Regional inequality in Mexico's post-pandemic economic recovery

Roberto Yoan Castillo Dieguez
(corresponding author)
Faculty of Accounting and Administration,
Autonomous University of Querétaro, Mexico
Email: roberto.castillo@uaq.mx

Samuel Arturo Mongrut Montalván
Department of Finance, Universidad del Pacífico, Peru & Faculty of Accounting and Administration, Autonomous University of Querétaro, Mexico

José Antonio Climent Hernández
Department of Systems, Azcapotzalco Unit, Metropolitan Autonomous University, Mexico

Ma. Benilde Rincón García
Faculty of Accounting and Administration, Autonomous University of Querétaro, Mexico

Graciela Lara Gómez
Faculty of Accounting and Administration, Autonomous University of Querétaro, Mexico

Keywords:
economic recovery,
regional inequality,
Covid-19 impact,
tourism dependence,
economic diversification

This study analysed how the Covid-19 pandemic has affected pre-existing regional economic disparities in Mexico by comparing economic recovery between tourism and non-tourism regions. Using a blend of econometrics and machine learning methods, this study examines major economic variables, such as gross domestic product, unemployment rate (UR), and foreign direct investment (FDI), between 2018 and 2023. This approach combines synthetic control method, random forest, and hierarchical clustering to address the differential regional recovery given economic structural variations.

Findings highlight a glaring economic dichotomy: the southeast and southwest regions of the USA that mainly are reliant on tourism experienced massive economic shocks and much longer recovery timelines owing to their dependence on tourism and informal labor markets. Conversely, more diverse economies, such as the northeast, maintained a more stable recovery through consistent FDI, which fostered some level of economic protection. SHAP and random forest analyses revealed that UR and FDI are key drivers of economic recovery. The unemployment rate negatively impacts tourism-dependent regions, whereas foreign direct investment fosters recovery in non-tourism-dependent regions.

These results indicate the need for economic diversification to act as a buffer from external shocks. Effective regional policies should work toward decreasing the reliance on tourism in vulnerable regions while making those areas more competitive for investment, thereby contributing to a more balanced recovery. Such an approach may cultivate structural resilience, thereby reducing the negative impact of future shocks and more broadly controlling for economic growth across Mexico.

Online first publication date: 25 September 2025

Introduction

The Covid-19 pandemic caused a considerable global economic shock that altered growth patterns and disrupted long-standing stability. In Mexico, the pandemic exposed weaknesses in the public health system and exacerbated economic disparities between regions through health system privatization. Lockdown measures, supply chain disruptions, and international trade interruptions had varying effects across Mexico, potentially widening existing economic and social gaps.

Over many years, the structural differences between north and central Mexico have been deeply investigated by scholars and policymakers, highlighting regional disparities in Mexico. Similarly, the economic dynamism of these regions, relating to their involvement in global value chains and a broad industrial base, is contrasted against the reliance of the southern regions on tourism and low-level informal labor, resulting in a cycle of poor economic performance over time. The Covid-19 pandemic has deepened the existing inequalities and made the economic recovery process complex, highlighting the need to figure out what lies behind the different ratios of economic recovery across regions.

This study aimed to analyse discrepancies in key economic indicators, such as regional gross domestic product (GDP), employment, and gross value added (GVA), that may impact the post-pandemic systemic recovery from 2020 to 2023. A methodological approach that incorporates a combination of reasoning, evidence, observation, difference-in-difference methods, and autodidactic algorithmic systems was developed. This methodological integration enables the observation of the immediate impacts of the pandemic on regional economies and the identification of non-linear patterns influencing mid-term recovery.

In finance, these events are commonly analysed using the event study methodology. However, this concept has been expanded here to evaluate the macroeconomic consequences of the pandemic (event). This method enables the identification of deviating developments visible as differences between expected economic growth (EG), derived from pre-Covid-19 trends, and actual EG. The synthetic control method (SCM) provides a rigorous framework to assess differences in how regions respond to health crises by constructing a synthetic counterfactual that closely mirrors the characteristics of regions that were less exposed to the health crisis. This allows for a precise comparison between areas that experienced substantial economic shocks and those whose structural attributes insulated them from major disruptions. Finally, through machine learning algorithms (random forest and clustering), complex patterns can be identified, grouping regions based on their recovery trajectories.

The remainder of this paper is organized as follows: the relevant literature presents a review of regional inequality and economic recovery. Then a detailed explanation of the methodology is provided, focusing on econometric and machine learning approaches, followed by the presentation and analysis of the results. Finally, the conclusions and suggestions for future research are presented.

Review of the literature

The regional inequalities in Mexico have been a recurring theme in academic literature, with great differences between the north and south (Castañeda 2024, Monroy-Gómez–Villagómez-Ornelas 2024). While the northern regions (i.e., those that have been more industrialized and integrated into global value chains) have displayed greater economic dynamism than their southern counterparts, the latter have mainly relied on the dependence of the primary sector, informal employment, and tourism (León-Cruz et al. 2024, Medina-Rivas et al. 2022, Mendoza-Velázquez et al. 2020, Preciado–Torrero 2021).

Before the pandemic, these differences were evident in GDP, the state competitiveness index, and the employment rate (IMCO 2023). However, the Covid-19 pandemic has worsened pre-existing inequalities, particularly in states dependent on tourism and informal work, such as the southern states (Casado 2021).

Numerous academic studies have investigated the effects of the pandemic. Kim et al. (2023) reported that the upturn in the most industrialized parts of the world has been faster than in less diversified regions, as they have been able to adjust to shifts in global demand and a rebound in international trade. Gleaming alongside statistics on the recent rebound of manufacturing jobs, regions that are less economically diversified are emerging more slowly from their malaise, highlighting the need for public policy built around infrastructure and economic development in these sectors.

The use of event studies in economic analysis facilitates the evaluation of the impact of abnormal events on macroeconomic variables. This approach is commonly employed in financial studies; however, recent studies have adapted this method to add new variables, such as the Covid-19 pandemic effect (Borusyak et al. 2024, Cheruiyot 2020, Zhang–Tian 2024). In addition, the difference-in-difference method has been the mainstay of public policy analysis to assess the different impacts of treatment (Bianchi et al. 2023, Kumar et al. 2023, Thakur–Das 2022). This method is used in this study to investigate the effects of the pandemic in multiple regions of Mexico.

In the regional economy, machine learning techniques have shown good efficacy in demonstrating non-linear relationships between complex patterns and large volumes of data. Among these techniques, random forest and cluster analysis have proven to be effective tools for capturing non-linear relationships between economic variables and understanding the behavior of the economy in crisis scenarios (Goulet 2020). This study employed these different methodologies in combination to provide a holistic picture of the inequalities present during Mexico's post-pandemic economic recovery.

Methodology

Event study with abnormal variations

The study employed a standard financial event study methodology to identify unexpected changes in important regional macroeconomic drivers, e.g., regional GDP, employment, and GVA. The Covid-19 pandemic represents the exogenous event of interest, which suddenly occurred in March 2020, altering productive systems worldwide, including the Mexican economy. The analysis period spanned from the 1st quarter of 2018 until the 4th quarter of 2023 to fully capture the impact dynamics and recovery process. This long-term context allows to capture not only the short-term impact of the shock but also the heterogeneity in recovery dynamics across regions.

Definition of the event

The economic repercussions of the pandemic were modeled as an external shock to the economy, with its effects examined over the quarters following its onset. This study provides a valuable framework for analysing economic recovery across different geographical regions and assessing variations in post-pandemic revival among states.

Analysis periods

To compare the periods before, during, and after the event, three phases were analysed:

Before the event (2018–2019): This phase is a baseline that provides the average of trends in the economy, which will inform predictions of what behavior would have looked like.

During the event (2020): This stage defines the immediate effects of the pandemic, featuring sharp drops in key economic indicators.

After the event (2021–2023): In this phase, the growth rates are modeled differently by region to assess the distinct recovery paths that may occur based on economic structure and ability to adapt.

Calculation of abnormal variations

Abnormal variations (AVs) are defined as observed growth minus estimated growth for individual regions, where the latter is an estimate of projected growth based on historical trends before the Covid-19. Mathematically, these variations are expressed as follows:

$$VI_{it} = R_{it} - E(R_{it}) \quad (1)$$

where VI_{it} denotes the unexpected variation for region i in time t; R_{it} is the observed real EG, and $E(R_{it})$ is the estimated EG.

SCM

SCM is an analytical method for evaluating the treatment effect on a group by constructing a synthetic version of the treated unit through a weighted combination of untreated units. It is used to create a counterfactual for each studied region, enabling a direct, like-for-like comparison of economic performance between affected regions and those with greater structural resilience.

The SCM model is designed to minimize differences in the pretreatment period between the treated unit and the synthetic control (SC) unit. To achieve this, it employs an automatic penalization function that reduces discrepancies in the matching process.

$$\min \sum_t (Y_{it} - \sum_{j \neq i} w_j Y_{jt})^2 + \lambda \sum_{j,k} (w_j Y_{jt} - w_k Y_{kt})^2 \quad (2)$$

where Y_{it} represents the value of the outcome variable for the treated region i at time t; w_j denotes the weights assigned to each control unit j that build the counterfactual, and λ is the penalty coefficient that balances the difference between the treated unit and the SC.

Validity and robustness of the SC model

Temporal and spatial placebo tests were conducted to verify the reliability of the estimates:

Placebo over time: This analysis involves varying the treatment period to a time prior to the pandemic (e.g., 1st quarter of 2018) and checking whether or not the counterfactual and treated unit are significantly different before the actual event. If no substantial divergence is observed before 2020, it confirms the reliability of the counterfactual estimates and strengthens the validity of the model.

Space placebo: "Treatment" is (falsely) assigned to regions less affected than the treated region. The effects estimated from the SC in the treated region can be compared with a distribution of placebo or counterfactuals. Therefore, this approach can provide at least a measure of relative significance of the impact of the pandemic on the treated region.

Penalized and differentiated SC model

Penalized SCM was used to optimize the adjustment by introducing a penalty coefficient that balances the accuracy of the synthetic control unit before and after the correction process. Meanwhile, differentiated SCM was independently applied to analyse variables where trends over time are more relevant than absolute levels. This approach enables the estimation of economic recovery effects without imposing unrealistic linearity assumptions on the underlying economic relationships.

Machine learning

Given the significant complexity and non-linearity that drives the regional economic recovery process, this study employed advanced machine learning techniques to identify critical factors and classify regions by their recovery trajectories – random forest and clustering.

Random forest

The random forest algorithm was used to identify the most significant variables that explain the differences in the recovery speed between regions. This non-parametric technique allows to model complex interactions between variables without the need to assume predetermined linear relations.

Clustering

Clustering analysis was employed to group regions according to their post-pandemic economic behavior, allowing to identify common patterns between regions that rapidly recovered and those that experienced a persistent lag in their recovery.

Data

The data used in this study were obtained from reliable official sources, such as the National Institute of Statistics and Geography, the Bank of Mexico, and the Secretariat of Finance and Public Credit. Quarterly data series were used, from the 1st quarter of 2018 to the 4th quarter of 2023, which allows to accurately capture fluctuations in the key indicators of EG and employment throughout the analysis period.

Results

Before proceeding with volatility modeling using the GARCH model and an AV equation, it is necessary to confirm whether the analysed time series – foreign direct investment (FDI), unemployment rate (UR), and economic growth (EG) – are stationary. Stationarity is a fundamental condition that must be met to ensure that fluctuations in the series remain stable over time. If stationarity is not achieved, uncontrolled temporal trends may lead to spurious results, compromising the reliability of the model.

To assess stationarity, the augmented Dickey–Fuller (ADF) unit root test was conducted on each series. The findings indicated that EG exhibited stationarity at the level form (i.e., without differencing), as evidenced by a P-value of 3.79e−15, which was well below the conventional threshold of 0.05. This result confirms the rejection of the null hypothesis of the ADF test, indicating that EG is stationary in its original form. In contrast, the time series analysis of FDI and UR revealed that neither variable exhibited stationarity in their raw form. The *P*-values for FDI and UR were 0.727 and 0.160, respectively, exceeding the conventional threshold of 0.05. Consequently, the null hypothesis could not be rejected, suggesting that these variables require differencing to achieve stationarity.

As the FDI and the UR did not exhibit stationarity, both series were differentiated once, applying the transformation of first-order differences. This technique is commonly employed to convert non-stationary series into stationary ones, eliminating long-term trends and allowing fluctuations to be more stable over time.

After differentiation, the ADF test was again applied to the transformed series. The results indicated that both differentiated series were stationary, with *P*-values of 2.61e−09 and 4.66e−26 for the differentiated series of FDI and UR, respectively, confirming that both series are stationary in their first differences (see Table 1).

Table 1

ADF test results

Variable	Statistical ADF	*P*-value	Result
FDI (at level)	−1.068	0.7275	non-stationary
Unemployment rate (in level)	−2.336	0.1605	non-stationary
Economic growth (in level)	−9.092	3.79e−15	stationary
FDI (differentiated)	−6.773	2.61e−09	stationary
Unemployment rate (differentiated)	−13.952	4.66e−26	stationary

With the three time series transformed into stationary, it was possible to proceed to the next step of the analysis, which included modeling of conditional volatility using the GARCH model and calculation of AVs. This validation ensures that the results obtained in later stages of the analysis are based on stable series and that the econometric assumptions of the model are robust.

GARCH model results

The GARCH model was used to implement the volatility dynamics of the major economic variables affected by the pandemic, namely, FDI, UR, and EG. It enables modeling of conditional volatility in financial and economic series such that a measure of changeability to exogenous shocks, such as Covid-19, is available. The following information summarizes the results of the model for each variable.

Results of the GARCH model

The main features provided by the GARCH model for all the time series are presented in Table 2. As the coefficients of ARCH (α) and GARCH (β) represent the temporal dependence of volatility on each variable, their sum ($\alpha + \beta$) shows how long the impact of a given volatility will last.

Table 2

Summary of the GARCH model results

Variable	Average (μ)	α (ARCH)	β (GARCH)	Persistence (α + β)	Remarks
Economic growth (in level)	−1.0132e+07 (p = 0.797)	0.3453 (p < 0.001)	0.6547 (p < 0.001)	1.000	High persistence in volatility, indicating long-term effects
FDI (differentiated)	3.8881e−04 (p < 0.001)	0.0100 (p = 0.878)	0.9700 (p < 0.001)	0.980	Persistently high volatility, although the contribution of α is low
Unemployment rate (differentiated)	0.5610 (p = 0.151)	0.7050 (p = 0.016)	0.2950 (p < 0.001)	1.000	Less-persistent volatility and more subject to short-term fluctuations

The overall persistence (α + β) in all variables is close to 1, recommending that volatility shocks are persistent, which is particularly true due to the high dependence of FDI and UR on previous shocks. This persistence implies a lasting reaction in these series to the uncertainty created by the pandemic, indicating considerable stabilization difficulties in the short and medium terms.

Summary of the results of the GARCH model

The analysis of conditional volatility reveals significant fluctuations in the time series throughout the study period. Figure 1 presents the per-series behavior:

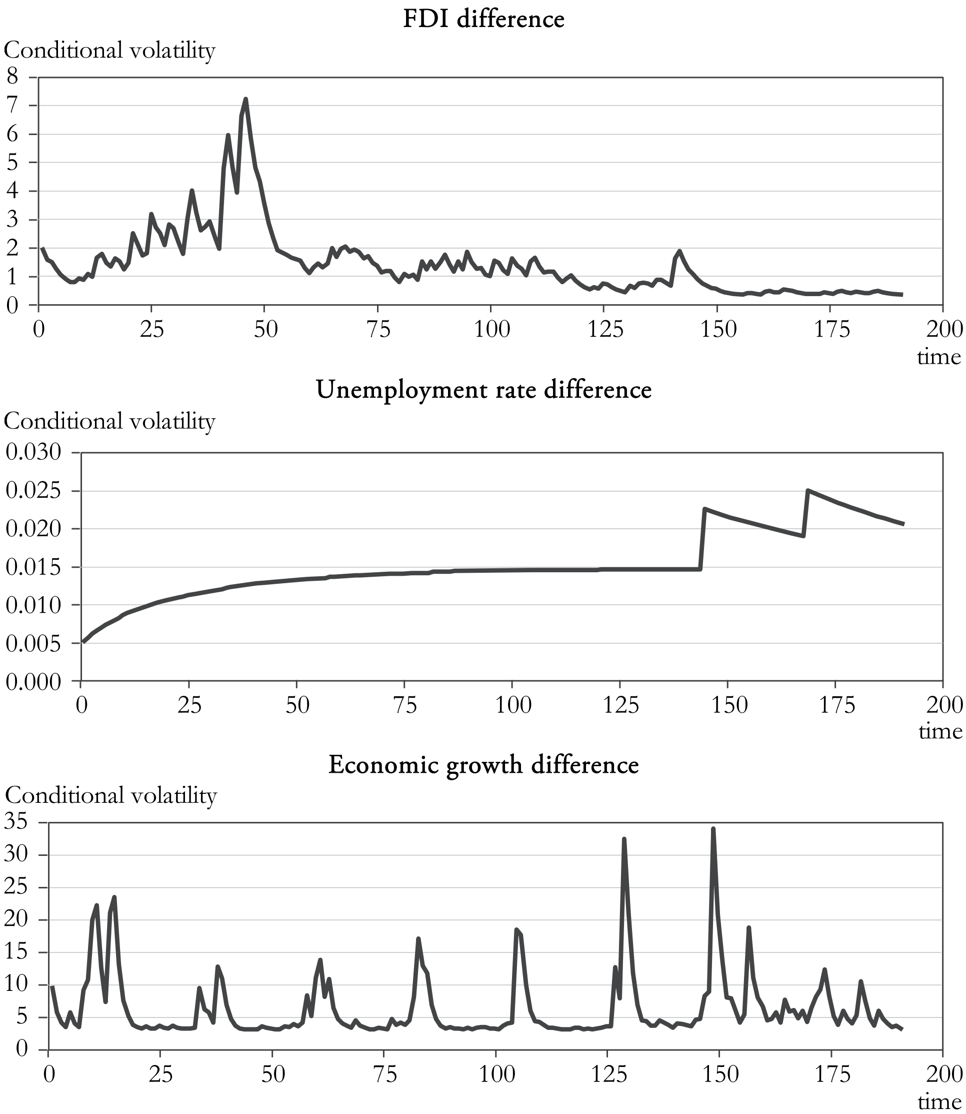
FDI: The volatility of FDI exhibited very high peaks at the onset of the pandemic, suggesting heightened uncertainty surrounding foreign direct investment in Mexico during this period. These peaks reflect the sensitivity of FDI to external factors and global uncertainty, particularly in sectors highly interconnected with international capital flows. Over time, volatility gradually stabilizes, indicating that investment flows are adapting to a new market equilibrium.

UR: The conditional volatility of the unemployment rate (UR) shows a sustained increase following the crisis, with a gradual return to stability rather than an immediate correction. This pattern mirrors the slow adjustment dynamics of the labor market, which typically lags behind other economic sectors in recovering from shocks. The persistent volatility in UR suggests underlying structural weaknesses in the labor market, exacerbated by the crisis, with a slow and delayed recovery.

EG: EG has cyclical volatility dynamics, where its cyclical peaks reflect repeated efforts at recovery. These cycles may be associated with stimulus fiscal and monetary policies or unbalanced recovery between sectors of the economy. This lower persistence in volatility suggests that while individual shocks to this variable remain persistent, the effects of each shock on recovery tend to be less persistent than those for any of the other variables examined.

Figure 1

Conditional volatility of regional macroeconomic variables



The GARCH model results indicate that the response of all economic variables to the pandemic shock varies over time and in terms of volatility persistence. The coefficient of standard deviation shows high persistence in FDI and UR, suggesting that the effects of the pandemic are long-lasting (Camino-Mogro–Armijos 2022). In contrast, EG exhibits a more cyclical behavior, fluctuating around its mean and stabilizing over time without enduring long-term effects.

Cumulative abnormal returns (CAR) and statistical tests

Several statistical tests were applied on the CAR of each variable (*t*-test, adjusted Patell, generalized rank test, and permutation test) to analyse their significance. Using previously derived results, we can determine whether the CAR values are significantly different from zero and study their sign and magnitude over various horizons in terms of economic recovery.

Table 3

Summary of cumulative abnormal returns (CAR) and statistical test results

Variable	Test	Statistic	*P*-value	CAAR value	Pos CAR[a)]	Number of CARs
FDI abnormal variation	*t*-test	5.025	2.36e−05	1.041	26:4	30
	adjusted Patell	5.111	3.20e−07	-	-	-
	generalized rank	0.000	-	-	-	-
	permutation test	1.041	0.995	-	-	-
Unemployment rate abnormal variation	*t*-test	3.406	1.95e−03	0.809	21:9	30
	adjusted Patell	3.464	5.33e−04	-	-	-
	generalized rank	0.000	-	-	-	-
	permutation test	0.809	0.315	-	-	-
Economic growth abnormal variation	*t*-test	−12.250	5.48e−13	−2.745	0:30	30
	adjusted Patell	−12.459	0.000	-	-	-
	generalized rank	0.000	-	-	-	-
	permutation test	−2.745	0.426	-	-	-

a) Pos CAR (positive CAR): indicates the number of observations with positive abnormal returns versus negative ones in the sample.

The **FDI** showed positive skewness in the CAAR values. A positive to negative ratio of 26:4 implies that the trends on FDI recover from early pandemic hiccups. This type of behavior is consistent with existing literature where FDI adjusts relatively quickly to new market conditions after shocks owing to its high degrees of flexibility, particularly in the case of sectors and flow types that attract the highest proportion of international capital (Micheli 2020). The recovery is statistically significant, as indicated by the large *P*-values (<0.001) in the *t* and adjusted Patell tests. These statistical tests are useful in event studies to assess the crisis and post-crisis effects on an unstable economic series (Krawczyk et al. 2023, Moudud-Ul-Huq et al. 2020).

The **UR** also has a positive CAR (CAAR = 0.809) and a positive to negative ratio of 21:9 in CARs. This shows a rebound in the labor market, but with less fluctuation compared with FDI. Both the *t* and adjusted Patell tests are statistically significant ($P < 0.01$), indicating a reduction in the abnormal variability that provides post-crisis stabilizing dynamics.

The negative CAAR (−2.745) and the positive-to-negative ratio of 0:30 in CARs indicate an unfavorable response for EG as this important indicator is either moving slowly toward recovery or is unable to fully recover compared with other indicators. This negatively significant result is highly statistically significant in all tests ($P < 0.001$), indicating that EG remains considerably below its pre-crisis levels, highlighting the prolonged impact of the economic downturn. Thus, prolonged crises and external shocks can make EG more sensitive (Chu et al. 2023, Lee 2024), particularly with regard to the structural constraints of domestic demand and production.

In summary, the results indicate that the recovery of FDI and UR has been statistically significant and positive, suggesting that these indicators have successfully adapted to the new market conditions. In contrast, EG exhibits a pronounced negative trend, implying that the broader economy has been more severely impacted by the pandemic, possibly due to structural constraints in production and consumption recovery.

This divergence in recovery rates among economic indicators aligns with findings from international studies on post-crisis economic recoveries, which show that FDI and UR tend to stabilize more quickly than EG due to their distinct underlying structures (Crescenzi–Ganau 2024, Sánchez–Cuadrado-Roura 2024).

Given the presence of structural barriers to growth, policy measures should not remain as temporary relief strategies but should instead be integrated into a broader framework aimed at fostering long-term economic improvement and addressing the persistent state of negative growth observed in this study.

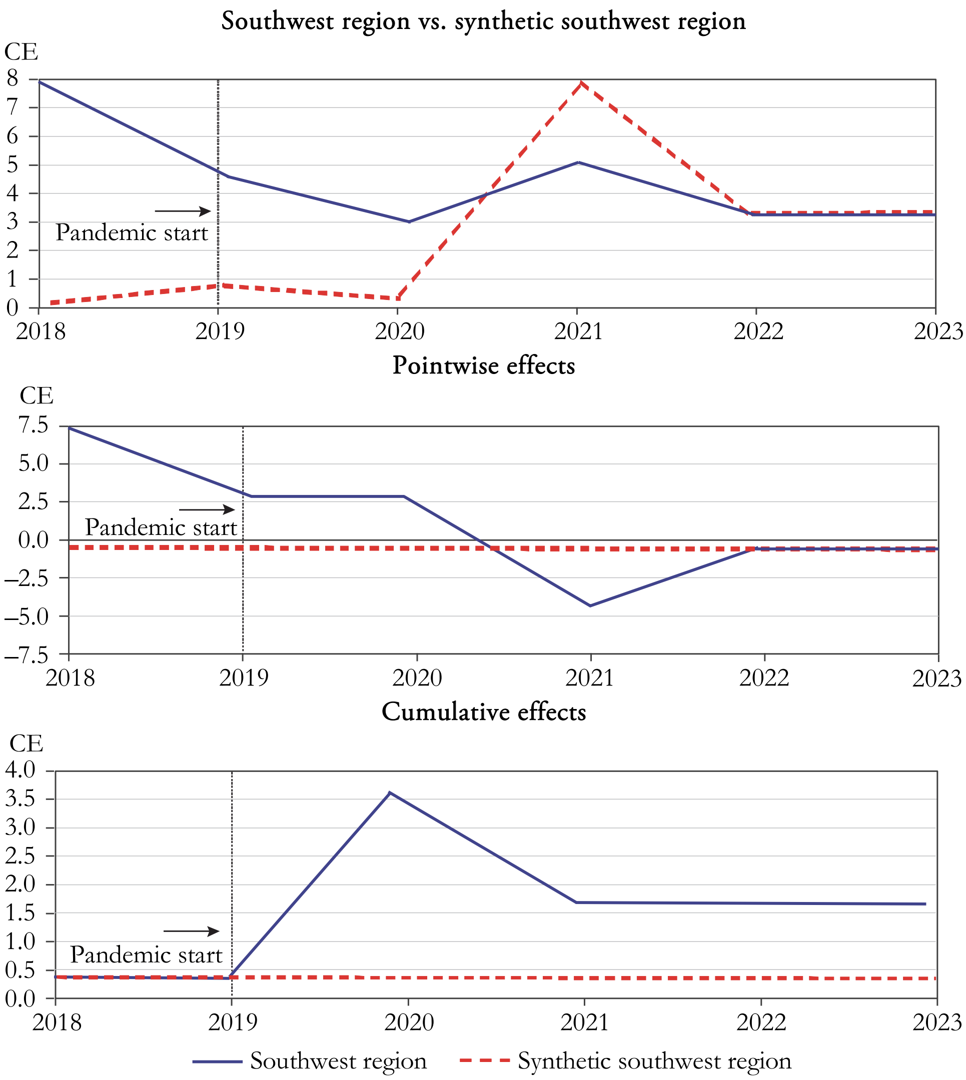
The effect of the pandemic on the regional economic disparities in Mexico

To consolidate the results about the post-pandemic economic recovery in Mexico, a comparison was made between a tourism-dependent region and one with a less tourism-dependent economy. This comparison allows for the evaluation of the differential effects of the pandemic in regions with different economic structures, particularly in terms of vulnerability and resilience to external shocks (Ding et al. 2024, Fabio–Pietro 2020, Sakyi-Nyarko et al. 2022).

Figure 2 presents the evolution of EG in the Southwest region and its SC over the analysis period: We found evidence that the EG in this region significantly decreased relative to a counterfactual estimate after the pandemic. It is characteristic of this behavior that the dependence on tourism thereby amplifies effects on external shocks as it does not have an economy capable of diversifying and cushioning such a blow (Charfeddine–Dawd 2024, Falk et al. 2023, Langer–Schmude 2023).

Figure 2

Economic recovery trends: southwest region vs. synthetic control



Note: the economic recovery of the southwest region compared to its synthetic control, constructed from a weighted combination of other regions to assess the pandemic's impact.

Table 4 presents the weights assigned to each region such that they can be combined to create the SC and allow us to understand the nature of the counterfactual. The weights represent the optimal combination of less tourism-dependent regions selected by the model to replicate the economic trajectory that the southwest region would have followed in the absence of Covid-19. This supports the argument that the region's existing economic structure, along with its varying levels of resilience compared to the control regions, contributes to its heightened vulnerability during crisis events (Kitsos et al. 2023, Wang et al. 2023).

Table 4

Weights of the regions in the synthetic control of the southwest region

Region	Weight
North-central	0.138935
Central-southern	0.141018
Northeast	0.144671
Northwest	0.142181
Western	0.139827
Eastern	0.119119
Southeast	0.174249

Also, the root mean squared error of prediction (RMSPE) in the pre-pandemic and post-pandemic periods supports the correctness of the model. Specifically, the RMSPE after the intervention (i.e., the pandemic shock) is significantly lower than the RMSPE before the intervention, for which the actual EG and equilibrium growth deviated, as shown in Table 5. This increase in RMSPE indicates the substantial impact that the pandemic had on the economic stability of the region. (Fan et al. 2024, Glebocki Keefe–Hepp 2024).

Table 5

Root mean squared prediction error (RMSPE) for the southwest region

Unit	Pre_rmspe	Post_rmspe	Post/Pre
Southwest region	6.597686	2.087529	0.316403

Table 6 presents a summary of the major economic indicators for the southwest region compared with its synthetic counterfactual. The outcomes show significant differences in the EG, UR, and FDI. However, they also confirm how reliant the region's economy is on external shocks.

Table 6

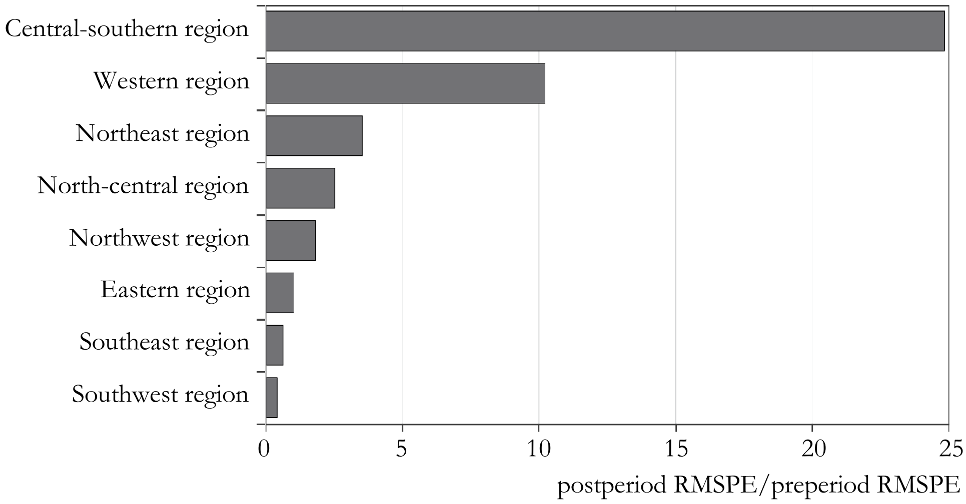
Comparison of the economic variables between the southwest region and its synthetic counterfactual

Variable	Southwest region	Synthetic southwest region	WMAPE	Importance
EG	6.570	0.230	6.340	0.52
UR	0.020	0.060	0.040	0.20
FDI	1.9355e+08	1.1845e+09	9.9103e+08	0.28

Figure 3 comparison of RMSPE in the post-pandemic and pre-pandemic periods across different regions to assess model performance. In this graph, the southwest region exhibits the lowest RMSPE ratio, indicating that deviations after the intervention are more stable. This result is crucial for validating the robustness of the synthetic estimate for the southwest region compared to other less affected regions or those with different production structures.

Figure 3

Comparison of post-pandemic RMSPE ratios across the treated and placebo regions



Note: "treated regions" refer to those directly affected by the intervention (i.e., the pandemic shock), while "placebo regions" are used as controls in the synthetic control method.

Figure 4 presents a comparison of the actual growth rate of the real GDP in the non-tourism-dependent northeast region and its synthetic counterfactual on the entire analysis period. After the onset of the pandemic in 2020, the observed GDP growth rate (represented by the blue line) declined significantly, diverging from pre-pandemic trends. The synthetic counterfactual (represented by the red dotted line) illustrates the projected economic trajectory that the northeast region would have followed in the absence of the pandemic.

However, the deviation of the northeast region from the counterfactual is less pronounced compared to tourism-dependent regions, indicating that its economy was less vulnerable to the pandemic shock – likely due to its more diversified economic structure. The point and cumulative effects show that, although there was an initial drop in economic growth (EG), the region demonstrated a steady recovery in subsequent quarters. This suggests a greater capacity for economic adjustment and resilience, allowing it to absorb the crisis impact more effectively than tourism-dependent regions.

Figure 4

Graph of the northeast region vs. synthetic control

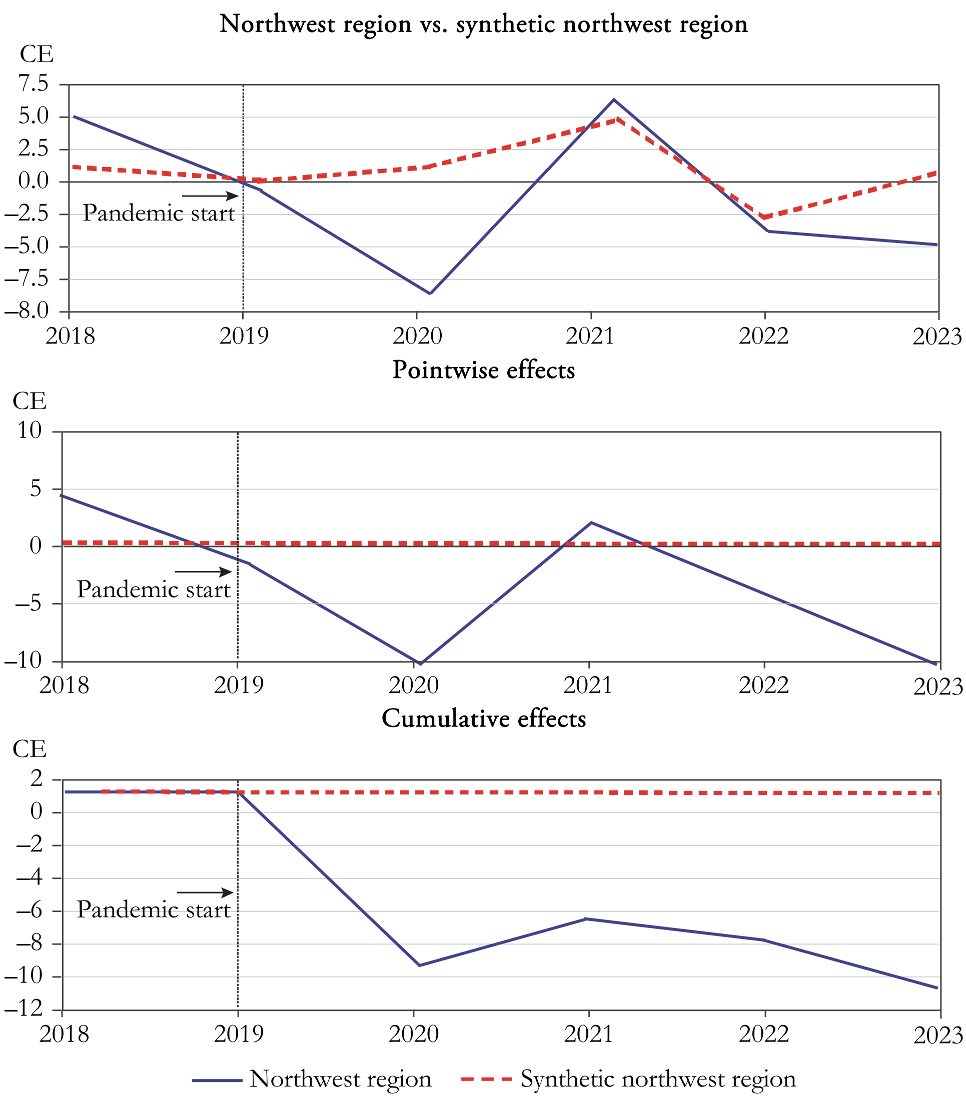


Table 7 shows the weights allocated to each region while creating the synthetic counterfactual for the northeast region. The weights represent a balanced combination of economic diversity and low tourism dependency. This selection by the model enables the construction of a counterfactual trajectory that simulates how the northeast region's economy would have evolved over time in the absence of the pandemic. This reinforces the model's robustness by providing a strong basis for comparison. These results align with the more balanced economic structure of this region, further supporting the idea that lower tourism dependence contributes to greater economic stability and resilience against external shocks, such as those induced by the pandemic.

Table 7

Weights of the regions in the synthetic control of the northeast region

Region	Weight
North-central	0.142840
Central-southern	0.142841
Northwest	0.142842
Western	0.142840
Eastern	0.142825
Southeast	0.142865
Southwest	0.142947

Table 8 presents the RMSPEs for the pretreatment and post-treatment periods in the northeast region. The higher post-treatment RMSPE compared to the pretreatment RMSPE indicates that the pandemic had a significant impact on the region's economic performance. However, this effect was notably smaller than in tourism-dependent regions, where economic disruptions were more pronounced. The relatively mild divergence in RMSPE suggests that the northeast region's more diversified economy provided greater resilience to external shocks.

Table 8

Root mean squared prediction error (RMSPE) for the northeast region

Unit	Pre_RMSPE	Post_RMSPE	Post/Pre
Northeast region	1.541139	5.041737	0.316403

Table 9 presents the comparative values of the main economic variables for the northeast region and its synthetic counterfactual. The difference in EG is significant, whereas the differences in UR and FDI are not statistically significant. The WMAPE and relative importance analysis indicate that EG had the greatest weight in estimating the pandemic's effects on the northeast region.

Table 9

Comparison of economic variables between the southwest region and its synthetic counterfactual

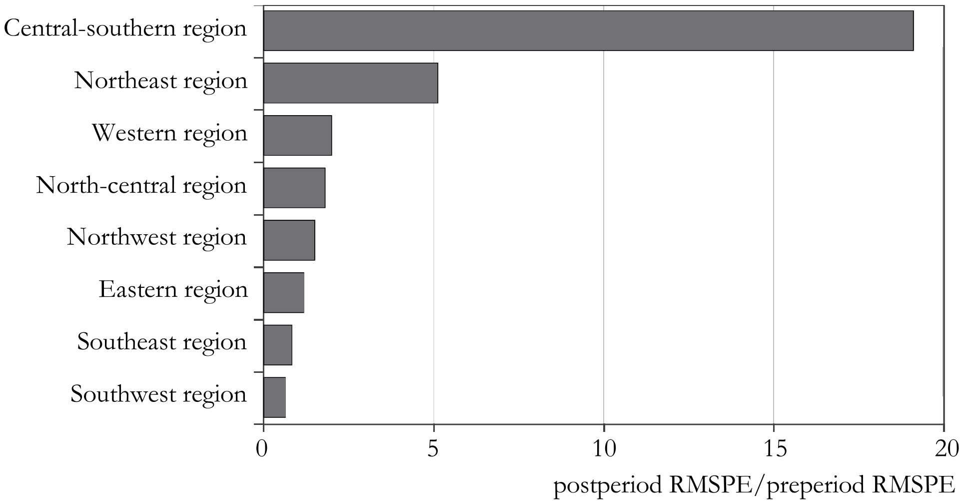
Variable	Northeast region	Synthetic northeast region	WMAPE	Importance
EG	1.6	0.76	2.59	1.0
UR	0.04	0.06	0.03	0.0
FDI	1.859626e+09	9.613063e+08	1.109442e+09	0.0

Figure 5 compares the RMSPE ratios of placebo regions, plotted against each other for the same markets (MSEPS values are presented in their natural logarithms).

The central-southern region exhibits the highest post-pandemic to pre-pandemic RMSPE ratio, indicating that its economic trajectory has significantly diverged from the synthetic counterfactual since the pandemic. In contrast, the northeast region shows a more moderate increase in this ratio, suggesting that while the pandemic impacted its growth, its lower reliance on tourism provided some degree of resilience against the economic shock. This evidence supports the idea that economies with greater sectoral diversification tend to be more resistant to crises (Trippl et al. 2024).

Figure 5

Comparison of the post-pandemic RMSPE ratios across the treated and placebo regions



Our results provide evidence of a clear post-pandemic economic divide between tourism- and non-tourism-dependent regions in Mexico. The different economic paths of tourism regions, particularly the southwest, indicated their susceptibility to external shocks. Conversely, regions where the economy is more diversified, such as the northeast, steadily and resiliently adjusted to the post-pandemic environment. Economic diversification can provide a buffer against global shocks and the economic impact of lockdown measures, but such benefits are by no means guaranteed, highlighting the need for policies to develop structural resilience in regions with high direct or indirect domestic exposure (Crescenzi–Ganau 2024, Trippl et al. 2024).

Regional analysis of post-pandemic economic recovery

The variable importance analysis and SHAP method were used to assess the impact of UR and FDI on EG in Mexico's regions after the pandemic. Figure 6 presents the feature importance analysis, showing that UR had the greatest effect in tourism-dependent regions, as these areas are more sensitive to external shocks (e.g., southwest region) (Lee 2023, Martínez Rodríguez–Moreno 2023, Monterrubio 2022). Contrarily, in the northeast, a region less affected by tourism dependence on external income and expenditures, the growth effects of unemployment are smaller. Figure 7 illustrates the SHAP values, confirming that in these areas, higher unemployment values have a more significant negative impact on growth.

Figure 6

Importance of variables in the random forest model

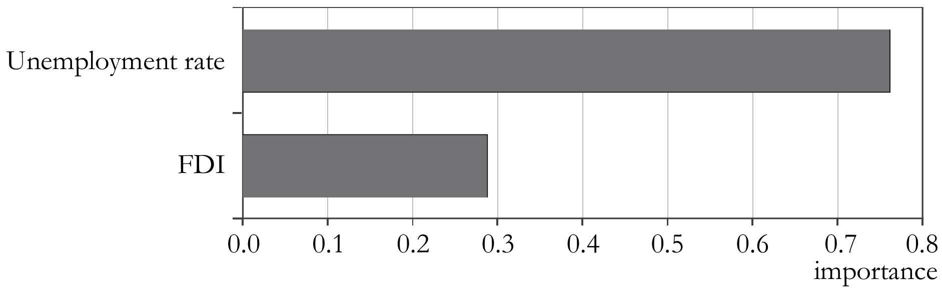
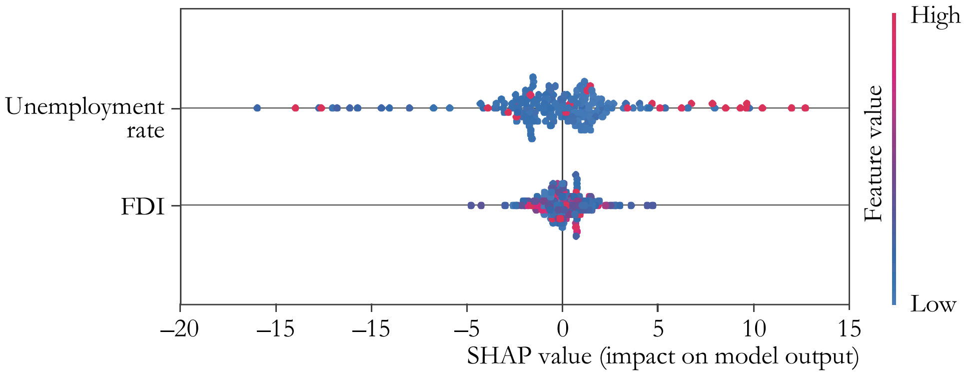


Figure 7

SHAP values to evaluate the impact of variables on model predictions



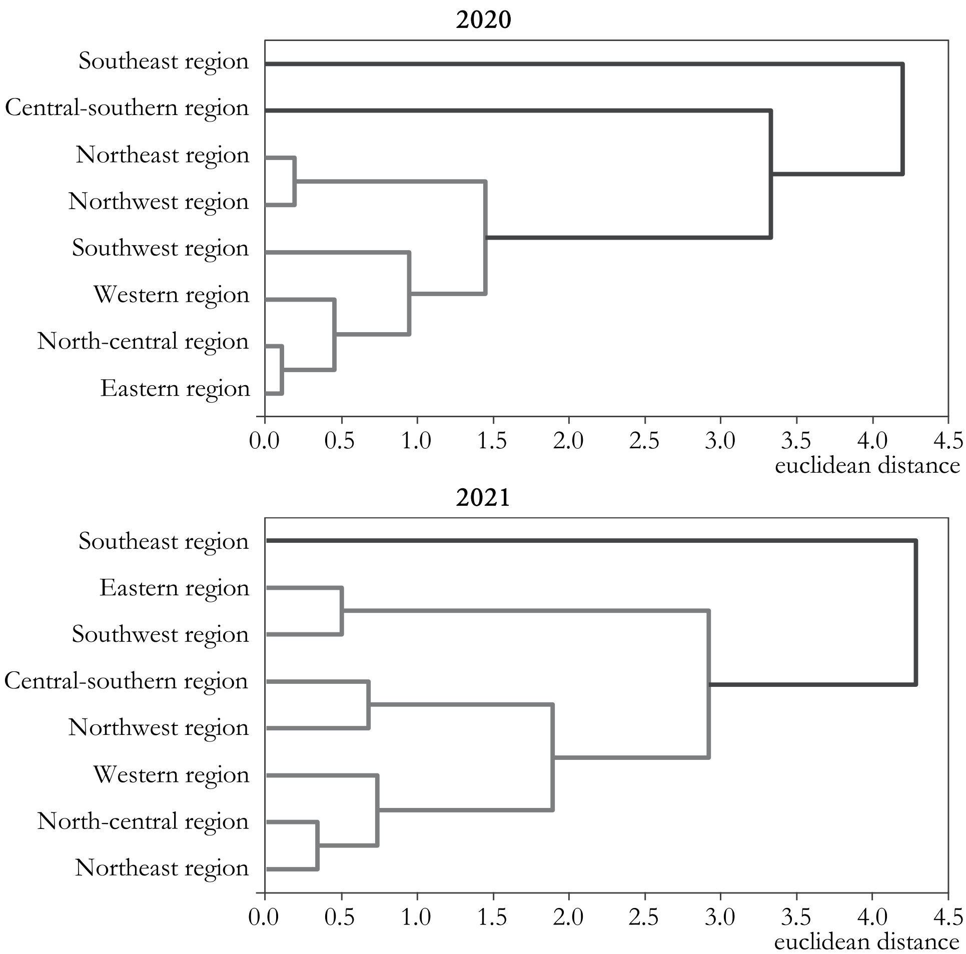
FDI promotes economic recovery in regions independent of tourism. In these areas, FDI reduces the negative impact of the pandemic, increasing resilience to future shocks (Kammoun–Ben Romdhane 2024, Ngo 2024). In contrast, tourism-dependent regions, which are less attractive to FDI, experience a slower recovery in economic growth (Anuradha et al. 2024). These findings indicate the need to diversify the economies in tourism-dependent regions to reduce their vulnerability. Policies that attract investment and strengthen other economic sectors would help improve the recovery and economic stability of these regions.

Comparative analysis of hierarchical clustering and random forest for regional economic recovery

The hierarchical clustering dendrogram for 2020 and 2021 shows how the regions in Mexico clustered into groups with similar responses (economically) as the pandemic unfolded. The separate cluster of the southeast and southwest regions in 2020 (see Figure 8) indicates the vulnerability of these tourism-dependent regions to pandemic-associated economic shocks. Clusters of separated regions, such as the northeast and some in central-southern regions, may indicate more resilience over this period owing to diversification.

Figure 8

Hierarchical clustering of regions



A small change in clustering by 2021 (see Figure 8) suggests tentative signs of recovery among tourism-dependent regions, particularly the southwest region, which appears to be shifting toward economic patterns more characteristic of diversified regions. This shift could indicate that these regions are gradually adapting to new economic conditions, possibly through diversification or an increase in foreign investment inflows. This historical trend highlights the importance of economic structure for regional resilience, as more diversified regions tend to experience a smoother recovery path.

Conclusions

The onset of the Covid-19 pandemic and its residual effects have already taken a severe toll on economies worldwide, including Mexico. The effects of the health crisis on regional economic inequality in the country have drawn attention as a result of this study. Using an econometric and machine learning approach, the relative importance of different factors in explaining the regional economic recovery was examined, emphasizing that UR and FDI are key determinants of EG. The findings present important implications to better understand the differences in the response to an unprecedented crisis and to establish public policies that reduce vulnerability to external shocks in the future.

The economic structure of each region has been crucial for their post-pandemic resilience, and this is the first takeaway. More diverse economies, such as the northeast region, quickly recovered steadily and strongly due to their minimal reliance on tourism and their ability to draw foreign investment. This is because the industrial and manufacturing sectors in these regions are part of global value chains that enabled them to recover more swiftly after the global demand shocks and the resumption of international trade. In contrast, regions that demonstrated high reliance on tourism, such as the Southeast and Southwest regions tended to be more susceptible to the impacts of tourism activity loss and the precariousness of informal jobs, a considerable aspect of their economy.

Meanwhile, the random forest and SHAP analyses revealed that the two important variables observed in regional economic recovery are UR and FDI. For tourist regions, the UR exerted a notable drag on EG, which shows how protracted unemployment and job volatility throttled recovery. In contrast, FDI played a crucial role in sustaining economic recovery and mitigating the adverse effects of the pandemic in more diversified economies, where it contributed to maintaining and enhancing sectoral diversification. This implies that a region's attractiveness for investment is a key factor in its economic resilience, particularly during crises.

The hierarchical clustering approaches implemented in 2020 and 2021 identified groups of regions based on their economic response to the pandemic, particularly distinguishing between tourism-dependent economies and those with more diversified structures. In 2020, some tourism-dependent regions clustered closely together, indicating that the economic shock had a similarly severe impact across these areas. By 2021, the southwest region exhibited early signs of improvement as it moved toward clusters of more diversified economies. This shift suggests that certain tourism-dependent regions have begun adapting to evolving economic conditions – either by diversifying their economic activities or by attracting FDI into non-tourism sectors.

These results highlight the importance of maintaining and expanding economic diversification policies in regions affected by the pandemic. Overdependence on highly sensitive sectors, such as tourism, limits resilience to global shocks. The pandemic has demonstrated that regions with more diversified economies experienced a lower overall economic impact and exhibited significantly greater adaptive and resilient capacities. Therefore, it is essential that regional policies go beyond short-term recovery efforts and focus on reducing structural economic weaknesses by encouraging investment in diverse sectors and promoting job formalization.

To sum up, although the economic recovery from the pandemic has not been homogeneous in Mexico, structural characteristics regarding regions have been key to its development. Development policies should focus on reducing economic dependence on highly vulnerable sectors, such as tourism, and fostering conditions that attract investment in activities that generate formal and stable employment. In an economic context, the pandemic experience leaves a fundamental teaching about building a much stronger and more diversified regional economy that allows various regions of the country to cope with external shocks in a more resilient and stable manner. These efforts would lead not only to faster and more sustainable recovery but also to fairer and more balanced development in the long run.

REFERENCES

BIANCHI, F.–KUNG, H.–TIRSKIKH, M. (2023): The origins and effects of macroeconomic uncertainty *Quantitative Economics* 14 (3): 855–896. https://doi.org/10.3982/QE1979

BORUSYAK, K.–JARAVEL, X.–SPIESS, J. (2024): Revisiting event-study designs: robust and efficient estimation *Review of Economic Studies* 91 (6): 3253–3285. https://doi.org/10.1093/restud/rdae007

CAMINO-MOGRO, S.–ARMIJOS, M. (2022): Short-term effects of Covid-19 lockdown on foreign direct investment: evidence from Ecuadorian firms *Journal of International Development* 34 (4): 715–736. https://doi.org/10.1002/jid.3598

CASADO, J. M. (2021): De crisis sanitaria a crisis económica y laboral: patrones espaciales del impacto de la Covid-19 en el empleo formal de México *Investigaciones Geográficas* 104. https://doi.org/10.14350/rig.60212

CASTAÑEDA, D. (2024): Moderate opulence: the evolution of wealth inequality in Mexico in its first century of independence *Explorations in Economic History* 92: 101567. https://doi.org/10.1016/j.eeh.2023.101567

CHARFEDDINE, L.–DAWD, I. (2024): Analysis of the resilience of the Turkey tourism industry to exogenous shocks: new evidence from a NARDL model *Tourism Recreation Research* 49 (3): 529–544. https://doi.org/10.1080/02508281.2022.2026615

CHERUIYOT, K. (2020): Heterogeneous relationships between income levels and associated correlates in Gauteng province, South Africa: quantile regression approach *Development Southern Africa* 37 (6): 871–887. https://doi.org/10.1080/0376835X.2019.1701415

CHU, L. K.–TRUONG, H. H. D.–DUNG, H. P. (2023): Unveiling the influence of economic complexity and economic shocks on output growth volatility: evidence from a global sample *Eurasian Economic Review* 13 (3–4): 637–676. https://doi.org/10.1007/s40822-023-00246-8

CRESCENZI, R.–GANAU, R. (2024): Inward FDI and regional performance in Europe after the great recession *Cambridge Journal of Regions, Economy and Society* 18 (1): 167–192. https://doi.org/10.1093/cjres/rsae038

DING, C.–GAO, X.–XIE, Z. (2024): Analysing the differential impact of the Covid-19 pandemic on the resilience of the tourism economy: a case study of the Chengdu-Chongqing urban agglomeration in China *International Journal of Disaster Risk Reduction* 102: 104255. https://doi.org/10.1016/j.ijdrr.2024.104255

FABIO, M.–PIETRO, P. (2020): *Resilience and convergence: short vs. long-run regional effects of economic crises and macroeconomic policies* pp. 431–452. https://doi.org/10.14650/98286

FALK, M. T.–HAGSTEN, E.–LIN, X. (2023): Persistence of an external shock to domestic tourism demand *Scandinavian Journal of Hospitality and Tourism* 23 (5): 434–448. https://doi.org/10.1080/15022250.2024.2332300

FAN, F.–WENG, Z.–TIAN, J. (2024): Impact of government support policies on regional economic resilience under the Covid-19 outbreak *Technological and Economic Development of Economy* 30 (1): 74–106. https://doi.org/10.3846/tede.2024.20397

GLEBOCKI KEEFE, H.–HEPP, R. (2024): The effects of European fiscal discipline measures on current account balances. *International Economics and Economic Policy* 21: 251–283. https://doi.org/10.1007/s10368-024-00587-y

GOULET, P, C. (2020): The macroeconomy as a random forest *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3633110

KAMMOUN, S.–BEN ROMDHANE, Y. (2024): Impact of government responses to Covid-19 on the resilience of FDI attractiveness factors in the Asian region *Journal of Economic and Administrative Sciences* 40 (1): 37–56. https://doi.org/10.1108/JEAS-03-2022-0085

KIM, A.–LIM, J.–COLLETTA, A. (2023): How regional economic structure matters in the era of Covid-19: resilience capacity of U.S. states *The Annals of Regional Science* 70 (1): 159–185. https://doi.org/10.1007/s00168-022-01134-w

KITSOS, T.–GRABNER, S. M.–CARRASCAL-INCERA, A. (2023): Industrial embeddedness and regional economic resistance in Europe *Economic Geography* 99 (3): 227–252. https://doi.org/10.1080/00130095.2023.2174514

KRAWCZYK, D.–MARTYNETS, V.–OPANASIUK, Y.–REKUNENKO, I. (2023): Socio-economic development of European countries in times of crisis: ups and downs *Sustainability* 15 (20): 14820. https://doi.org/10.3390/su152014820

KUMAR, S.–GORODNICHENKO, Y.–COIBION, O. (2023): The effect of macroeconomic uncertainty on firm decisions *Econometrica* 91 (4): 1297–1332. https://doi.org/10.3982/ECTA21004

LANGER, M.–SCHMUDE, J. (2023): Covid-19 as a paradigm shift? Insights from the degrowth debate in tourism *European Journal of Tourism Research* 36: 3607. https://doi.org/10.54055/ejtr.v36i.3137

LEE, J. (2023). Social vulnerability and local economic outcomes during the Covid-19 pandemic. *Regional Studies, Regional Science*, 10(1), 845–869. https://doi.org/10.1080/21681376.2023.2274097

Lee, K.-K. (2024): Korea's economic growth and the growth model in the changing global economy *The Japanese Political Economy* 50 (2): 161–184. https://doi.org/10.1080/2329194X.2024.2325981

LEÓN-CRUZ, J. F.–ROMERO, D.–RODRÍGUEZ-GARCÍA, H. I. (2024): Spatial and temporal changes in social vulnerability to natural hazards in Mexico. *ISPRS International Journal of Geo-Information* 13 (3): 66. https://doi.org/10.3390/ijgi13030066

MARTÍNEZ RODRÍGUEZ, M. C.–MORENO, C. N. (2023): Perspectives for resilience, social inclusion, and sustainable tourism in Mexico. In: AGUILAR-RIVERA, N.–BORSARI, B.–R. B. DE BRITO, P.–ANDRADE GUERRA, B. (eds.): *SDGs in the Americas and Caribbean region. Implementing the UN sustainable development goals – regional perspectives* pp. 49–74., Springer, Cham. https://doi.org/10.1007/978-3-031-16017-2_49

MEDINA-RIVAS, C. M.–RODRÍGUEZ-TAPIA, L.–MORALES-NOVELO, J. A.–REVOLLO-FERNÁNDEZ, D. A. (2022): Spatial inequality of domestic water consumption in Mexico city *Water Resources and Economics* 40: 100210. https://doi.org/10.1016/j.wre.2022.100210

MENDOZA-VELÁZQUEZ, A.–GERMAN-SOTO, V.–MONFORT, M.–ORDÓÑEZ, J. (2020): Club convergence and inter-regional inequality in Mexico, 1940–2015 *Applied Economics* 52 (6): 598–608. https://doi.org/10.1080/00036846.2019.1659491

MONROY-GÓMEZ, L.–VILLAGÓMEZ-ORNELAS, P. (2024): Stratification economics in the land of persistent inequalities *The American Journal of Economics and Sociology* 83 (1): 157–175. https://doi.org/10.1111/ajes.12527

MONTERRUBIO, C. (2022): The informal tourism economy, Covid-19 and socioeconomic vulnerability in Mexico *Journal of Policy Research in Tourism, Leisure and Events* 14 (1): 20–34. https://doi.org/10.1080/19407963.2021.2017726

MOUDUD-UL-HUQ, S.–AKTER, R.–BISWAS, T. (2020): Impact of financial crisis on credit risk: pre- and post-financial crises in an emerging economy *FIIB Business Review* 9 (2): 118–132. https://doi.org/10.1177/2319714520923952

NGO, N.-Q.-N. (2024): How do monetary policy and FDI impell economic recovery? – international evidence *SSRN Electronic Journal* https://doi.org/10.2139/ssrn.4969078

PRECIADO, V. H. T.–TORRERO, N. O. M. (2021): Economic and crime cycles synchronization across states in México: a dynamic factor model approach *Regional Statistics* 11 (4): 3–27. https://doi.org/10.15196/RS110401

SAKYI-NYARKO, C.–AHMAD, A. H.–GREEN, C. J. (2022): The gender-differential effect of financial inclusion on household financial resilience *The Journal of Development Studies* 58 (4): 692–712. https://doi.org/10.1080/00220388.2021.2013467

SÁNCHEZ, A. M.–CUADRADO-ROURA, J. R. (2024): Sectoral–regional resilience and productive specialization: a comparison among the last three crises *The Annals of Regional Science* 73 (1): 61–86. https://doi.org/10.1007/s00168-024-01261-6

THAKUR, D.–DAS, I. (2022): Statistical assessment of spatio-temporal impact of Covid-19 lockdown on air pollution using different modelling approaches in India, 2019–2020 *Regional Statistics* 12 (3): 54–84. https://doi.org/10.15196/RS120303

TRIPPL, M.–FASTENRATH, S.–ISAKSEN, A. (2024): Rethinking regional economic resilience: preconditions and processes shaping transformative resilience *European Urban and Regional Studies* 31 (2): 101–115. https://doi.org/10.1177/09697764231172326

WANG, X.–XU, S.–WANG, D. (2023): Analysis of regional resilience network from the perspective of relational and dynamic equilibrium *Journal of Cleaner Production* 425: 138859. https://doi.org/10.1016/j.jclepro.2023.138859

ZHANG, X.–TIAN, C. (2024): Measurement and influencing factors of regional economic resilience in China *Sustainability* 16 (8): 3338. https://doi.org/10.3390/su16083338

INTERNET SOURCES

ANURADHA, A.–CHOWDHURY, P. R.–BISWAS, A. (2024): Post-pandemic recovery measures to ensure socio economic sustainability through FDI in smart tourism *European Economic Letters (EEL)* 14 (1 SE-Articles): 1498–1509. https://www.eelet.org.uk/index.php/journal/article/view/1207 (downloaded: January 2024)

INSTITUTO MEXICANO PARA LA COMPETITIVIDAD A. C. [IMCO] (2023): *Índice de competitividad estatal* Instituto Mexicano para la Competitividad A.C. https://imco.org.mx/indices/indice-de-competitividad-estatal-2023/ (downloaded: January 2024)

MICHELI, J. (2020): FDI, regional development and structural change. The case of three states in El Bajío, Mexico *Análisis Económico* 35 (90): 199–220. https://www.redalyc.org/articulo.oa?id=41365966009 (downloaded: January 2024)