tourism and hospitality

MDPI

Article

Accommodation-Based Tourism Demand and Rural Labour-Market Dynamics in Kazakhstan: Evidence from a Regional Panel

Gulzhan Abdramanova [1], Bakyt Syzdykbayeva [1,2] and Zhanar Dulatbekova [1,*]

[1] Higher School of Economics and Tourism, L. N. Gumilyov Eurasian National University, Astana 010008, Kazakhstan; abdramanovag.25@gmail.com (G.A.); bakyt_syzdykbaeva@mail.ru (B.S.)
[2] Department of Tourism and Service, Kazakh Academy of Sport and Tourism, Almaty 050022, Kazakhstan
* Correspondence: dulatbekovazhanar1@gmail.com

Abstract

Tourism-related opportunities may be unevenly distributed within regions where visitor activity and economic linkages are spatially concentrated. This study examines associations between regionwide accommodation-visitor growth and rural labour-market changes in Kazakhstan. Using a harmonised panel of 14 regions over 2003–2025, we estimate first-difference regressions with year fixed effects, region-clustered standard errors, and wild-cluster bootstrap inference. The principal sample contains 293 observations (290 with regional economic and agricultural controls). Neither rural employment intensity nor rural unemployment shows a statistically significant contemporaneous association with accommodation-visitor growth. Distributed-lag models fail to reject joint null hypotheses for current and lagged tourism-demand coefficients at the 5% level. Targeted temporal, territorial, and control-based checks likewise yield no statistically significant contemporaneous associations at the 5% level. However, exploratory specifications reveal statistically significant negative associations with rural employment intensity and, in restricted annual samples, rural unemployment. These findings are sensitive to sample selection and measurement horizon and do not establish causal effects. Accommodation visitors are measured regionwide, including urban destinations, while labour-market indicators cover rural residents across sectors. Consequently, the estimates identify neither tourism-specific employment effects nor mechanisms transmitting tourism-related opportunities to rural communities. The findings highlight spatial scale, indicator coverage, and statistical uncertainty.

Keywords: tourism demand; accommodation establishments; rural labour market; employment intensity; rural unemployment; territorial transmission; spatial scale; regional panel data; Kazakhstan

Check for updates

Academic Editor: Lewis Ting On Cheung

Received: 4 October 2026
Revised: 29 October 2026
Accepted: 1 November 2026
Published: 4 November 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

Tourism has been widely examined as a potential driver of regional economic diversification, particularly through employment opportunities, enterprise development, and economic linkages between tourism businesses and other sectors (Ashley et al., 2001; Mitchell & Ashley, 2010). Research on tourism and community development further emphasises that the distribution of these opportunities depends on local participation, institutional arrangements, and the capacity of residents and enterprises to benefit from tourism-related activity (Scheyvens, 2007; Sharpley & Telfer, 2014). At a broader geographical scale, empirical studies demonstrate that the economic implications of tourism vary across territories

and are shaped by spatial spillovers, agglomeration, productive structures, and urban–rural conditions (Faber & Gaubert, 2019; Comerio et al., 2020; Kim et al., 2021; Wang & Bai, 2023). Consequently, growth in tourism demand within an administrative region does not necessarily indicate where the associated economic opportunities arise or how they are distributed among its residents.

This distinction is especially relevant to rural development. Rural tourism can contribute to economic diversification by connecting accommodation services with agriculture, food production, cultural resources, transport, local enterprises, and other activities (Sharpley, 2002; Saxena et al., 2007; Cawley & Gillmor, 2008). Tourism-related opportunities may reach surrounding communities through employment, local procurement, entrepreneurship, and participation in tourism value chains (Meyer, 2007; Mitchell, 2012). However, the extent of these connections depends on local production capacity, business relationships, ownership, and community participation (Adiyia & Vanneste, 2018; Y. Li et al., 2024; Chambwe & Saayman, 2025). Increased visitor activity therefore does not, by itself, establish whether rural residents obtain additional employment or other economic benefits.

Regional accommodation-visitor counts encompass both urban and rural destinations, whereas rural labour-market indicators cover residents across all economic sectors. Although their association can be examined at the regional level, these aggregate measures cannot identify tourism-specific rural employment or establish how visitor-related opportunities are distributed within regions.

The concept of territorial transmission provides an interpretive perspective on how tourism-related economic opportunities arising within a wider region may reach rural residents and enterprises through employment, procurement, entrepreneurship, and local supply chains (Meyer, 2007; Mitchell, 2012; Adiyia & Vanneste, 2018). Because these channels are not directly observed, the concept is used to frame the cross-scale relationship rather than to claim that territorial transmission mechanisms are empirically identified.

Kazakhstan provides a relevant setting for examining this cross-scale relationship. Its administrative regions encompass combinations of urban centres, tourism destinations, and extensive rural territories, while local conditions for rural tourism development vary geographically. For the full study period, the available official statistics did not permit the construction of a consistent, territorially comparable series linking tourism activity in rural destinations with tourism-sector employment among rural residents. Accordingly, this study examines whether annual changes in regionwide accommodation-based tourism demand co-move with changes in rural employment intensity and rural unemployment within Kazakhstan's analytical regions.

This study contributes to the tourism and regional-development literature in three respects. First, it examines a cross-scale relationship between regionwide accommodation-based tourism demand and labour-market conditions among rural residents, complementing evidence from individual destinations and communities. Second, it provides longitudinal evidence from a harmonised panel of 14 analytical regional units in Kazakhstan over 2003–2025, assessing whether annual associations are sensitive to regional economic and agricultural conditions and lagged tourism-demand changes. Third, it reports the statistical uncertainty surrounding these associations and clarifies what aggregate indicators can and cannot reveal about rural labour-market dynamics. The analysis does not identify causal effects or the mechanisms of territorial transmission.

2. Literature Review

2.1. Tourism-Led Regional Development and the Territorial Distribution of Economic Benefits

Tourism can contribute to regional economic development, but its economic implications are not necessarily distributed uniformly across space. Evidence from tourism-

intensive areas in Mexico shows substantial local economic gains and positive spillovers to manufacturing, while aggregate effects differ from the accumulation of local gains (Faber & Gaubert, 2019). Positive tourism and economic spillovers have also been identified across Japanese prefectures, where neighbouring tourism activity is associated with regional economic dynamics (Comerio et al., 2020). Tourism-industry productivity likewise varies with agglomeration conditions and spatial spillovers across regions (Kim et al., 2021). Regional heterogeneity has also been documented in the relationship between tourism development and the urban–rural income gap, with rural infrastructure acting as a mediating mechanism (Wang & Bai, 2023). These findings indicate that tourism-related economic outcomes depend partly on territorial conditions and the spatial organisation of tourism activity.

Tourism activity may be concentrated in particular destinations within broader administrative regions. Evidence of spatial clustering in rural tourism links this variation to tourism-market potential, transport capacity, economic development, topographical conditions, and policy factors (Du et al., 2024). This concentration is relevant to the interpretation of regionwide tourism indicators.

Rural tourism research provides a more local perspective on the relationship between visitor activity and rural development. Rural tourism has been promoted as a source of income and employment in rural areas and as a means of diversifying tourism activity and extending tourism benefits beyond established destinations (Sharpley, 2002). Integrated rural tourism further emphasises the connection of tourism with the economic, social, cultural, natural, and human structures of rural localities and with networks linking local resources, activities, products, and communities (Saxena et al., 2007). Similar research highlights the importance of combining local resources and stakeholder participation in ways that add value within rural economies (Cawley & Gillmor, 2008).

The presence of tourism demand, however, does not establish the extent to which local residents and enterprises participate in the resulting economic activity. Tourism-related linkages with neighbouring communities may operate through employment, local sourcing and procurement, small-enterprise development, outsourcing, and other forms of partnership (Meyer, 2007). Value-chain analysis similarly demonstrates the importance of tracing tourism expenditure through the local economy to determine which groups participate in and benefit from tourism activity (Mitchell, 2012). Evidence from western Uganda shows that local supply-chain linkages can be constrained by production and supply conditions, while opportunities for local suppliers to participate in tourism value chains depend on the organisation of relationships between tourism businesses and surrounding communities (Adiyia & Vanneste, 2018). Local ownership and meaningful community participation in tourism value chains have also been identified as important components of rural community development (Chambwe & Saayman, 2025).

Research on traditional villages further shows that rural tourism may generate economic and social benefits while also creating pressures and disruptions. Sustainable outcomes depend on the preservation of community values, stakeholder interests, livelihood diversity, and meaningful community participation rather than on tourism growth alone (Y. Li et al., 2024). These findings reinforce the distinction between the volume of tourism activity and the way tourism-related opportunities are organised and distributed within rural territories.

Rural tourism potential in Kazakhstan also varies at the subnational level. Rural districts of Aktobe Oblast have been assessed in terms of natural resources, social characteristics, infrastructure, cultural and recreational resources, and environmental conditions, with substantial differences identified across districts (Saparov et al., 2024). Such evidence demonstrates territorial variation in the conditions for rural tourism development but does

not establish whether growth in regionwide tourism demand is accompanied by labour-market changes among rural residents across Kazakhstan.

2.2. Tourism, Local Economic Linkages and Rural Labour-Market Outcomes

Tourism-related activity can influence labour markets through direct employment in tourism establishments and through additional demand for goods and services supplied by other sectors. However, tourism-sector employment and broader labour-market conditions should be distinguished. Growth in tourism employment may occur without producing comparable changes in employment conditions across the wider population, particularly when tourism activity is concentrated in specific destinations or relies on workers and suppliers from outside surrounding rural communities. Research on tourism work also shows that employment outcomes involve not only the number of jobs but also skills, working conditions, mobility, career opportunities, and other characteristics of tourism employment (Ladkin et al., 2023). Regional tourism can contribute to employment generation while still being associated with relatively low-paid or precarious occupations and limited training and career opportunities for some workers (Kronenberg & Fuchs, 2021).

Rural tourism provides employment opportunities within particular destinations, but these opportunities are not necessarily distributed uniformly among rural residents. Evidence from a rural tourism destination in China shows that tourism supported employment in the study area, while the tourism workforce remained disproportionately young, male, and relatively low-skilled, and labour mobility extended beyond the immediate locality (T. Li et al., 2021). Such findings illustrate why the existence of tourism-related jobs within a rural destination does not imply an equivalent improvement in labour-market conditions for the entire rural population of a larger administrative region.

Destination-level tourism development may also be associated with changes in the structure of rural employment. Evidence from the Jinshitan scenic area in China documents substantial growth in the share of non-agricultural employment alongside the expansion of tourism and wider processes of touristification and rural revitalisation (Yang et al., 2021). This type of evidence demonstrates that tourism can become closely connected to employment restructuring in particular destinations. However, such locally observed changes cannot be assumed to occur uniformly across all rural areas contained within broader regions, especially where tourism activity is spatially concentrated.

Participation in rural tourism employment also depends on household and local conditions. Rural households differ in their willingness and capacity to engage in tourism-related activities, and empirical evidence shows that market characteristics, household understanding of tourism, income sources, and proximity to tourism services can influence decisions to participate in rural tourism (Luo et al., 2022). These factors indicate that visitor demand does not translate automatically into labour-market participation among nearby rural households. The relationship depends on whether residents have access to tourism markets, relevant information, suitable economic opportunities, and the conditions required to provide tourism-related services.

Temporal conditions provide another source of variation in the relationship between tourism and employment. Tourism-related employment does not necessarily follow the same cyclical pattern as employment in other sectors, and tourism activity may under some conditions contribute to the stabilisation of cyclical employment (Kožić & Sever, 2022). Investment, enterprise development, and the establishment of local supply relationships may also require time before changes in tourism demand are reflected in labour-market conditions. Major external shocks can nevertheless alter these relationships considerably. The COVID-19 pandemic produced substantial losses in tourism employment and labour income across countries, with particularly high vulnerability among women, young workers,

and lower-income groups (Sun et al., 2022). These considerations motivate examining both contemporaneous and lagged associations and assessing the sensitivity of annual estimates to the pandemic period.

2.3. Research Gap

Destination- and household-level studies document tourism-related employment restructuring and differences in local participation (Yang et al., 2021; T. Li et al., 2021; Luo et al., 2022). However, these findings concern geographical units and populations that differ from broader administrative regions, where tourism activity and rural residents may be spatially separated. Regional tourism research further indicates that economic outcomes vary across territories and depend on local conditions (Faber & Gaubert, 2019; Comerio et al., 2020; Kim et al., 2021; Wang & Bai, 2023). Nevertheless, evidence from destination-level employment and tourism value chains does not establish whether annual changes in accommodation-based tourism demand measured for broader administrative regions are systematically associated with labour-market changes among rural residents within the same regions. This empirical gap concerns the correspondence between tourism demand and labour-market outcomes across geographical scales, rather than whether tourism can generate employment in individual rural destinations.

The present study addresses this gap using a territorially harmonised regional panel to assess contemporaneous and delayed associations while accounting for selected changes in regional economic and agricultural conditions. This approach allows the strength, precision, and sensitivity of aggregate regional–rural associations to be evaluated across alternative specifications. However, the estimated relationships do not identify causal effects or the employment, procurement, and enterprise linkages through which tourism-related opportunities may reach rural communities (Meyer, 2007; Mitchell, 2012; Adiyia & Vanneste, 2018; Chambwe & Saayman, 2025).

3. Methodology

3.1. Data and Variables

The empirical analysis uses annual regional data for Kazakhstan covering 2003–2025. The dataset is organised as a balanced panel of 14 analytical regional units and 23 years, yielding 322 region-year observations before first differencing. The analysis relies on four principal series: visitors served by accommodation establishments, rural population, employed rural residents, and unemployed rural residents. These variables are used to construct one measure of regional accommodation-based tourism demand and two measures of rural labour-market conditions. Additional regional indicators of economic activity, agricultural production, rural demographic composition, and accommodation supply are incorporated in the extended specifications and sensitivity analyses.

Statistical series were obtained from official publications and datasets of the Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan (BNS). Annual visitors served by accommodation establishments are used as an operational proxy for regionwide accommodation-based tourism activity, based on the official BNS dataset Number of Served Visitors in Accommodation Establishments (BNS, n.d.-f). This indicator captures recorded accommodation activity across each analytical region but does not distinguish between establishments located in urban and rural areas. Accordingly, it represents regionwide accommodation-based tourism demand rather than a direct measure of rural tourism.

Rural population is measured as the number of residents officially classified as living in rural areas and is obtained from the BNS population series (BNS, n.d.-h). Rural population is retained in the analysis because it forms the denominator of the rural employment-

intensity indicator. Rural population is not treated as a primary labour-market outcome. However, changes in its underlying count are examined separately in diagnostic decompositions to distinguish movements in the numerator and denominator of the employment-intensity indicator.

Rural employment and unemployment data were compiled from successive official statistical publications issued by the institutions responsible for state statistics in Kazakhstan, as no single continuous source was used for the complete 2003–2025 period. Observations for 2003–2012 were obtained from successive editions of Economic Activity of the Population of Kazakhstan (Agency of the Republic of Kazakhstan on Statistics, 2008, 2013). Observations for 2013–2022 were taken from successive editions of Employment in Kazakhstan, available through the BNS statistical collection archive (BNS, n.d.-d). Values for 2023–2025 were obtained from subsequent BNS employment and unemployment publications and spreadsheets containing regional data for employed and unemployed rural residents (BNS, n.d.-b, n.d.-c). Where publication periods overlapped, observations were taken from the source assigned to the corresponding year in the data-compilation protocol rather than averaged across editions.

Additional regional indicators were compiled from official BNS statistical series to account for factors that may vary across analytical regions over time. The gross regional product volume index and agricultural production volume index are used to represent changes in regional economic and agricultural activity. Demographic conditions are examined using the share of the rural population aged 15–64, while the number of accommodation establishments located in rural areas provides an additional indicator of rural accommodation supply. The latter variable measures the availability of accommodation establishments rather than the number of visitors or tourism-related jobs in rural areas. Because the additional statistical series differ in temporal coverage and completeness, the corresponding controlled and sensitivity specifications are estimated using the available observations.

Territorial harmonisation was required because Kazakhstan's administrative structure changed during the study period. The panel is therefore organised around 14 analytical regional units: Akmola, Aktobe, Almaty–Zhetysu, Atyrau, West Kazakhstan, Zhambyl, Karagandy–Ulytau, Kostanay, Kyzylorda, Mangystau, South Kazakhstan/Turkistan, Pavlodar, North Kazakhstan, and East Kazakhstan–Abai. Astana and Almaty are not included as separate analytical units because this study focuses on regional rural labour-market outcomes. The combined regional names refer to harmonised analytical units rather than administrative regions existing under those names throughout the entire study period. In particular, Karagandy–Ulytau denotes the historical Karagandy Region before the 2022 territorial reform and the aggregation of Karagandy and Ulytau regions thereafter. Shymkent is not included as a separate analytical unit in the panel; its 2018 territorial transition is addressed separately below (President of the Republic of Kazakhstan, 2018).

The establishment of Abai, Zhetysu, and Ulytau regions on 8 June 2022 required harmonisation of the long-run panel (President of the Republic of Kazakhstan, 2022). Before this reform, the territories subsequently assigned to Ulytau, Zhetysu, and Abai formed part of the Karagandy, Almaty, and East Kazakhstan regions, respectively. Accordingly, the corresponding historical regional series are used for 2003–2021 without introducing separate observations for the newly established regions. For 2022–2025, the constituent regions are aggregated to reconstruct the earlier territorial coverage of Karagandy, Almaty, and East Kazakhstan. Underlying counts are aggregated before labour-market ratios are calculated. These changes are distinct from the 2007 settlement-classification adjustments, which concern urban–rural statistical comparability rather than the creation of new regions.

For the analytical regions reconstructed following the 2022 administrative changes, the harmonised gross regional product (GRP) and agricultural production volume indices

are researcher-constructed approximations based on constituent BNS statistical series and lagged nominal weights. These approximations are not directly published by the BNS as volume indices for the reconstructed analytical territories. Where harmonised indices could not be constructed from the available data, the corresponding values were treated as missing and were not imputed.

The South Kazakhstan/Turkistan analytical unit requires separate treatment because Shymkent was included in the regional series before its separation in 2018 but is not included in the subsequent Turkistan Region series. The resulting change in territorial coverage creates a discontinuity in the 2018 annual comparison, so the 2018 first-difference observation for this analytical unit is excluded from the principal estimation sample (President of the Republic of Kazakhstan, 2018). The unit itself remains in the main panel, while its complete exclusion is examined as a territorial-comparability sensitivity analysis. The annual change dated 2019 is retained because the observations for 2018 and 2019 refer to the post-separation territory of Turkistan Region.

The 2022 annual change is also excluded from the principal estimation sample for all analytical units to separate the administrative-transition year from the harmonised post-reform observations. The 2003 observations are retained as the baseline required to calculate changes dated 2004. First differencing of the 322-observation level panel produces a maximum of 308 annual changes for 2004–2025. Excluding the 14 observations dated 2022 and the 2018 observation for the South Kazakhstan/Turkistan analytical unit leaves 293 observations across 14 analytical regions in the principal estimation sample.

3.2. Variable Construction and Transformations

Regional accommodation-based tourism demand is represented by the number of visitors served by accommodation establishments ($AccomVis$). Because regions differ substantially in the scale of accommodation activity, the econometric analysis uses the annual first difference in the natural logarithm:

$$\Delta \ln AccomVis_{it} = \ln AccomVis_{it} - \ln AccomVis_{i,t-1} \quad (1)$$

where $AccomVis_{it}$ denotes the number of visitors served by accommodation establishments in analytical region i in year t. The transformation captures annual proportional changes in recorded regional accommodation demand and reduces the influence of persistent differences in tourism scale across regions.

Rural employment intensity ($RurEmpInt$) is calculated as the number of employed rural residents per 1000 rural residents:

$$RurEmpInt_{it} = \frac{RurEmp_{it}}{RurPop_{it}} \times 1000 \quad (2)$$

where $RurEmp_{it}$ is the number of employed rural residents and $RurPop_{it}$ is the total rural population in region i in year t. This indicator adjusts the number of employed rural residents for differences in rural-population size across analytical regions. It is not interpreted as a conventional employment rate because its denominator is the total rural population rather than the working-age population or labour force.

For the econometric analysis, rural employment intensity enters as the first difference in its natural logarithm:

$$\Delta \ln RurEmpInt_{it} = \ln RurEmpInt_{it} - \ln RurEmpInt_{i,t-1} \quad (3)$$

This transformation measures the annual proportional change in rural employment intensity. The factor of 1000 used in the definition of rural employment intensity affects the level of the indicator but not its log first difference because this constant cancels when consecutive logarithms are differenced.

The rural unemployment rate (*RurUnempRate*) is calculated as unemployed rural residents relative to the observed rural labour force:

$$RurUnempRate_{it} = \frac{RurUnemp_{it}}{RurEmp_{it} + RurUnemp_{it}} \times 100 \quad (4)$$

where $RurUnemp_{it}$ denotes the number of unemployed rural residents and $RurEmp_{it} + RurUnemp_{it}$ represents the observed rural labour force. The resulting indicator is expressed as a percentage and corresponds to the share of unemployed persons in the employed-plus-unemployed rural population.

Because the unemployment indicator is already expressed as a percentage, the econometric analysis uses its annual change in percentage points rather than a logarithmic transformation:

$$\Delta RurUnempRate_{it} = RurUnempRate_{it} - RurUnempRate_{i,t-1} \quad (5)$$

where $\Delta RurUnempRate_{it}$ denotes the annual percentage-point change in the rural unemployment rate between years $t-1$ and t. Unlike the other principal variables, the unemployment rate is not logarithmically transformed. All variables entering logarithmic transformations elsewhere in the analysis are strictly positive for non-missing observations, so no constant offsets or zero-value substitutions are required.

In addition to the principal tourism-demand and rural labour-market variables, the analysis incorporates two indicators of regional economic conditions. The gross regional product volume index (*GRPVolidx*) is used to represent changes in overall regional economic activity, while the agricultural production volume index (*AgriVolidx*) captures changes in agricultural output. These indicators are included because broader regional economic developments and agricultural conditions may be associated with changes in rural labour-market outcomes independently of accommodation-based tourism demand. For the econometric analysis, both indices are transformed into logarithmic measures relative to an index value of 100:

$$GRPGrowth_{it} = \ln\left(\frac{GRPVolidx_{it}}{100}\right) \quad (6)$$

$$AgriGrowth_{it} = \ln\left(\frac{AgriVolidx_{it}}{100}\right) \quad (7)$$

where $GRPVolidx_{it}$ and $AgriVolidx_{it}$ denote the original regional volume indices for analytical region i in year t. The logarithmic transformation equals zero when the index is 100, while values above and below 100 correspond to positive and negative logarithmic changes, respectively. Unlike the tourism-demand indicator, these variables are calculated as the natural logarithms of annual volume indices divided by 100, using either directly reported BNS values or the harmonised approximations described in Section 3.1. As the indices already measure changes relative to the previous year, their logarithms are not first-differenced again. Their inclusion allows the estimated association between accommodation-demand growth and rural labour-market outcomes to be examined conditional on contemporaneous changes in regional economic and agricultural activity.

Demographic conditions are examined using the share of rural residents aged 15–64 years (*RurAge1564Share*). This variable is calculated by the authors from the Bureau of National Statistics of Kazakhstan (BNS) data on rural population by single year of age. For each analytical region and year, the population aged 15–64 is divided by the total rural population reported in the same age-specific dataset and multiplied by 100. Where necessary, population counts are first aggregated to the harmonised regional boundaries before calculating the share. Both the numerator and denominator of the age-share indicator are taken from the same BNS age-specific rural-population dataset. The population total

used to calculate this share is not substituted for the separately compiled *RurPop* series used in the employment-intensity measure.

The indicator is measured as of 1 January of each year and is available for 2009–2025. It is relevant because the employment-intensity outcome relates the number of employed rural residents to the total rural population rather than exclusively to the working-age population. To assess the sensitivity of the results to changes in demographic composition, we additionally consider the annual first difference in the age-share indicator:

$$\Delta RurAge1564Share_{it} = RurAge1564Share_{it} - RurAge1564Share_{i,t-1} \quad (8)$$

where $RurAge1564Share_{it}$ denotes the authors' calculated share of rural residents aged 15–64 in analytical region i, measured as of 1 January of year t. The transformation captures the change in demographic composition between 1 January of year $t-1$ and 1 January of year t and retains the measurement scale of the original series (percentage points). This indicator is not included in the principal controlled specification because its available observations are more limited. Instead, demographic sensitivity is evaluated separately by comparing models estimated on an identical sample with and without the age-share variable.

The availability of accommodation infrastructure in rural areas is represented by the number of rural accommodation establishments (*RurAccom*). This variable provides a more geographically specific indicator than the regionwide visitor series, although it measures the supply of accommodation establishments rather than the volume of tourism demand in rural destinations. To examine whether changes in rural accommodation supply are associated with rural labour-market conditions, the variable is transformed using the annual first difference in its natural logarithm:

$$\Delta \ln RurAccom_{it} = \ln RurAccom_{it} - \ln RurAccom_{i,t-1} \quad (9)$$

where $RurAccom_{it}$ denotes the number of accommodation establishments located in rural areas of analytical region i in year t. The logarithmic transformation measures proportional annual changes in the number of establishments and facilitates comparison with the corresponding transformation of regionwide accommodation-based tourism demand. The rural accommodation-supply analysis covers changes dated 2014–2021, subject to the availability of the additional model variables and the territorial restrictions described in Section 3.1. Because establishment counts do not reveal visitor volumes, employment, or the participation of local residents, the variable is interpreted as a supplementary supply-side indicator rather than a direct measure of rural tourism activity.

The construction of rural employment intensity in Equation (2) provides the basis for an additional decomposition. Because employment intensity is defined as the ratio of employed rural residents to the total rural population, its logarithmic change can be expressed as the difference between the logarithmic changes in these two underlying counts. The relationship is given by

$$\Delta \ln RurEmpInt_{it} = \Delta \ln RurEmp_{it} - \Delta \ln RurPop_{it} \quad (10)$$

This identity distinguishes changes in the number of employed rural residents from changes in the population denominator. Declining employment intensity may therefore reflect a reduction in rural employment, an increase in rural population relative to employment, or a combination of both. To examine these components, the analysis estimates separate auxiliary regressions for changes in rural employment and rural population using the same explanatory variables and analytical observations. These decompositions assist in interpreting the employment-intensity estimates but do not identify the mechanisms through which tourism demand may influence either component. The definitions, transformations, and data sources for the variables used in the analysis are summarised in Table 1.

Table 1. Variable definitions and transformations.

Variable	Definition	Measurement/Transformation	Data Source
AccomVis	Visitors served by accommodation establishments, persons	$\Delta\ln(AccomVis)$; proxy for regional accommodation-based tourism activity	BNS, accommodation visitor statistics (BNS, n.d.-f)
RurPop	Rural population, persons	Denominator of rural employment intensity	BNS, population statistics (BNS, n.d.-h)
RurEmp	Employed rural residents, persons	Component of rural employment intensity and unemployment rate	Official labour-market statistics (Agency of the Republic of Kazakhstan on Statistics, 2008, 2013; BNS, n.d.-b, n.d.-c, n.d.-d)
RurUnemp	Unemployed rural residents, persons	Component of rural unemployment rate	Official labour-market statistics (Agency of the Republic of Kazakhstan on Statistics, 2008, 2013; BNS, n.d.-b, n.d.-c, n.d.-d)
RurEmpInt	Rural employed persons per 1000 rural residents	$1000 \times RurEmp/RurPop$; modelled as $\Delta\ln(RurEmpInt)$	Authors' calculations from BNS data
RurUnempRate	Rural unemployment rate, %	$100 \times RurUnemp/(RurEmp + RurUnemp)$; modelled as $\Delta RurUnempRate$	Authors' calculations from BNS data
GRPVolidx	Gross regional product volume index (previous year = 100)	$\ln(GRPVolidx/100)$; regional economic control	BNS, GRP volume indices; authors' harmonisation (BNS, n.d.-i)
AgriVolidx	Agricultural production volume index (previous year = 100)	$\ln(AgriVolidx/100)$; agricultural activity control	BNS, agricultural volume indices; authors' harmonisation (BNS, n.d.-e)
RurAge1564Share	Share of rural residents aged 15–64, %	100 × rural population aged 15–64/total rural population; annual change in sensitivity models	BNS, population by single year of age; authors' calculations (BNS, n.d.-g)
RurAccom	Accommodation establishments in rural areas, number	$\Delta\ln(RurAccom)$; rural accommodation-supply sensitivity	BNS, accommodation-establishment statistics; authors' aggregation (BNS, 2014, n.d.-a)

Note: Δ denotes an annual first difference, and ln denotes the natural logarithm. BNS denotes the Bureau of National Statistics of Kazakhstan. Changes in unemployment and age shares are measured in percentage points. Harmonised GRP and agricultural volume indices include authors' approximations for reconstructed analytical regions where applicable. *RurAge1564Share* is measured as of 1 January and is available for 2009–2025. *RurAccom* is compiled for 2013–2021, with annual changes analysed for 2014–2021.

3.3. Econometric Specification

The empirical analysis examines whether annual changes in regionwide accommodation-based tourism demand are associated with annual changes in rural labour-market conditions within the same analytical regions. The specification is estimated in first differences to focus on year-to-year within-region variation and to remove time-invariant regional characteristics that may be correlated with both tourism activity and labour-market conditions (Wooldridge, 2010). The estimation strategy initially considers contemporaneous associations and subsequently examines whether these relationships are sensitive to selected time-varying regional conditions and lagged changes in tourism demand.

Two outcome-specific models are initially estimated. Rural employment intensity is examined using its annual logarithmic change, while rural unemployment is measured through annual percentage-point changes in the unemployment rate. The initial specification for rural employment intensity is expressed as

$$\Delta \ln\left(RurEmpInt_{it}\right) = \alpha + \beta \Delta \ln(AccomVis_{it}) + \lambda_t + \varepsilon_{it} \quad (11)$$

where $\Delta \ln\left(RurEmpInt_{it}\right)$ is the annual log change in rural employment intensity in analytical region i in year t, and $\Delta \ln(AccomVis_{it})$ is the annual log change in visitors served by accommodation establishments. The coefficient β captures the contemporaneous within-region association between changes in accommodation-based tourism demand and changes in rural employment intensity, conditional on common year effects. Because both variables are expressed as logarithmic changes, the coefficient has an elasticity-type interpretation. Positive values of this coefficient indicate an association between stronger regional accommodation-demand growth and higher growth in rural employment intensity, whereas negative values indicate an association in the opposite direction.

Rural unemployment is estimated separately using the same explanatory variable and annual effects. Unlike employment intensity, the unemployment indicator is expressed as a percentage and enters the regression as an annual change in percentage points. The corresponding specification is

$$\Delta RurUnempRate_{it} = \alpha + \beta \Delta \ln(AccomVis_{it}) + \lambda_t + \varepsilon_{it} \quad (12)$$

where $\Delta RurUnempRate_{it}$ denotes the annual percentage-point change in rural unemployment in analytical region i in year t. In this specification, β represents the percentage-point change in rural unemployment associated with a one-unit change in the logarithmic growth of accommodation-establishment visitors. The coefficient therefore has a different measurement scale from the coefficient in Equation (11). Negative values of the unemployment coefficient indicate an association between stronger accommodation-demand growth and a reduction in rural unemployment, although the statistical significance and precision of the estimated relationship must be evaluated separately.

Year fixed effects, represented by λ_t, are included in both models to account for common annual shocks affecting all analytical regions. These effects absorb nationwide developments that vary over time but are shared across regions, including macroeconomic conditions and other common disturbances. Separate regional intercepts are not included in the principal first-difference regressions because time-invariant regional components are removed through differencing. Nevertheless, this transformation does not eliminate regional influences that change over time, making it important to consider additional time-varying explanatory variables.

The initial specifications are therefore extended by incorporating measures of regional economic and agricultural activity. Changes in gross regional product and agricultural production may be associated with rural labour-market conditions independently of tourism-demand fluctuations, particularly in regions where agricultural activity remains an important component of the rural economy. The extended specification includes the logarithmic

transformations of the gross regional product and agricultural production volume indices defined in Equations (6) and (7). For both rural labour-market outcomes, the controlled model is expressed as

$$Y_{it} = \alpha + \beta\Delta\ln(AccomVis_{it}) + \gamma_1 GRPGrowth_{it} + \gamma_2 AgriGrowth_{it} + \lambda_t + \varepsilon_{it} \quad (13)$$

where Y_{it} denotes either the annual logarithmic change in rural employment intensity, $\Delta\ln(RurEmpInt_{it})$, or the annual percentage-point change in rural unemployment, $\Delta RurUnempRate_{it}$, depending on the dependent variable. *GRPGrowth* and *AgriGrowth* represent the logarithmic transformations of the corresponding regional volume indices relative to 100, while γ_1 and γ_2 denote their estimated coefficients. The coefficient β measures the contemporaneous association between accommodation-based tourism demand and the rural labour-market outcome after conditioning on these two observed regional factors and common year effects. The inclusion of these variables allows the analysis to assess whether the initially estimated association is sensitive to concurrent changes in regional economic and agricultural conditions. However, the controlled model does not eliminate all potential time-varying confounding or establish that the estimated tourism-demand coefficient represents a causal effect.

The initial specifications use 293 annual observations, while the controlled models use 290 observations with complete economic and agricultural indices. To distinguish the effects of sample composition from those of introducing additional controls, the initial model is also re-estimated on the common 290-observation sample. To distinguish changes in rural employment from changes in the population denominator, the controlled specification is also estimated separately for the logarithmic changes in employed rural residents and rural population, using the same analytical sample and the decomposition specified in Equation (10).

The relationship between tourism demand and rural labour-market conditions may also involve temporal adjustments. Changes in visitor activity may be accompanied by investment, enterprise development, or changes in labour demand that do not occur entirely within the same year. To examine whether the contemporaneous specification overlooks delayed associations, the controlled model is extended to include one-year and two-year lags of accommodation-demand growth. The distributed-lag specification is written as

$$Y_{it} = \alpha + \sum_{k=0}^{K} \beta_k \Delta\ln(AccomVis_{i,t-k}) + \gamma_1 GRPGrowth_{it} + \gamma_2 AgriGrowth_{it} + \lambda_t + \varepsilon_{it} \quad (14)$$

where β_0 represents the contemporaneous association between accommodation-demand growth and the rural labour-market outcome, while β_1 and β_2 represent the associations with tourism-demand changes one and two years earlier, respectively. The maximum lag length K is set to one or two years. The one-lag specification includes the contemporaneous and first-lag terms, whereas the two-lag specification additionally includes the second lag. Economic and agricultural controls and year fixed effects are retained in both specifications. The lagged variables are constructed from the annual log changes in accommodation-establishment visitors rather than from the logarithmic levels of visitor counts.

The one-lag model is estimated using 261 observations, while the two-lag model contains 232 observations. These samples are smaller than the principal annual sample because lagged estimation requires consecutive admissible annual changes for the relevant periods. Observations for which the required lag sequence crosses a predefined territorial discontinuity or excluded annual comparison are not used. The resulting specifications therefore preserve the territorial-comparability restrictions while allowing the association with tourism-demand growth to be examined over a longer temporal horizon.

To distinguish changes associated with lag specification from those arising from differences in sample composition, the contemporaneous controlled model and the one-lag model

are additionally re-estimated on the same 232 observations used in the two-lag model. This common-sample comparison holds the region-year observations constant while varying the included tourism-demand lag terms. The results are reported in Supplementary Table S4.

The distributed-lag analysis considers both the individual coefficients and their combined association over the included periods. The sum of the contemporaneous and lagged tourism-demand coefficients provides a summary of their estimated cumulative association within the specified lag window. Joint hypothesis tests are also conducted to examine whether all included tourism-demand coefficients are simultaneously equal to zero. These tests complement the individual coefficient estimates because a lack of statistical significance for a particular lag does not necessarily establish the absence of an association across the included periods. The lagged models remain observational specifications and do not identify the causal timing of tourism-related labour-market adjustments.

For interpretation, the specifications are organised according to their analytical purpose rather than the statistical significance of their estimates. The main annual analysis comprises M1, the initial contemporaneous model; M2, its matched-sample counterpart; and M3, the model incorporating regional economic and agricultural controls. M3 serves as the principal adjusted benchmark, while M1 and M2 distinguish the effects of adding controls from those of changing the estimation sample. The one- and two-year distributed-lag models (L1 and L2) are secondary, substantively motivated extensions examining delayed annual associations. Temporal, territorial, regional-trend, demographic, and rural accommodation-supply checks assess sensitivity to selected assumptions and additional variables. The restricted-demand and overlapping three-year specifications, together with retrospective reclassification exclusions and additional inference diagnostics, are exploratory. Their coefficients may describe different conditional associations or measurement horizons and are not interchangeable with the principal annual estimates. Statistical significance is not used to select a preferred specification, and exploratory p-values are not adjusted for multiple testing.

The coefficient in the rural employment-intensity model has an elasticity-type interpretation because accommodation demand and employment intensity enter as first differences in natural logarithms. For small proportional changes, β approximately represents the percentage change in employment intensity associated with a 1% change in regional accommodation demand. The unemployment coefficient has a different interpretation because the dependent variable is measured in percentage points. For a proportional change in accommodation demand of size g, the corresponding association with the rural unemployment rate is calculated as

$$\beta \ln(1 + g) \tag{15}$$

where g denotes the proportional change in accommodation demand expressed as a decimal. For example, a 10% increase in accommodation-establishment visitors corresponds to $g = 0.10$, so the estimated percentage-point association is obtained by multiplying the unemployment coefficient by $\ln(1.10)$. This conversion preserves the distinction between changes in the logarithm of accommodation demand and changes in the unemployment rate measured in percentage points. For rural employment intensity, the corresponding exact percentage change is obtained as

$$100[\exp(\beta \ln(1 + g)) - 1] \tag{16}$$

3.4. Diagnostic and Robustness Checks

Before estimation, the uniqueness of region-year observations, panel balance, validity of logarithmic inputs, and positivity of rural labour-force denominators are verified. The source-data diagnostics also identify one unusually high rural employment-intensity observation

for the historical Karagandy Region in 2007 (870.49 employed persons per 1000 rural residents). The rural employment count is consistent with the official 2003–2007 labour-market collection (Agency of the Republic of Kazakhstan on Statistics, 2008). The rural-population series refers to the beginning of the calendar year, whereas rural employment counts are reported for the corresponding reference year. Around the 2007 reclassification of urban and rural settlements, these different temporal reference points may affect the comparability of the numerator and denominator of employment intensity, even when both underlying values are correctly reported. The observation is therefore retained in the principal estimation sample without correction or automatic exclusion. Retrospective sensitivity checks examine the possible influence of the 2007–2008 settlement-status transition by excluding the corresponding annual changes for the historical Karagandy Region and, separately, for all analytical regions (Supplementary Table S3). These checks assess sensitivity to a potential statistical discontinuity but do not establish measurement error.

Inference is based on standard errors clustered at the analytical-region level. Cluster-robust inference allows regression disturbances to be heteroskedastic and correlated within regions over time, which is important in longitudinal regional data where repeated observations from the same territorial unit are unlikely to be statistically independent (Cameron & Miller, 2015). Because the principal models contain only 14 regional clusters, conventional large-sample cluster-robust inference may provide unreliable finite-sample approximations when used alone (Cameron et al., 2008).

Wild-cluster bootstrap inference is therefore used as an additional small-cluster procedure. Wild-cluster bootstrap methods are designed to improve statistical inference when regression disturbances are clustered and the number of clusters is limited (Cameron et al., 2008). The procedure generates a reference distribution for hypothesis testing by applying random weights at the cluster level, thereby accounting for the dependence among observations within each analytical region. Further research shows that the performance of cluster-robust and wild-bootstrap procedures can depend on the number and relative size of the clusters, reinforcing the need for caution when inference is based on a small number of groups (MacKinnon & Webb, 2017). Accordingly, the wild-cluster bootstrap is used here as a finite-sample inference check rather than as a remedy that eliminates the limitations associated with having only 14 regional clusters.

Wild-cluster bootstrap p-values are calculated using 9999 replications with clustering at the analytical-region level. The principal specifications retaining 14 analytical regions use Rademacher weights, whereas the 13-cluster specification excluding South Kazakhstan/Turkistan uses Webb weights. Rademacher weights are drawn from a two-point distribution, while Webb weights are drawn from a six-point distribution. The choice of weights may affect bootstrap-based statistical inference but does not change the estimated regression coefficients or the estimation sample. Supplementary Table S3 additionally examines the sensitivity of selected specifications to the use of alternative bootstrap weights.

Finite-cluster t-based inference is also reported for the estimated coefficients, using a t-reference distribution with $G - 1$ degrees of freedom, where G denotes the number of regional clusters. For selected coefficients, 95% confidence intervals are obtained by inverting the wild-cluster bootstrap hypothesis tests rather than by applying conventional t-critical values to the clustered standard errors. These intervals provide an additional assessment of the range of coefficient values compatible with the data under the bootstrap procedure. Wild-cluster bootstrap p-values are reported alongside the cluster-t p-values, while the lagged specifications are additionally evaluated using joint tests of the included tourism-demand coefficients. Statistical interpretation therefore considers the coefficient estimates, their clustered standard errors, and the available bootstrap evidence rather than relying on conventional large-sample significance tests alone.

Temporal sensitivity is assessed by excluding annual changes dated 2020–2021 to account for pandemic-related disruptions ($N = 265$) and by restricting the estimation period to changes through 2021 to exclude the harmonised post-2022 observations ($N = 251$ in the initial model and $N = 248$ in the controlled model). Both restrictions are examined using the initial annual specification, with additional matched-sample comparisons of models with and without regional economic and agricultural controls reported in Supplementary Table S4.

Territorial robustness is assessed by excluding the South Kazakhstan/Turkistan analytical unit entirely. This specification addresses the possibility that the 2018 separation of Shymkent creates comparability concerns extending beyond the single first-difference observation removed from the principal sample. Unlike the temporal restrictions, this check is conducted using the specification that includes the regional economic and agricultural controls. The resulting model contains 270 observations across 13 analytical regions, compared with 290 observations across 14 regions in the corresponding controlled specification. The same explanatory variables, year fixed effects, and region-clustered inference procedures are retained to assess whether the estimated association is sensitive to the inclusion of this territorially affected analytical unit.

Differences in underlying regional trends are examined in an additional specification. Although first differencing removes time-invariant regional components, it does not eliminate persistent changes in regional characteristics that may evolve differently across analytical regions. To examine sensitivity to such variation, analytical-region indicators are added to the first-difference regression alongside the economic and agricultural controls and year fixed effects. In this formulation, the regional indicators correspond to region-specific linear trends in the underlying level relationships rather than conventional region fixed effects in the original level specification. The model is estimated using the same 290 observations as the principal controlled specification, permitting the estimated accommodation-demand coefficient to be compared without changes in sample composition.

Demographic sensitivity is assessed by incorporating the annual change in the share of rural residents aged 15–64. This variable is relevant because rural employment intensity is defined relative to the total rural population, and changes in the demographic composition of that population may be associated with changes in the prevalence of employment. The demographic indicator is available for a smaller analytical sample, containing 206 observations across 14 regions. To distinguish the contribution of the additional variable from changes caused by sample availability, the controlled specification is first re-estimated on these 206 observations without the demographic indicator. The annual change in the age-share variable is then introduced using the same observations, allowing the accommodation-demand coefficients to be compared directly.

As a geographically more specific supply indicator, the annual logarithmic change in rural accommodation establishments is included in a separate sensitivity analysis for 2014–2021. Establishment counts do not measure rural visitor demand, tourism employment, or rural residents' participation in tourism-related activities and are therefore not substitutes for the principal tourism-demand measure.

The rural accommodation-supply sample contains 108 observations across 14 analytical regions. Of the 112 potential region-year observations for 2014–2021, the 2018 observation for the South Kazakhstan/Turkistan analytical unit is excluded because of the territorial discontinuity, while three observations dated 2021 are excluded because the required model inputs are incomplete. The 2019 observation for South Kazakhstan/Turkistan is retained because it refers to the post-separation territorial basis. The controlled model is first estimated without the rural accommodation-supply indicator and is subsequently extended by adding its annual logarithmic change. Both specifications use the same 108 observations,

together with year fixed effects and the regional economic and agricultural controls, so that the comparison is not affected by differences in the estimation sample.

Retrospective sensitivity checks examine potential discontinuities in rural labour-market and population statistics around the 2007 settlement-classification changes (Republic of Kazakhstan, 2006; Zhumasultanov, 2011). The first diagnostic excludes annual changes dated 2007 and 2008 for the historical Karagandy Region (RegionID 7), while the second applies the same exclusions to all 14 analytical regions. Three-year specifications omit windows containing the excluded changes, and stricter distributed-lag diagnostics also exclude observations whose lag windows involve the affected years. These checks assess sensitivity to possible statistical discontinuities without attributing observed changes to settlement reclassification. The original data, principal estimation samples, and baseline coefficients remain unchanged.

3.5. Additional Exploratory Specifications

Two additional exploratory analyses are conducted to examine sensitivity to the distribution of annual accommodation-demand changes and to the measurement horizon. The first retains observations satisfying $|\Delta \ln(AccomVis)| \leq \ln(1.5)$, corresponding to ordinary annual changes between approximately −33.3% and +50%. This cutoff is used as an exploratory sensitivity threshold rather than as a statistically identified breakpoint or an established data-cleaning rule. It excludes 45 of the original 293 annual observations, leaving 248 observations without controls and 245 after incorporating regional economic and agricultural controls. Because the restriction conditions on the observed explanatory variable and materially changes the estimation sample, these coefficients describe associations within a restricted range of tourism-demand changes and should not be treated as directly interchangeable with the principal annual estimates.

In a further diagnostic, the sample is restricted to observations with ordinary annual changes in rural population within ±10%, leaving 234 observations. Since rural population forms the denominator of employment intensity, this additional restriction is not an independent robustness test and cannot establish a demographic transmission mechanism. Further exploratory diagnostics assess the dependence of the trimmed associations on the selected cutoff. The controlled annual specifications are re-estimated using alternative absolute log-change thresholds of $\ln(1.25)$ and $\ln(2)$, in addition to the original $\ln(1.5)$ restriction. The original restricted specification is also re-estimated after excluding annual changes dated 2020–2021. These checks are conducted to evaluate sensitivity to sample selection rather than to select a preferred specification on the basis of statistical significance. The corresponding results and the distribution of excluded observations across years and regions are reported in Supplementary Table S4.

The second analysis examines overlapping three-year logarithmic changes in accommodation-based tourism demand and rural employment intensity. The corresponding economic and agricultural controls are accumulated over the same three-year periods, while an additional specification includes analytical-region indicators to account for differences in underlying regional trends. The resulting sample contains 232 observations across 14 analytical regions, with the predefined territorial exclusions retained. The three-year specifications therefore address a different measurement horizon rather than independently replicating the contemporaneous annual estimates. Their overlapping windows, differences in eligible years, and sensitivity to the inclusion of region indicators limit direct comparisons with M3 and the annual distributed-lag models. These specifications are treated as exploratory, and their wild-cluster bootstrap p-values have not been adjusted for multiple comparisons. The detailed results are reported in Supplementary Tables S1 and S2.

4. Results

4.1. Descriptive Statistics and Variable Dynamics

Table 2 reports descriptive statistics for the 322 region-year observations in levels and the 293 admissible annual first differences. The controlled specifications use 290 observations with complete regional economic and agricultural indices. Regional accommodation demand varies substantially, with visitor counts ranging from 7838 to 1,040,088 per region–year. Rural employment intensity averages 514.05 employed persons per 1000 rural residents, while the rural unemployment rate averages 5.57%. These pooled statistics should not be interpreted as population-weighted national labour-market indicators.

Table 2. Descriptive statistics of the principal tourism-demand and rural labour-market variables.

Variable	Obs.	Mean	Std. Dev.	Min
Panel A. Level variables				
AccomVis	322	179,230.33	162,999.04	7838
RurPop	322	513,969.52	401,767.40	77,938
RurEmp	322	256,150.15	183,548.98	26,500
RurUnemp	322	14,501.47	10,256.36	2735
RurEmpInt	322	514.047	73.827	302.296
RurUnempRate	322	5.569	1.732	2.220
Panel B. Variables in the principal first-difference sample				
$\Delta\ln(AccomVis)$	293	0.1012	0.2824	−1.0834
$\Delta\ln(RurEmpInt)$	293	0.0005	0.0722	−0.3706
$\Delta RurUnempRate$	293	−0.1619	0.6092	−3.6873

Note: *RurEmpInt* denotes employed rural residents per 1000 rural residents, while *RurUnempRate* denotes unemployed rural residents as a percentage of employed plus unemployed rural residents. Δln denotes the first difference in the natural logarithm, and Δ*RurUnempRate* is expressed in percentage points. Panel A is based on the complete 322-observation panel for 2003–2025. Panel B is based on the 293-observation principal estimation sample. The additional economic, agricultural, demographic, and rural accommodation-supply variables used in the extended specifications are defined in Table 1 and have differing analytical sample coverage. Source: Authors' calculations from BNS data.

Annual accommodation-demand growth exhibits greater variation (mean = 0.1012; SD = 0.2824) than changes in rural employment intensity (mean = 0.0005; SD = 0.0722). Rural unemployment declines by an average of 0.162 percentage points per region-year, although annual changes vary considerably, as shown in Table 2.

4.2. Results of the Main Regression Models

The results of the main regression models are presented in Table 3. The analysis initially examines contemporaneous associations between annual changes in regionwide accommodation-based tourism demand and the two rural labour-market outcomes. The initial models contain 293 observations across 14 analytical regions and include year fixed effects, with standard errors clustered at the analytical-region level. To assess the sensitivity of the estimates to additional regional conditions, the models are subsequently re-estimated on a common 290-observation sample before and after incorporating the economic and agricultural controls. This sequence distinguishes changes associated with sample composition from those arising from the inclusion of additional explanatory variables.

In the initial model (M1), accommodation-demand growth is not significantly associated with rural employment intensity ($\beta = 0.00131$, wild-bootstrap $p = 0.9644$; 95% CI $[-0.05027, 0.05193]$). For a 10% increase in accommodation visitors, the confidence interval corresponds to employment-intensity changes of approximately −0.478% to 0.496%. This

range indicates considerable statistical uncertainty rather than evidence that the relationship is necessarily absent.

Table 3. Results of the main regression models for rural labour-market outcomes.

Statistic	Rural Employment Intensity	Rural Unemployment Rate
Panel A. Initial annual models (M1)		
$\Delta\ln(AccomVis)$, coefficient	0.00131	−0.22329
Clustered standard error	0.02203	0.16352
Cluster-t *p*-value	0.9535	0.1953
Wild-cluster bootstrap *p*-value	0.9644	0.2315
95% wild-bootstrap CI	[−0.05027, 0.05193]	[−0.54830, 0.22186]
Observations	293	293
R^2	0.2965	0.2062
Panel B. Matched-sample models (M2)		
$\Delta\ln(AccomVis)$, coefficient	0.00138	−0.22451
Clustered standard error	0.02176	0.16342
Cluster-t *p*-value	0.9503	0.1927
Wild-cluster bootstrap *p*-value	0.9617	0.2298
Observations	290	290
R^2	0.2973	0.2060
Panel C. Controlled models (M3)		
$\Delta\ln(AccomVis)$, coefficient	−0.00106	−0.16529
Clustered standard error	0.02179	0.15957
Cluster-t *p*-value	0.9619	0.3192
Wild-cluster bootstrap *p*-value	0.9689	0.3391
95% wild-bootstrap CI	[−0.04940, 0.04915]	[−0.47189, 0.27374]
GRPGrowth, coefficient	0.08921	−2.15212
GRPGrowth, clustered SE	0.09389	1.29790
GRPGrowth, wild-bootstrap *p*-value	0.3686	0.1111
AgriGrowth, coefficient	−0.01075	0.59200
AgriGrowth, clustered SE	0.04726	0.32791
AgriGrowth, wild-bootstrap *p*-value	0.8772	0.0660
Observations	290	290
R^2	0.2991	0.2232

Note: The dependent variables are $\Delta\ln(RurEmpInt)$ for rural employment intensity and $\Delta RurUnempRate$ (percentage points) for rural unemployment. All models include year fixed effects and standard errors clustered by analytical region (14 clusters). Cluster-t *p*-values use $G - 1$ degrees of freedom, while wild-cluster bootstrap *p*-values are based on 9999 replications with Rademacher weights. The reported 95% confidence intervals for accommodation-demand coefficients in M1 and M3 are obtained by inverting wild-cluster bootstrap tests; corresponding intervals were not calculated for M2. M1 uses 293 observations, whereas M2 and M3 use the same 290 observations. *GRPGrowth* and *AgriGrowth* are included only in M3. R^2 refers to the complete regression model, including year fixed effects and applicable controls, rather than to the explanatory contribution of accommodation-demand growth alone.

In the initial rural unemployment model, the accommodation-demand coefficient is negative but statistically insignificant ($\beta = -0.22329$, wild-bootstrap $p = 0.2315$; 95% CI [−0.54830, 0.22186]). A 10% increase in accommodation visitors corresponds to an estimated decline of approximately 0.0213 percentage points in rural unemployment, but the confidence interval includes associations of either sign. The estimate therefore does not establish a systematic unemployment-reducing relationship.

Restricting the initial models to the common 290-observation sample produces coefficients of 0.00138 for rural employment intensity and −0.22451 for rural unemployment,

both statistically insignificant under wild-cluster bootstrap inference. Their similarity to the full-sample estimates indicates that excluding the three observations with incomplete control data has little influence on the initial associations (Table 3).

After introducing regional economic and agricultural controls (M3), the accommodation-demand coefficient is −0.00106 for rural employment intensity (wild-bootstrap $p = 0.9689$; 95% CI [−0.04940, 0.04915]) and −0.16529 for rural unemployment ($p = 0.3391$; 95% CI [−0.47189, 0.27374]). Both estimates remain statistically indistinguishable from zero and compatible with associations of either sign. The small sign reversal in the employment-intensity coefficient does not establish a negative relationship.

Neither regional economic nor agricultural growth is statistically distinguishable from zero at the 5% level under wild-cluster bootstrap inference in either outcome model (Table 3). The agricultural coefficient in the rural unemployment model approaches this threshold ($p = 0.0660$), but the estimate remains statistically uncertain and should not be interpreted causally.

The construction of rural employment intensity also requires attention to changes in its underlying components. As shown in Equation (10), the logarithmic change in employment intensity equals the change in the number of employed rural residents minus the change in the rural population. Auxiliary regressions using the same 290-observation controlled sample yield an accommodation-demand coefficient of 0.01401 for rural employment and 0.01507 for rural population, with wild-cluster bootstrap p-values of 0.5512 and 0.1250, respectively. Neither component coefficient is statistically distinguishable from zero at the 5% level. Their difference equals the employment-intensity coefficient of −0.00106, confirming that the slightly negative controlled estimate cannot be interpreted independently of the population denominator.

4.3. Lagged Associations Between Tourism Demand and Rural Labour-Market Outcomes

Table 4 presents the controlled distributed-lag models with one-year and two-year lags of accommodation-demand growth, estimated using 261 and 232 observations, respectively. For rural employment intensity, the one-lag model produces a contemporaneous coefficient of −0.00305 and a first-lag coefficient of −0.03825, with wild-cluster bootstrap p-values of 0.9107 and 0.1025, respectively. The sum of the two coefficients is −0.04130, with a bootstrap p-value of 0.1424. The two-lag model yields negative coefficients for contemporaneous and lagged accommodation-demand growth, with a cumulative estimate of −0.07647 ($p = 0.0784$). Although the first-lag coefficient and the cumulative estimate approach the 10% significance threshold in the two-lag specification, the joint bootstrap test does not reject the hypothesis that all tourism-demand coefficients are zero ($p = 0.4443$). These results therefore do not establish a statistically distinguishable lagged association at the 5% level.

The rural unemployment models display a different coefficient pattern. In the one-lag specification, the contemporaneous coefficient is −0.15504, while the first-lag coefficient is 0.08393. Neither estimate is statistically distinguishable from zero under wild-cluster bootstrap inference. In the two-lag model, the contemporaneous coefficient becomes more negative at −0.32657. Although its cluster-t p-value is 0.0392, the wild-cluster bootstrap p-value is 0.0969, indicating that statistical significance is sensitive to the inferential procedure. The cumulative coefficient is −0.23865 (wild-bootstrap $p = 0.1931$), and the joint wild-bootstrap p-value is 0.2981. Together with the employment-intensity results, these findings indicate that the joint null hypotheses for contemporaneous and lagged tourism-demand coefficients cannot be rejected at the 5% level, although the estimates vary with lag structure and sample composition.

Table 4. Distributed-lag associations between accommodation-based tourism demand and rural labour-market outcomes.

Statistic	Employment Intensity L1	Employment Intensity L2	Rural Unemployment L1	Rural Unemployment L2
Contemporaneous $\Delta\ln(AccomVis)$	−0.00305	−0.00530	−0.15504	−0.32657
Clustered SE	0.02183	0.02408	0.15529	0.14245
Wild-bootstrap *p*-value	0.9107	0.8454	0.3611	0.0969
First lag	−0.03825	−0.03923	0.08393	0.04808
Clustered SE	0.02065	0.02060	0.11445	0.07921
Wild-bootstrap *p*-value	0.1025	0.0872	0.4881	0.5119
Second lag	—	−0.03194	—	0.03984
Clustered SE	—	0.02679	—	0.15789
Wild-bootstrap *p*-value	—	0.2905	—	0.8041
Sum of tourism coefficients	−0.04130	−0.07647	−0.07111	−0.23865
Clustered SE	0.02528	0.03720	0.20839	0.20890
Wild-bootstrap *p*-value	0.1424	0.0784	0.7393	0.1931
Joint wild-bootstrap *p*-value	0.2865	0.4443	0.5647	0.2981
Observations	261	232	261	232
Regional clusters	14	14	14	14

Note: The dependent variables are $\Delta\ln(RurEmpInt)$ and $\Delta RurUnempRate$, respectively. L1 includes contemporaneous and one-year lagged accommodation-demand growth, while L2 additionally includes a two-year lag. All specifications control for regional economic and agricultural growth and include year fixed effects. Standard errors are clustered by analytical region (14 clusters). Individual wild-bootstrap *p*-values are reported below the corresponding coefficients, using 9999 replications. The sum of tourism coefficients represents the cumulative contemporaneous and lagged association. The joint wild-bootstrap *p*-value tests the null hypothesis that all included tourism-demand coefficients, including the contemporaneous coefficient, are jointly equal to zero. Rural unemployment changes are measured in percentage points. The estimates represent statistical associations rather than causal effects.

Common-sample comparisons ($N = 232$; Supplementary Table S4) leave the contemporaneous employment-intensity coefficients close to zero and statistically insignificant under wild-bootstrap inference. For rural unemployment, however, the controlled contemporaneous coefficient changes from −0.16529 in the full sample to −0.33058 in the common lag sample, indicating sensitivity to sample composition. The latter estimate is significant under cluster-t inference ($p = 0.0419$) but not under wild-bootstrap inference ($p = 0.1011$). The joint wild-bootstrap tests in the lagged models remain insignificant at the 5% level.

4.4. Robustness and Sensitivity Checks

Table 5 summarises targeted temporal, territorial, regional-trend, demographic, and rural accommodation-supply sensitivity checks. Excluding pandemic-period observations, restricting annual changes through 2021, or omitting South Kazakhstan/Turkistan does not produce statistically significant accommodation-demand associations for either outcome under wild-cluster bootstrap inference. Employment-intensity coefficients remain close to zero, while the rural unemployment coefficients remain negative but statistically uncertain. Additional controlled temporal comparisons and retrospective settlement-reclassification diagnostics are provided in Supplementary Tables S4 and S3, respectively.

The controlled M3 models also yield statistically insignificant contemporaneous accommodation-demand coefficients under both temporal restrictions ($N = 265$ and $N = 248$; Supplementary Table S4). Allowing for region-specific trends on the original 290-observation controlled sample likewise produces insignificant coefficients for employment intensity ($\beta = 0.00107$, wild-bootstrap $p = 0.9728$) and unemployment ($\beta = -0.16923$,

$p = 0.4308$). These checks address selected sources of temporal and regional heterogeneity but do not eliminate reverse causality or unobserved time-varying confounding.

Table 5. Robustness and sensitivity checks for the association between accommodation-based tourism demand and rural labour-market outcomes.

Specification	*N*	*G*	Employment Intensity β (SE)	Bootstrap *p*-Value	Rural Unemployment β (SE)	Bootstrap *p*-Value
Panel A. Temporal and territorial robustness						
Excluding 2020–2021	265	14	−0.00222 (0.02176)	0.9249	−0.27657 (0.20935)	0.2286
Through 2021	251	14	0.00557 (0.02372)	0.8615	−0.25798 (0.18868)	0.2395
Excluding South Kazakhstan/Turkistan	270	13	0.00001 (0.02246)	0.9996	−0.16790 (0.16437)	0.3499
Region-specific trends	290	14	0.00107 (0.02398)	0.9728	−0.16923 (0.20379)	0.4308
Panel B. Demographic sensitivity						
Matched sample, without demographic control	206	14	−0.01056 (0.01844)	0.5965	0.00314 (0.09037)	0.9708
Including demographic control	206	14	−0.01046 (0.01855)	0.6036	0.00081 (0.09393)	0.9927
Panel C. Rural accommodation-supply sensitivity						
Matched sample, without rural supply indicator	108	14	−0.00447 (0.02262)	0.8420	0.00976 (0.09150)	0.9151
Including rural supply indicator	108	14	−0.00369 (0.02327)	0.8827	0.01836 (0.09602)	0.8417
$\Delta\ln(RurAccom)$ coefficient in augmented model	108	14	0.02973 (0.04706)	0.5584	0.32912 (0.21237)	0.1127

Note: Except for the final row, the reported coefficient is for $\Delta\ln(AccomVis)$. The dependent variables are $\Delta\ln(RurEmpInt)$ and $\Delta RurUnempRate$, with the latter expressed in percentage points. Parentheses contain standard errors clustered by analytical region. All specifications include year fixed effects. The temporal restrictions use the initial annual model, whereas the territorial exclusion, region-specific-trend, demographic, and rural accommodation-supply specifications include economic and agricultural controls. The specification excluding South Kazakhstan/Turkistan uses 13 clusters and Webb weights; all other specifications use 14 clusters and Rademacher weights. Wild-cluster bootstrap *p*-values are based on 9999 replications.

The additional variable-based checks examine whether the accommodation-demand coefficients change after accounting for demographic composition or rural accommodation supply. On the matched 206-observation sample, including changes in the share of rural residents aged 15–64 leaves the coefficients essentially unchanged and statistically insignificant for both outcomes. Similarly, adding changes in rural accommodation-establishment counts on the matched 108-observation sample does not produce significant accommodation-demand associations. The rural supply coefficients are also insignificant for employment intensity (wild-bootstrap $p = 0.5584$) and unemployment ($p = 0.1127$). Establishment counts represent accommodation supply rather than tourism demand or employment among rural residents.

The targeted checks and retrospective exclusions around the 2007–2008 settlement reclassification leave the principal contemporaneous annual coefficients statistically insignificant at the 5% level. In particular, excluding the affected annual changes for Karagandy–Ulytau alone or for all analytical regions does not change the non-rejection result in the controlled M3 specifications (Supplementary Table S3). This conclusion applies to the principal annual models and should not be generalised to all exploratory specifications, some of which yield statistically significant associations.

The retrospective Karagandy-only two-lag diagnostics illustrate sensitivity to the treatment of excluded annual changes. When observations dated 2007–2008 are excluded without imposing the stricter lag-window restriction, the cumulative employment-intensity coefficient is −0.08073 ($N = 230$; wild-bootstrap $p = 0.0539$). Under the strict lag-window restriction, which additionally excludes observations whose tourism-demand lags involve the affected years, the cumulative coefficient is −0.08075 ($N = 228$; $p = 0.0473$). The corresponding joint wild-bootstrap test remains insignificant ($p = 0.3489$; Supplementary Table S3). The cumulative contrast and joint test assess different null hypotheses, while the change in cumulative significance across these closely related diagnostic samples reinforces the need for caution in interpreting isolated exploratory results.

4.5. Exploratory Sensitivity Results

The supplementary exploratory results are reported in Tables S1 and S2. Restricting unusually large annual changes in accommodation-based tourism demand produces negative associations with both rural labour-market outcomes. In the controlled restricted sample, the accommodation-demand coefficient is −0.04945 for rural employment intensity and −0.51368 for rural unemployment, with wild-cluster bootstrap p-values of 0.0047 and 0.0330, respectively. Both estimates are statistically distinguishable from zero at the 5% level, in contrast to the principal annual results.

The additional trimming-threshold diagnostics reveal substantial specification sensitivity (Supplementary Table S4). Under the tighter $\ln(1.25)$ restriction, the controlled employment-intensity coefficient is −0.02458 (wild-bootstrap $p = 0.6825$), while the rural unemployment coefficient is −1.09840 ($p = 0.0593$), based on 194 observations. Under the wider $\ln(2)$ restriction, the corresponding coefficients are −0.01469 ($p = 0.4210$) and −0.11242 ($p = 0.3640$), based on 280 observations. Thus, the 5% statistical significance observed under the original $\ln(1.5)$ cutoff is not preserved under either alternative threshold.

The original trimmed coefficients remain similar and statistically significant when 2020–2021 are additionally excluded (Supplementary Table S4). However, this does not establish robustness to the choice of trimming threshold. The $\ln(1.5)$ restriction removes 45 annual observations, including 20 from 2020 to 2021 and 14 from 2006 to 2007 (Supplementary Table S4). This concentration indicates that the trimmed estimates partly reflect the exclusion of major tourism-demand adjustments in particular historical periods.

Restricting annual rural-population changes to ±10% weakens the trimmed coefficients to −0.01655 for employment intensity (wild-bootstrap $p = 0.5119$) and −0.31302 for rural unemployment ($p = 0.1339$). In the original trimmed sample, the employed-count and rural-population coefficients are 0.01245 ($p = 0.8495$) and 0.06190 ($p = 0.0848$), respectively (Supplementary Table S1). Because rural population is the denominator of employment intensity, these diagnostics do not establish a demographic transmission mechanism. Moreover, negative employment-intensity and unemployment coefficients have different labour-market interpretations.

The overlapping three-year specifications yield negative employment-intensity associations in the controlled model ($\beta = -0.04681$, wild-bootstrap $p = 0.0010$) and after adding region indicators ($\beta = -0.04355$, $p = 0.0467$; 95% CI $[-0.11818, -0.00083]$). In the latter model, however, the cluster-t p-value is 0.0564, indicating sensitivity to the inferential procedure. The employed-count and rural-population coefficients are individually insignificant under wild-bootstrap inference ($p = 0.3549$ and 0.3781), which is compatible with a significant coefficient for their difference because its uncertainty also depends on covariance. These associations do not establish that tourism-demand growth reduces the number of employed rural residents.

The retrospective exclusions around 2007–2008 preserve some, but not all, exploratory findings (Supplementary Table S3). The Karagandy-only restriction preserves significance in the selected negative associations. Under the nationwide restriction, however, the trimmed employment-intensity coefficient becomes insignificant ($\beta = -0.01804$, wild-bootstrap $p = 0.1429$), whereas the trimmed unemployment coefficient remains significant ($\beta = -0.49450$, $p = 0.0225$). The three-year employment-intensity association remains significant without region indicators ($p = 0.0055$), but not with them ($p = 0.1239$). These exclusions change both sample composition and temporal coverage and cannot establish whether settlement reclassification caused the observed changes.

In the nationwide-restricted three-year decomposition, the employed-count and rural-population coefficients are individually significant (wild-bootstrap $p = 0.0077$ and 0.0160), while their difference, representing employment intensity, is not ($p = 0.1239$). This is

statistically consistent because the variance of a difference also depends on the covariance between its components.

Overall, the exploratory models do not produce specification-invariant statistically significant associations. In particular, the negative coefficients found under the $\ln(1.5)$ trimming rule fail to reach the 5% significance level under either the $\ln(1.25)$ or $\ln(2)$ threshold (Supplementary Table S4). The trimmed employment-intensity association also loses significance after restricting annual rural-population changes to $\pm 10\%$ (Supplementary Table S1), while selected three-year results vary with the treatment of the 2007–2008 observations and the inclusion of region indicators (Supplementary Table S3). These differences demonstrate that statistical significance depends on sample composition and model specification, rather than providing consistent evidence of a negative association. Given the limited number of clusters, overlapping three-year windows, and multiple unadjusted exploratory comparisons, isolated statistically significant coefficients should be interpreted as hypothesis-generating rather than confirmatory evidence. None of these specifications identifies causal effects or mechanisms of territorial transmission.

5. Discussion

The principal controlled models provide statistically uncertain associations between accommodation-demand growth and both rural labour-market outcomes. For employment intensity, the 95% wild-bootstrap confidence interval is compatible with changes of approximately -0.47% to 0.47% for a 10% increase in accommodation visitors. For rural unemployment, the corresponding interval allows changes ranging from approximately -0.045 to 0.026 percentage points. Thus, the estimates do not establish systematic contemporaneous relationships, but neither do they provide sufficiently precise evidence to exclude associations of potentially meaningful magnitude. The differing coefficient patterns in the additional specifications further limit conclusions about the direction and size of these relationships.

The principal controlled annual and distributed-lag specifications do not provide statistically significant evidence at the 5% level under wild-cluster bootstrap inference, although their confidence intervals do not rule out associations of potentially meaningful magnitude. Selected exploratory trimmed and overlapping three-year models yield negative coefficients, but their statistical significance is not consistently preserved when the trimming threshold, rural-population-change restriction, treatment of the 2007–2008 settlement-status transition, or regional-trend specification is modified. For example, the trimmed controlled employment-intensity coefficient is statistically significant under the $\ln(1.5)$ restriction (wild-bootstrap $p = 0.0047$), but not under $\ln(1.25)$ ($p = 0.6825$) or $\ln(2)$ ($p = 0.4210$). This variation does not invalidate the exploratory estimates, but it prevents interpreting them as evidence of a robust negative relationship between accommodation-demand growth and rural employment intensity. The contrasting results warrant further investigation rather than a single causal or directional interpretation.

Supplementary Table S4, Panel D, shows that 34 of the 45 observations excluded by the exploratory $\ln(1.5)$ restriction fall in 2006–2007 or 2020–2021. The earlier exclusions involve large increases in accommodation visitors, whereas the later ones include sharp decreases in 2020 and predominant increases in 2021. This concentration means that trimming changes not only the range of observed demand changes but also the temporal composition of the estimation sample, potentially contributing to the sensitivity of the exploratory coefficients to alternative thresholds. Although the trimmed coefficients remain similar when the remaining 2020–2021 observations are additionally excluded, that comparison cannot isolate the contribution of the large changes already removed by trimming. These patterns do not establish measurement error or a causal explanation for the estimated associations.

The alternative bootstrap-weight comparisons reported in Supplementary Table S3 do not change the 5% significance classifications of the examined specifications. The leave-one-region-out diagnostics for the principal M3 model (Supplementary Table S4, Panel E) show that the rural unemployment coefficients remain negative across regional exclusions, whereas the near-zero employment-intensity coefficient changes sign in several cases. These exercises assess sensitivity to bootstrap weighting and the influence of individual regions, respectively. They do not establish statistically robust negative associations or eliminate the uncertainty surrounding the principal estimates.

The aggregate results do not contradict destination-level evidence that tourism can generate employment. Tourism-related opportunities depend on seasonality, labour mobility, destination characteristics, and household participation (Kožić & Sever, 2022; Ladkin et al., 2023; T. Li et al., 2021; Luo et al., 2022). Research on spatially heterogeneous tourism effects and integrated tourism value chains likewise explains why local economic gains need not translate into measurable improvements across wider rural labour markets (Faber & Gaubert, 2019; Comerio et al., 2020; Kim et al., 2021; Saxena et al., 2007; Mitchell, 2012).

This study contributes evidence on the precision and specification sensitivity of associations measured across different geographical and sectoral scales. Territorial transmission provides an interpretive framework, but the available indicators do not identify the underlying employment, procurement, entrepreneurship, or commuting mechanisms.

Evidence from Aktobe Oblast illustrates subregional differences in tourism resources, infrastructure, and development potential (Saparov et al., 2024). Rural livelihoods may also combine agriculture, household production, self-employment, seasonal tourism activities, and other income sources. Tourism may therefore increase working hours or household earnings without necessarily changing recorded employment or unemployment status. Although some forms of self-employment are covered by official labour statistics, the aggregate indicators used here cannot identify tourism-related income diversification or changes in work intensity. These possibilities are relevant to interpreting the findings but cannot be tested directly with the available data.

The practical implication is that regional accommodation statistics alone provide an incomplete basis for evaluating tourism-related labour-market development in rural areas. More informative monitoring would combine visitor statistics with the location of tourism establishments and employment, workers' places of residence, local enterprise participation, procurement from rural suppliers, commuting flows, seasonal work, and the retention of visitor expenditure. Such information would make it possible to examine whether increases in accommodation demand are associated with new employment opportunities, additional income among already-employed residents, or changes concentrated in particular destinations. It would also allow the intermediate linkages identified in rural tourism and value-chain research to be assessed more directly.

Several limitations qualify the interpretation of the findings. First differencing removes time-invariant regional characteristics, and year fixed effects account for shocks common to the analytical regions. The addition of regional economic and agricultural controls and the specification allowing for regional trends address selected sources of variation, but they do not eliminate unobserved time-varying confounding, simultaneity, or reverse causality. The coefficients therefore remain conditional associations rather than causal effects. Statistical inference is also constrained by the availability of only 14 analytical regional clusters. Region-clustered standard errors, finite-cluster t-tests, and wild-cluster bootstrap procedures provide a more cautious assessment of uncertainty, but they cannot remove the limitations of the underlying sample structure.

Measurement limitations are equally important. The accommodation-demand indicator includes establishments throughout each analytical region and does not distinguish

visitors to rural destinations from those staying in urban areas. Rural employment intensity covers employed rural residents across all industries and uses the total rural population as its denominator, while the unemployment indicator refers to the wider rural labour force rather than tourism-sector workers. The additional rural accommodation-establishment series provides a more geographically specific supply indicator, but it does not measure rural visitor demand or tourism-related employment. Although one- and two-year lagged associations are examined, annual observations may still conceal seasonal adjustments, shorter-term changes, and longer investment or enterprise-development processes. These differences in geographical coverage, sectoral scope, and temporal measurement limit the conclusions that can be drawn about the transmission of tourism-related opportunities to rural residents.

Future research should combine district- or destination-level tourism data with employer–employee records, enterprise surveys, and household income information to examine the local linkages that cannot be identified here. Comparative studies could also assess whether the observed sensitivity to geographical scale and temporal specification extends beyond Kazakhstan.

6. Conclusions

This study examined associations between regionwide accommodation-based tourism demand and rural labour-market dynamics using a harmonised panel of 14 analytical regions in Kazakhstan over 2003–2025. The principal annual first-difference models do not show statistically significant contemporaneous associations with rural employment intensity or unemployment under wild-cluster bootstrap inference. This conclusion is maintained in the targeted temporal, territorial, controlled, and regional-trend checks. The principal distributed-lag models also fail to reject the joint null hypotheses at the 5% level.

These non-rejections do not establish that the underlying associations are zero or negligible, because the confidence intervals remain compatible with positive and negative associations of potentially meaningful magnitude. Although selected exploratory trimmed and overlapping three-year specifications yield statistically significant negative coefficients, their significance and magnitude depend on sample restrictions, temporal aggregation, and model structure. Consequently, these estimates should be regarded as exploratory patterns requiring further examination rather than robust evidence of a negative relationship. Taken together, the findings establish neither an absence of association nor a stable negative association across specifications.

Because accommodation demand is measured regionwide while labour-market indicators cover rural residents across economic sectors, the analysis does not identify tourism-specific employment effects or mechanisms of territorial transmission. Evaluating tourism-related opportunities in rural areas requires geographically disaggregated evidence on employment, commuting, local enterprise participation, procurement, and other economic linkages.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/tourhosp7110370/s1, Table S1, Exploratory annual associations under restrictions on accommodation-demand changes; Table S2, Exploratory associations based on overlapping three-year changes in accommodation demand and rural employment intensity; Table S3, Retrospective sensitivity to the 2007 settlement-status transition and additional inference diagnostics; and Table S4, Additional common-sample, temporal-control, and trimming-threshold sensitivity diagnostics.

Author Contributions: Conceptualisation, G.A. and Z.D.; methodology, Z.D.; software, Z.D.; validation, G.A., B.S. and Z.D.; formal analysis, G.A.; investigation, G.A.; resources, B.S.; data curation, G.A.; writing—original draft preparation, G.A.; writing—review and editing, B.S. and Z.D.; visualisation,

G.A. and Z.D.; supervision, B.S.; project administration, Z.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original statistical series used in this study are publicly available from the Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan through the data sources cited in Section 3.1 and the reference list. The harmonised analytical panel used for the estimations was constructed by the authors from these publicly available series.

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

References

Adiyia, B., & Vanneste, D. (2018). Local tourism value chain linkages as pro-poor tools for regional development in western Uganda. *Development Southern Africa*, *35*(2), 210–224. [CrossRef]

Agency of the Republic of Kazakhstan on Statistics. (2008). *Economic activity of the population of Kazakhstan (2003–2007)* [Statistical collection]. Available online: https://new.stat.gov.kz/upload/iblock/9e1/ua35hjn0wzncxzrnjia6fde2t66qy247/Экономическая%20активность.pdf (accessed on 27 August 2026).

Agency of the Republic of Kazakhstan on Statistics. (2013). *Economic activity of the population of Kazakhstan (2008–2012)* [Statistical collection]. Available online: https://stat.gov.kz/ru/release-calendar/?date=11.2013 (accessed on 27 August 2026).

Ashley, C., Roe, D., & Goodwin, H. (2001). *Pro-poor tourism strategies: Making tourism work for the poor—A review of experience*. Overseas Development Institute. Available online: https://odi.org/en/publications/pro-poor-tourism-strategies-making-tourism-work-for-the-poor-a-review-of-experience/ (accessed on 27 August 2026).

Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan (BNS). (n.d.-a). *Activities of accommodation establishments in the Republic of Kazakhstan [Annual statistical spreadsheets, 2013–2021]*. Available online: https://stat.gov.kz/industries/business-statistics/stat-tourism/spreadsheets/ (accessed on 16 October 2026).

Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan (BNS). (n.d.-b). *Employment and unemployment: Publications*. Available online: https://stat.gov.kz/en/industries/labor-and-income/stat-empt-unempl/publications/ (accessed on 27 August 2026).

Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan (BNS). (n.d.-c). *Employment and unemployment: Spreadsheets*. Available online: https://stat.gov.kz/en/industries/labor-and-income/stat-empt-unempl/spreadsheets/ (accessed on 27 August 2026).

Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan (BNS). (n.d.-d). *Employment in Kazakhstan: Statistical collection archive*. Available online: https://stat.gov.kz/ru/publication/collections/?name=16818&period=&year= (accessed on 27 August 2026).

Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan (BNS). (n.d.-e). *Indexes of physical volume of gross agricultural production by region*. Available online: https://stat.gov.kz/en/industries/business-statistics/stat-forrest-village-hunt-fish/dynamic-tables/ (accessed on 16 October 2026).

Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan (BNS). (n.d.-f). *Number of served visitors in accommodation establishments*. Available online: https://stat.gov.kz/api/iblock/element/57131/file/en/ (accessed on 26 August 2026).

Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan (BNS). (n.d.-g). *Population by single year of age (by years)*. Available online: https://stat.gov.kz/en/industries/social-statistics/demography/dynamic-tables/ (accessed on 16 October 2026).

Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan (BNS). (n.d.-h). *Population (by years)*. Available online: https://stat.gov.kz/api/iblock/element/38497/file/en/ (accessed on 26 August 2026).

Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan (BNS). (n.d.-i). *Volume index of gross regional product*. Available online: https://stat.gov.kz/en/industries/economy/national-accounts/dynamic-tables/ (accessed on 16 October 2026).

Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan (BNS). (2014, October 28). *Activities of accommodation establishments in the Republic of Kazakhstan (January–December 2013)*. Available online: https://stat.gov.kz/industries/business-statistics/stat-tourism/spreadsheets/?year=2013&name=27161&period=&type= (accessed on 16 October 2026).

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, *90*(3), 414–427. [CrossRef]

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, *50*(2), 317–372. [CrossRef]

Cawley, M., & Gillmor, D. A. (2008). Integrated rural tourism: Concepts and practice. *Annals of Tourism Research*, *35*(2), 316–337. [CrossRef]

Chambwe, M., & Saayman, A. (2025). A framework for tourism value chain ownership in rural communities. *Community Development Journal*, *60*(2), 323–340. [CrossRef]

Comerio, N., Pacicco, F., & Serati, M. (2020). An analysis of sub-national tourism in Japan: Tourist and economic spillovers and their determinants. *Annals of Tourism Research*, *85*, 102881. [CrossRef]

Du, J., Zhao, B., & Feng, Y. (2024). Spatial distribution and influencing factors of rural tourism: A case study of Henan Province. *Heliyon*, *10*(7), e29039. [CrossRef] [PubMed]

Faber, B., & Gaubert, C. (2019). Tourism and economic development: Evidence from Mexico's coastline. *American Economic Review*, *109*(6), 2245–2293. [CrossRef]

Kim, Y. R., Williams, A. M., Park, S., & Chen, J. L. (2021). Spatial spillovers of agglomeration economies and productivity in the tourism industry: The case of the UK. *Tourism Management*, *82*, 104201. [CrossRef]

Kožić, I., & Sever, I. (2022). Can tourism activity stabilize cyclical employment? *Annals of Tourism Research Empirical Insights*, *3*(1), 100043. [CrossRef]

Kronenberg, K., & Fuchs, M. (2021). Aligning tourism's socio-economic impact with the United Nations' sustainable development goals. *Tourism Management Perspectives*, *39*, 100831. [CrossRef]

Ladkin, A., Mooney, S., Solnet, D., Baum, T., Robinson, R., & Yan, H. (2023). A review of research into tourism work and employment: Launching the Annals of Tourism Research curated collection on tourism work and employment. *Annals of Tourism Research*, *100*, 103554. [CrossRef]

Li, T., Li, Q., & Liu, J. (2021). The spatial mobility of rural tourism workforce: A case study from the micro analytical perspective. *Habitat International*, *110*, 102322. [CrossRef]

Li, Y., Ismail, M. A., & Aminuddin, A. (2024). How has rural tourism influenced the sustainable development of traditional villages? A systematic literature review. *Heliyon*, *10*(4), e25627. [CrossRef] [PubMed]

Luo, W., Timothy, D. J., Zhong, C., & Zhang, X. (2022). Influential factors in agrarian households' engagement in rural tourism development. *Tourism Management Perspectives*, *44*, 101009. [CrossRef]

MacKinnon, J. G., & Webb, M. D. (2017). Wild bootstrap inference for wildly different cluster sizes. *Journal of Applied Econometrics*, *32*(2), 233–254. [CrossRef]

Meyer, D. (2007). Pro-poor tourism: From leakages to linkages. A conceptual framework for creating linkages between the accommodation sector and 'poor' neighbouring communities. *Current Issues in Tourism*, *10*(6), 558–583. [CrossRef]

Mitchell, J. (2012). Value chain approaches to assessing the impact of tourism on low-income households in developing countries. *Journal of Sustainable Tourism*, *20*(3), 457–475. [CrossRef]

Mitchell, J., & Ashley, C. (2010). *Tourism and poverty reduction: Pathways to prosperity*. Earthscan. [CrossRef]

President of the Republic of Kazakhstan. (2018, June 19). *On some issues of the administrative-territorial structure of the Republic of Kazakhstan* [Decree No. 702]. Adilet Legal Information System. Available online: https://adilet.zan.kz/rus/docs/U1800000702 (accessed on 27 August 2026).

President of the Republic of Kazakhstan. (2022, May 3). *On some issues of the administrative-territorial structure of the Republic of Kazakhstan* [Decree No. 887]. Adilet Legal Information System. Available online: https://adilet.zan.kz/rus/docs/U2200000887 (accessed on 27 August 2026).

Republic of Kazakhstan. (2006, November 4). *On introducing an amendment and an addition to the Law of the Republic of Kazakhstan "On the Administrative-Territorial Structure of the Republic of Kazakhstan"* [Law No. 184]. Adilet Legal Information System. Available online: https://adilet.zan.kz/rus/docs/Z060000184_ (accessed on 16 October 2026).

Saparov, K., Omirzakova, M., Yeginbayeva, A., Sergeyeva, A., Saginov, K., & Askarova, G. (2024). Assessment for the sustainable development of components of the tourism and recreational potential of rural areas of the Aktobe Oblast of the Republic of Kazakhstan. *Sustainability*, *16*(9), 3838. [CrossRef]

Saxena, G., Clark, G., Oliver, T., & Ilbery, B. (2007). Conceptualizing integrated rural tourism. *Tourism Geographies*, *9*(4), 347–370. [CrossRef]

Scheyvens, R. (2007). Exploring the tourism–poverty nexus. *Current Issues in Tourism*, *10*(2–3), 231–254. [CrossRef]

Sharpley, R. (2002). Rural tourism and the challenge of tourism diversification: The case of Cyprus. *Tourism Management*, *23*(3), 233–244. [CrossRef]

Sharpley, R., & Telfer, D. J. (Eds.). (2014). *Tourism and development: Concepts and issues* (2nd ed.). Channel View Publications. [CrossRef]

Sun, Y.-Y., Li, M., Lenzen, M., Malik, A., & Pomponi, F. (2022). Tourism, job vulnerability and income inequality during the COVID-19 pandemic: A global perspective. *Annals of Tourism Research Empirical Insights*, *3*(1), 100046. [CrossRef]

Wang, Y., & Bai, H. (2023). The impact and regional heterogeneity analysis of tourism development on urban-rural income gap. *Economic Analysis and Policy*, *80*, 1539–1548. [CrossRef]

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). MIT Press.

Yang, J., Yang, R., Chen, M.-H., Su, C.-H., Zhi, Y., & Xi, J. (2021). Effects of rural revitalization on rural tourism. *Journal of Hospitality and Tourism Management*, *47*, 35–45. [CrossRef]

Zhumasultanov, T. (2011). Rasselenie naseleniya strany po gorodskoi i selskoi mestnostyam [Distribution of the country's population between urban and rural areas]. *Ekonomika i Statistika*, (3), 4–11. Available online: https://stat.gov.kz/upload/iblock/187/79em6y8401zj6jxzphgc2zxecdi9dobk/EandS1103End.pdf (accessed on 16 October 2026).

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.