livelihoods and demographic trends.

In general, Malthus (1978) posited that population growth would inevitably outstrip food production, leading to resource scarcity, famine, and misery. Initially, Nepal seemed to fit the Malthusian narrative. Rapid population growth, coupled with limited technological advancements, strained food production capacity, particularly in the hills and mountains. Malthus underestimated technological progress (Boserup, 1965). The introduction of Green Revolution technologies (HYVs, fertilizers) temporarily boosted food production, challenging the Malthusian prediction in the short term (Wrigley, 2004). Nepalese agriculture's persistent reliance on conventional farming techniques, combined with the growing effects of climate change, continues to threaten future food security (Shrestha & Gurung, 2018), indicating that Malthusian warnings still hold some validity. The catastrophic 2015 earthquake exposed the fragility of Nepal's agricultural networks when supply chains collapsed (Gauchan et al., 2017), momentarily supporting Malthus' predictions about resource shortages during emergencies. Nevertheless, the nation's ability to avert a full-blown famine through prompt grain imports after the disaster revealed shortcomings in Malthus' original propositions, demonstrating how modern trade mechanisms can overcome temporary production deficits.

Marx (1867) focused on the social relations of production, arguing that unequal access to land, resources, and power creates exploitation and hinders agricultural development. Nepal's history of feudal land ownership patterns, caste-based inequalities, and limited access to credit and technology for smallholder farmers align with the Marxist critique. Landlessness and marginalization of certain ethnic groups exacerbate food insecurity. Land reform efforts in Nepal have been largely unsuccessful due to political resistance and implementation challenges (Bhandari & Linghorn, 2011). This has perpetuated unequal access to land, hindering agricultural productivity and reinforcing the Marxist critique of class-based exploitation in agriculture. The dominance of large landowners in certain regions contributes to the marginalization of small farmers and limits their access to markets and resources. The inequitable distribution of land and resources limits the productivity of smallholder farmers, constricting overall national food production capacity (Blaikie et al., 2002).

Boserup (1965) argued that population pressure stimulates agricultural innovation and intensification. Necessity is the mother of invention. Boserup's theory has some relevance in Nepal. Population growth in certain areas has driven farmers to adopt more intensive farming practices, such as terracing, intercropping, and the use of organic fertilizers. However, Boserup's model assumes access to resources and technology, which is not always the case in Nepal. Environmental degradation, climate change, and limited access to credit and extension services can constrain the ability of farmers to intensify production (Raji et al., 2024). In the Terai region, population density has encouraged the adoption of HYVs and irrigation, leading to increased rice production. However, this intensification has also led to environmental problems such as soil degradation and water scarcity, highlighting the limitations of Boserup's theory when resource management is inadequate. While population pressure has encouraged intensification, the absence of sustainable management practices and resource limitations constrain further increases in food production (Hossain & Debnath, 2019).

Schultz (1964) argued that traditional agriculture is not inherently inefficient but operates efficiently within its constraints. He emphasized the importance of investing in human capital, technology, and infrastructure to increase productivity and transition to modern agriculture. In the context of Nepal, Schultz’s concepts can be used to tackle issues like low agricultural productivity, restricted access to contemporary technology, and dependence on subsistence-level farming. By focusing on education, training, and the adoption of innovative agricultural practices, Nepal could enhance its agricultural output and support rural development.

Chayanov (1966) focused on the unique characteristics of peasant economies, where household labor and consumption needs drive agricultural production decisions. Chayanov's theory is highly relevant in Nepal, where a significant proportion of farmers are smallholders engaged in subsistence or semi-subsistence agriculture (Macfarlane, 1976). Household labor availability, family size, and consumption requirements heavily influence cropping patterns and production levels. Remittances have altered household labor dynamics, leading to a decline in agricultural labor and a shift towards less labor-intensive crops or even land abandonment in some areas. This supports Chayanov's emphasis on household-level decision-making in agricultural production (Fricke, 1986). The decline in agricultural labor due to remittances can lead to decreased production of traditional crops, shifting towards more market-oriented but less labor-intensive alternatives, potentially compromising food diversity and nutritional security.

2. Research Gap

Despite extensive theoretical and empirical insights from Malthus, Marx, Boserup, Schultz, and Chayanov, existing studies on Nepalese agriculture often examine population dynamics, land distribution, or technological adoption in isolation, without systematically integrating demographic, socioeconomic, and agricultural factors to explain national agricultural productivity. Moreover, prior research frequently relies on fragmented datasets, short time spans, or qualitative observations, limiting the ability to quantify the combined effects of these variables on agricultural value added (% of GDP). This study addresses this gap by employing a comprehensive, longitudinal dataset spanning 1960–2023, incorporating demographic indicators (population density, growth, fertility), economic variables (GNI, remittances, urbanization), and agricultural measures (crop, livestock, and food production indices, fertilizer use, land allocation). Using robust econometric techniques, including pooled OLS and regularized regression models (Ridge, Lasso, Elastic Net), the research systematically evaluates the relative importance and interplay of these factors, providing a more holistic and quantifiable understanding of the determinants of agricultural productivity in Nepal. In this background, the study aims to assess agricultural value added (% of GDP) in Nepal, which has been influenced by various socioeconomic, agricultural, and demographic variables.

3. Methodology

3.1 Data

This study used secondary sources of data on Nepal, generated from published reports and indexes (Martins et al., 2018). Specifically, data were collected from the World Bank and World Population Prospects websites, including population density, population growth rate, agriculture, forestry, and fishery value added, livestock production index, crop, food, and cereal production indexes, agricultural land, rural population, fertilizer consumption (kilograms per hectare of arable land), per capita PCA, and gross national income (see Annex).

3.2 Approach

This study adopted a hypothetico-deductive approach (Frankfort-Nachmias & Nachmias, 2008), a fundamental framework in scientific research that integrates both inductive and deductive reasoning to evaluate theories and expand understanding. The research process involves seven key steps: identifying the problem, developing a hypothesis, designing the study, defining measurements, collecting data, analyzing the data, and drawing general conclusions. Each step is influenced by theory and, in turn, influences theoretical understanding. A defining characteristic of the research process is its cyclical nature, usually beginning with a problem and culminating in provisional empirical generalizations.

The process starts with observations that lead to the development of a hypothesis (inductive reasoning). From this hypothesis, specific and testable predictions are generated (deductive reasoning). These predictions are then rigorously tested or further observed. If the results contradict the predictions, the hypothesis is modified or rejected. If the results support the predictions, the hypothesis is strengthened, although it remains subject to ongoing evaluation and is not considered definitively proven.

4. Hypothesis

Food production in Nepal is significantly influenced by population dynamics, agricultural sector performance, livestock and cereal production, rural population characteristics, economic indicators, and agricultural input usage.

4.1 Econometric Analysis

This hypothesis is designed to be tested using pooled ordinary least squares (OLS) regression analysis. The study regresses the food production index on the listed independent variables. If pooled OLS fails to adequately examine the interrelationship between variables, three additional regression analyses (Ridge, Lasso, and Elastic Net) will be used to assess the extent to which independent variables contribute to the dependent variable. The mathematical formulas for Ridge, Lasso, and Elastic Net regressions all build upon the ordinary least squares (OLS) regression formula but add penalty terms to address multicollinearity and improve model generalization (Hastie et al., 2009). The goal of OLS is to reduce the total of the squared differences (residuals) between the actual observed values and the values predicted by the model.

The formula for the estimated coefficients (β) is:

$\hat{\beta} = (X^T X)^{-1} X^T y$

Where:

$\hat{\beta}$ is the vector of estimated coefficients.

X is the design matrix (predictor variables).

y is the response vector (dependent variable).

X^T is the transpose of X.

$(X^T X)^{-1}$ is the inverse of the matrix product of X^T and X.

Ridge Regression. Ridge regression adds an L2 penalty term to the OLS objective function. This penalty is proportional to the sum of the squared coefficients.

*Minimize: $||y - X\beta||^2 + \lambda||\beta||^2$

Where:

$||y - X\beta||^2$ is the sum of squared residuals (same as in OLS).

λ (lambda) is the regularization parameter (a non-negative constant). It controls the strength of the penalty. Larger λ leads to more shrinkage.

$||\beta||^2$ is the L2 norm (sum of squared coefficients).

The solution for the estimated coefficients in Ridge regression doesn't have a closed-form solution like OLS, but it can be efficiently computed using numerical methods.

Lasso Regression. Lasso regression incorporates an L1 penalty into the OLS objective function, where the penalty is based on the sum of the absolute values of the model's coefficients.

*Minimize: $||y - X\beta||^2 + \lambda||\beta||_1$

Where:

$||y - X\beta||^2$ is the sum of squared residuals (same as in OLS).

λ (lambda) is the regularization parameter (a non-negative constant).

$||\beta||_1$ is the L1 norm (sum of absolute values of coefficients).

Like Ridge regression, Lasso doesn't have a closed-form solution and requires numerical methods for computation.

Elastic Net Regression. Elastic Net regression combines both the L1 and L2 penalty terms. The objective function is:

Minimize: $||y - X\beta||^2 + \lambda[\alpha||\beta||_1 + (1-\alpha)||\beta||^2]$

Where:

$||y - X\beta||^2$ is the sum of squared residuals (same as in OLS).

λ (lambda) is the regularization parameter (a non-negative constant).

α (alpha) is the mixing parameter ($0 \leq \alpha \leq 1$).

$\alpha = 0$: Pure Ridge regression.

$\alpha = 1$: Pure Lasso regression.

$0 < \alpha < 1$: Elastic Net regression (a blend of Ridge and Lasso).

$||\beta||_1$ is the L1 norm (sum of absolute values of coefficients).

$||\beta||^2$ is the L2 norm (sum of squared coefficients).

Again, Elastic Net requires numerical methods for solving the optimization problem.

All three methods build upon OLS regression by adding penalty terms to the objective function. The choice of method and the tuning parameters (λ and α) depend on the specific characteristics of the data and the goals of the analysis. Ridge regression is effective for handling multicollinearity, Lasso performs variable selection, and Elastic Net combines the benefits of both. The regularization parameters (λ) are typically chosen using cross-validation to optimize predictive performance.

5. Results and Discussions

This section presents the results of the analysis aimed at testing the hypothesis that food production in Nepal is influenced by a range of demographic, economic, and agricultural factors. The discussion integrates descriptive statistics, pooled OLS regression, and advanced regularized regression methods (Ridge, Lasso, and Elastic Net) to evaluate the contributions of population dynamics, agricultural sector performance, livestock and crop production, rural population characteristics, economic indicators, and input usage to agricultural productivity.

Prior to regression analysis, missing values in the dataset were addressed using a combination of imputation methods, including backfilling, linear interpolation, time series decomposition, and regression-based prediction, to ensure a complete and reliable dataset. The results highlight key patterns, identify robust predictors, and examine the relative importance of each factor, providing a comprehensive understanding of the determinants of food production in the Nepali context.

6. Methods of Imputation

Backfilling (Next Observation Carried Backward – NOCB)

Since the missing values in Nepal's agricultural data occur at the start of the series, the earliest available value, for example from 1965, can be used to fill in the preceding gaps. This approach is reasonable if we assume that the values in the early 1960s were similar to those in 1965. However, it should be used with caution, as it assumes no significant changes occurred during that period.

Linear Interpolation (Extrapolation)

The trend from the available data can be extrapolated to estimate the missing values for Nepal. Since data for several years after the missing period are available, a linear model can be fitted to the existing data and projected backward. This assumes a relatively constant linear trend. While simple to implement and understand, this method may be inaccurate if the variable does not follow a linear trend.

Time Series Decomposition and Extrapolation

A more sophisticated approach involves decomposing the time series for Nepal's agricultural indicators into trend, seasonal, and random components. The trend component can then be extrapolated backward, and the seasonal and random components added back (if applicable). This method is suitable if the data exhibit seasonality or a clear trend.

Regression with External Predictors

If other variables correlated with Nepal's agriculture value added (% of GDP) (AGR_GDP) are available—such as rural population percentage, fertilizer use, or agricultural land area—a regression model can be built to predict AGR_GDP based on these predictors. The model can then be used to impute missing values for 1960–1964.

The study combined all four approaches. For example, linear extrapolation was applied first and then refined using a regression model with external predictors. The following table summarizes the socioeconomic and agricultural variables for Nepal (Table 1).

Table 1 *Descriptive Statistics*

Variable	Mean	SD	Min	Max	Skew	Kurtosis	SE	VIF
Population Density	137.64	44.26	68.78	201.90	-0.08	-1.54	5.53	1.34
Sex Ratio	100.00	3.85	91.53	103.62	-1.00	-0.52	0.48	1.53
Population Growth Rate	1.70	0.80	-0.14	3.00	-0.88	-0.58	0.10	1.36
Total Fertility	4.43	1.49	1.98	6.09	-0.48	-1.46	0.19	1.33
Life expectancy	55.62	10.51	38.70	70.40	-0.13	-1.52	1.31	1.41
Agriculture Added Value (% of GDP)	44.28	15.11	20.91	69.01	0.20	-1.23	1.89	1.67

Variable	Mean	SD	Min	Max	Skew	Kurtosis	SE	VIF
Livestock Production Index	62.38	31.96	25.51	144.45	1.06	0.38	3.99	1.54
Crop Production Index	53.50	31.83	20.08	116.66	0.63	-1.00	3.98	1.44
Food Production Index	55.35	31.83	21.05	123.18	0.71	-0.78	3.98	1.48
Agriculture related Import	1.75	0.98	0.00	5.32	1.35	2.93	0.12	1.46
Agriculture Related Export	12.92	11.84	0.45	48.02	1.42	1.35	1.48	1.46
Urban (% of Total Population)	10.69	5.94	3.48	21.90	0.34	-1.34	0.74	1.84
Remittances (% of GDP)	7.67	10.28	0.03	27.63	0.85	-1.11	1.29	1.51
Dependency on 15-65	75.08	8.16	54.02	82.41	-1.31	0.40	1.02	1.52
Informal Economy (DGE method)	44.06	9.36	30.04	61.80	0.25	-1.20	1.17	1.37
Fertilizer Consumption	27.58	29.61	0.08	115.33	1.28	0.87	3.70	1.61
Gross National Income Per Capita	346.41	377.76	50.00	1430.00	1.52	1.05	47.22	1.59
Mean Years of Schooling	2.55	0.95	1.26	4.49	0.49	-0.88	0.12	1.21

Population density shows a relatively low skew (-0.08), suggesting a fairly symmetrical distribution. The sex ratio (100.00) indicates a near-equal balance between males and females. Population growth rate (1.70) is moderate, with a negative skew (-0.88), indicating a concentration of values below the mean. Total fertility (4.43) is relatively high, with a slightly negative skew (-0.48). Life expectancy (55.62) is moderate, with a near-symmetrical distribution.

Agricultural value added as a percentage of GDP (44.28) is substantial, with a slight positive skew (0.20). Livestock, crop, and food production indices exhibit positive skews, indicating a concentration of values below the mean, possibly suggesting a significant portion of the population engaged in low-productivity agriculture. Agriculture-related imports (1.75) and exports (12.92) show high positive skews, suggesting a concentration of values below the mean for imports and a more dispersed distribution for exports.

Urbanization (10.69% of the total population) is relatively low, with a positive skew (0.34). Remittances as a percentage of GDP (7.67%) are moderate, with a positive skew (0.85). The dependency ratio (75.08) is high, with a negative skew (-1.31), indicating a relatively large proportion of dependents. The informal economy (44.06%) is substantial, with a slight positive skew (0.25). Fertilizer consumption (27.58) shows a high positive skew (1.28), suggesting that a small proportion of farmers use a large amount of fertilizer. Gross national income per capita (346.41) is low, with a high positive skew (1.52), indicating a large disparity in income distribution. Mean years of schooling (2.55) is low, with a positive skew (0.49).

In this dataset, VIF values remain relatively low, suggesting minimal multicollinearity concerns. However, higher VIF values in certain variables, such as agriculture-related indicators, suggest potential redundancy or correlation among predictors. Addressing high VIF can improve model reliability by ensuring independent variables provide distinct contributions.

The study used multiple imputation methods to address missing values in Nepal's agricultural and socioeconomic data. Descriptive statistics show moderate population density, high fertility, and moderate life expectancy. Agriculture contributes substantially to GDP, with livestock, crop, and food production skewed toward lower productivity. Urbanization and schooling are low, while income and fertilizer use show high variability. VIF values indicate minimal multicollinearity, providing a reliable basis for regression analysis.

Outputs of Pooled OLS Regression

"Outputs of Pooled OLS Regression" present the estimated results, showing how the independent variables collectively and individually affect the dependent variable, providing a baseline analysis of their relationships.

Table 1 evaluates the pooling model that examines the relationship between various socio-economic and demographic factors and agricultural GDP (AGR_GDP). The model is based on an unbalanced panel dataset comprising 64 observations across a defined time span, allowing for the assessment of different predictors and their impact on agricultural GDP.

The use of an unbalanced panel (n = 64, T = 1–1) indicates variability in the time dimension across entities, which can affect estimation reliability. The residuals range from -26.60 to 30.27, suggesting some predictions deviate significantly from observed values, indicating potential issues such as heteroscedasticity or model misspecification. The model's R-squared value of 0.40865 indicates that approximately 41 percent of the variance in AGR_GDP is explained by the included variables, which is relatively low and suggests that important predictors may be missing. The adjusted R-squared value of 0.1901 highlights that further variable addition yields diminishing returns, raising concerns about model efficiency. The F-statistic of 1.87 with a p-value of 0.047126 indicates that the model as a whole is statistically significant at the 5 percent level, suggesting that the predictors collectively provide a better fit than a null model.

The coefficient estimates show mixed signs, and many lack statistical significance. For example, the coefficient for population density (POP_DENS) is negative but not significant, suggesting negligible declines in agricultural GDP as population density increases. The sex ratio (SEX_RATIO) coefficient is positive (1.1280) and approaches significance (p = 0.0563), hinting that a balanced sex ratio may enhance agricultural productivity.

The total fertility rate (TOT_FER) shows a negative relationship with AGR_GDP that is not significant, challenging expected theories regarding high fertility rates and productivity. Variables related to agricultural indices (LIVE_IND, CROP_IND, FOOD_IND), agricultural trade (AGR_IMP, AGR_EXP), and economic dependency show coefficients that, while informative, largely lack statistical significance. The coefficient of agricultural exports is notably negative (-0.2525), contrary to expectations that greater exports would boost GDP. The mean years of schooling (MEAN_SCH) has a positive but insignificant effect (2.0027, p = 0.2405) on agricultural productivity, suggesting that educational attainment may not directly impact agricultural GDP in this context.

Overall, the pooling model reveals some insights but also significant limitations. With low explanatory power and many non-significant predictors, the model suggests several areas for improvement. Future analyses could benefit from refining the variable selection process to include potentially impactful factors such as technological adoption, market access, and environmental variables. Attention should also be given to data quality to ensure a balanced panel approach or the application of imputation methods to address data sparsity. Finally, exploring the broader socio-economic context beyond the current variables may enrich our understanding of the dynamics affecting agricultural GDP. This model serves as a preliminary exploration that underscores the need for a more comprehensive approach to effectively capture the complexities of agricultural productivity.

The summary statistics alone do not provide a complete picture of the data. Further analysis, including correlation analysis, regression modeling, and visualization techniques, is needed to fully understand the relationships between these variables and their implications for agricultural transformation. The high positive skewness in several variables suggests the presence of outliers, which should be investigated further. In this context, the study further analyzes panel data (1960–2023) using three regressions (Ridge, Lasso, and Elastic Net), employing methods designed for panel data that account for the correlation structure and time dependence (Table 2).

Table 2 compares the outputs of three regularized regression methods—Ridge, Lasso, and Elastic Net—applied to Nepal's agricultural and socioeconomic dataset. The table highlights differences in variable selection, coefficient shrinkage, intercept values, and key predictors across the models. Ridge regression retains all variables with smaller coefficients due to shrinkage, while Lasso and Elastic Net perform variable selection by setting many coefficients to zero, simplifying the model. The table also identifies variables consistently important across all models, as well as those with less reliable or inconsistent effects.

Table 2 *Comparison of the Regression Outputs*

Feature	Ridge Regression	Lasso Regression	Elastic Net Regression	Analysis
Variable Selection	No variables are excluded. All variables have non-zero coefficients (though some are very small).	Many variables have coefficients set to zero (indicated by "." in the output), indicating variable selection.	Similar to Lasso, many variables have coefficients set to zero, suggesting variable selection.	Lasso and Elastic Net are explicitly performing variable selection, simplifying the model. Ridge retains all variables, which may be useful if all variables have some predictive power, or the goal is only to improve prediction, not interpretation.

Feature	Ridge Regression	Lasso Regression	Elastic Net Regression	Analysis
Coefficient Size/ Shrinkage	Coefficients are generally smaller in magnitude compared to a standard linear regression, due to the shrinkage effect of Ridge.	Coefficients of the remaining variables (those not set to zero) tend to be larger in magnitude than in Ridge, as they compensate for the excluded variables.	Coefficients of the remaining variables tend to be between those of Ridge and Lasso.	The shrinkage effect is evident across all three methods, but most pronounced in Lasso. Larger coefficients in Lasso and Elastic Net indicate that the selected variables have a stronger individual impact.
Intercept	Relatively large positive intercept (14.62).	Negative intercept (-0.64).	Larger negative intercept (-3.05) than lasso.	The intercepts differ significantly, reflecting the different variable selection and coefficient shrinkage effects.
Key Predictors (Consistent Across Models)	SEX_RATIO, POPA_GRAT, AGR_IMO, AGR_EXP, URB_PER, REM_GDP, GNI_PCA	SEX_RATIO, POPA_GRAT, AGR_IMO, AGR_EXP, URB_PER, REM_GDP, GNI_PCA.	SEX_RATIO, POPA_GRAT, AGR_IMO, AGR_EXP, URB_PER, REM_GDP, GNI_PCA.	These variables consistently appear as important predictors across all three models, suggesting that they are robust and reliable predictors of the outcome variable. These have the largest coefficients in Ridge.

Feature	Ridge Regression	Lasso Regression	Elastic Net Regression	Analysis
Inconsistent Predictors	POP_DENS, TOT_FER, LIFE_EXPT, LIVE_IND, CROP_IND, FOOD_IND, DEP_ECO, ING_DGE, FERT_CON, MEAN_SCH	POP_DENS, TOT_FER, LIFE_EXPT, LIVE_IND, FOOD_IND, DEP_ECO, ING_DGE, FERT_CON, MEAN_SCH	POP_DENS, TOT_FER, LIFE_EXPT, LIVE_IND, FOOD_IND, DEP_ECO, ING_DGE, FERT_CON, MEAN_SCH, CROP_IND	These are variables that are deemed less important by Lasso and Elastic Net, and have small coefficients in Ridge. These predictors might be less reliable, or their effects may be captured by the other variables.
CROP_IND	Very small negative coefficient.	CROP_IND still appears with a (small) coefficient, but is mostly discarded	Lasso regression removes it entirely.	The differences in CROP_IND behavior illustrates how model selection choices affect the interpretation. Elastic Net is more conservative than Lasso, retaining slightly more variables.
General Model Complexity	Most complex (all variables included).	Least complex (most variables excluded).	Intermediate complexity.	The models have different levels of complexity, with Ridge being the most complex and Lasso being the simplest. The choice of complexity depends on the goals of the analysis.

The comparison shows that Lasso and Elastic Net provide simpler, more interpretable models by excluding less important predictors, whereas Ridge retains all variables, which may improve predictive performance but complicates interpretation. Key predictors such as sex ratio, population growth, agricultural imports and exports, urban population, remittances, and per capita GNI consistently emerge across all models, indicating their robust influence on the outcome. The choice of model should balance predictive accuracy and interpretability based on the study's goals.

The study effectively addresses the hypothesis that food production in Nepal is influenced by population dynamics, agricultural sector performance, livestock and crop production, rural population characteristics, economic indicators, and input usage. Using a combination of descriptive statistics, pooled OLS regression, and regularized regression methods (Ridge, Lasso, and Elastic Net), the analysis evaluated the impact of these variables on agricultural GDP and the food production index. Results indicate that factors such as population growth, sex ratio, agricultural trade (imports and exports), urban population, remittances, and per capita GNI consistently emerge as significant predictors across multiple models, highlighting their robust influence on agricultural outcomes. Conversely, variables like population density, total fertility, crop and livestock production indices, fertilizer consumption, and mean years of schooling show limited or inconsistent effects, suggesting that their influence on food production may be indirect or context-dependent. Overall, the findings demonstrate that while many hypothesized factors are relevant, their significance varies, providing a nuanced understanding of the determinants of agricultural productivity in Nepal.

7. Conclusion and Policy Recommendations

This study confirms and partially refines the hypothesis that Nepal's food production is influenced by demographic, economic, and agricultural factors. Using multiple imputation, descriptive statistics, pooled OLS regression, and regularized regressions (Ridge, Lasso, and Elastic Net), the analysis identifies key determinants—such as sex ratio, population growth, agricultural trade, urban population, remittances, and per capita GNI—that consistently influence agricultural GDP, supporting the hypothesized relationships. Other factors included in the hypothesis, such as population density, total fertility, crop and livestock production indices, fertilizer consumption, and mean years of schooling, showed limited or inconsistent effects, suggesting their impact may be indirect or context-dependent.

Overall, the findings highlight the complex interplay of socio-economic and agricultural variables in shaping food production, underscoring the need for targeted policies that address demographic pressures, strengthen economic conditions, and promote efficient agricultural practices. Future research should consider additional variables, such as technological adoption, market access, and environmental factors, to develop a more comprehensive understanding of the determinants of agricultural productivity in Nepal.

References

Bhandari, R., & Linghorn, A. (2011). The four pillars of successful land reform: Can revolutionary Nepal stand up? *Himalayan Journal of Development and Democracy, 6*(2), 36–50. https://digitalrepository.unm.edu/cgi/viewcontent.cgi?article=1181&context=hprc

Blaikie, P., Cameron, J., & Seddon, D. (2002). *Nepal: Poverty and stagnation*. Adroit Publishers.

Boserup, E. (1965). *The condition of agricultural growth: The economics of agrarian change under population pressure*. Allen & Unwin.

Chayanov, A. V. (1966). *The theory of peasant economy* (D. Thorner, B. Kerblay, & R. E. F. Smith, Eds.). Richard D. Irwin, Inc.

Creswell, J. W. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.

Gauchan, D., Joshi, B. K., & Ghimire, K. (2017, December). Impact of 2015 earthquake on economy, agriculture and agrobiodiversity in Nepal. In *Proceedings of National Sharingshop: Rebuilding local seed system of native crops in earthquake affected areas of Nepal* (18 Dec 2017, Kathmandu, Nepal). NAGRC, Bioversity International, & Crop Trust. https://www.researchgate.net/publication/322364093_Impact_of_2015_Earthquake_on_Economy_Agriculture_and_Agrobiodiversity_in_Nepal

Frankfort-Nachmias, C., & Nachmias, D. (2008). *Research methods in the social sciences* (7th ed.).

Fricke, T. E. (1986). *Himalayan households: Tamang demography and domestic processes*. UMI Research Press.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). https://web.stanford.edu/~hastie/ElemStatLearn/

Hossain, F., & Debnath, S. (2019). Agricultural transformation and rural development in Nepal: Issues and challenges. *The Journal of Development Practice, 5*(1), 45–63.

Macfarlane, A. (1976). *Population and resources*. Adroit Publication.

Malthus, T. R. (1978). *An essay on the principle of population*. Johnson.

Martins, F. S., Cunha, J. A. C. da, & Serra, F. A. R. (2018). Secondary data in research – Uses and opportunities. *Revista Ibero-Americana de Estratégia, 17*(4), 1–4. https://doi.org/10.5585/ijsm.v17i4.2723

Marx, K. (1867). *Das kapital (capital)* (Vol. I). Verlag von Otto Meissner.

Pasa, R. B. (2017). Technological intervention in agriculture development. *Nepalese Journal of Development and Rural Studies, 14*(1–2), 86–97. https://doi.org/10.3126/njdrs.v14i1-2.19652

Pasa, R. B., Joshi, H. R., Acharya, M., & Adhikari, J. (2024). Appraising agriculture development strategy of Nepal from agricultural extension perspective. *Journey for Sustainable Development and Peace, 2*(2, Special Issue), 146–160. https://doi.org/10.3126/jsdpj.v2i2.69575

Raji, E., Ijomah, T., & Eyieyien, O. G. (2024). Improving agricultural practices and productivity through extension services and innovative training programs. *International Journal of Applied Research in Social Sciences, 6*(7), 1297–1309. https://doi.org/10.51594/ijarss.v6i7.1267

Schultz, T. W. (1964). *Transforming traditional agriculture*. Yale University Press.

Shrestha, S. K., & Gurung, G. P. (2018). Impacts of climate change in Nepalese agriculture. *Nepalese Journal of Development and Rural Studies, 15*, 114–123. https://doi.org/10.3126/njdrs.v15i0.31605

Wrigley, E. A. (2004). *Poverty, progress, and population*. Cambridge University Press.

Annex: Data Sources

1. Agriculture Land (%of Land Areas). https://data.worldbank.org/indicator/AG.LND.AGRI.ZS
2. Agriculture, Forestry, and Fishery Value Added (%of GDP). https://data.worldbank.org/indicator/NV.AGR.TOTL.ZS
3. Cereal Production. https://data.worldbank.org/indicator/AG.PRD.CREL.MT
4. Fertilizer Consumption: Kilograms per Hector of Arable Land https://data.worldbank.org/indicator/AG.CON.FERT.ZS
5. Food Production Index. https://data.worldbank.org/indicator/AG.PRD.FOOD.XD

6. GDP percent of per Capita PCA. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD
7. Gross National Income (GNI) per Capita Income. https://data.worldbank.org/indicator/NY.GNP.ATLS.CD
8. Livestock Production Index. https://data.worldbank.org/indicator/AG.PRD.LVSK.XD Crop Production Index. https://data.worldbank.org/indicator/AG.PRD.CROP.XD
9. Population Density, Population Growth Rate (Annual), and Life Expectancy at Birth; https://population.un.org/wpp/downloads?folder=Standard%20Projections&group=Most%20used
10. Rural Population (% of Total Population https://data.worldbank.org/indicator/SP.RUR.TOTL.ZS

~O~