Inter-state economic and social convergence was analyzed in Mexico's 32 states during the period 2005-2019. Absolute economic convergence was tested through real income and average real labor income and average hourly real labor income. The variables of life expectancy, survival rate, child survival rate and educational level of workers were used to test for absolute social convergence. In addition, using the spatial dependence technique, Moran's I, it was found that economic variables related to labor income present a high level of spatial correlation. In contrast, only social variables such as life expectancy and the educational level of workers are spatially correlated.

The neoclassical model establishes convergence as an approximation to each economy's long-term equilibrium per capita income level. Suppose similar long-term equilibria exist for different economies. In that case, this will imply convergence in their per capita income (albeit at different speeds, depending on their distance from such a long-term equilibrium).

Therefore, it should be emphasized that, within the Solow-Swan model, each economy is converged with its own long-term equilibrium. Convergence has been analyzed in a large number of studies at different levels: international (Barro and Sala-i-Martin, 1992, 1997; Mankiw et al., 1992; Quah, 1996), regional (Bivand and Brunstad, 2005; López-Bazo et al., 1999) and recently locally, but also with a social focus beyond the merely economic (Royuela and Artís, 2006; Royuela and García, 2015). The above is interesting because, although neo-classical models indeed seek to solve economic growth, this change in approach allows for a validation of the theory but is now applied to economic development problems on a regional level (Barro, 2016).

Kenny (2005) pointed out that improving GDP will help increase economic development indicators. Increasing their income may improve their quality of life, regardless of their income level.

The list of social indicators analyzed to test convergence is long and includes factors such as life expectancy, infant mortality, level of education, literacy, and environmental degradation, among others (Becker et al., 2005; Becker et al., 2005; Bourguignon and Morrisson, 2002a, 2002b; Goesling and Firebaugh, 2004; Neumayer, 2003). The results lead to mixed conclusions regarding convergence, depending on the time frame and the selection of countries and indicators.

In this research, the hypothesis of economic and social convergence on a regional level was analyzed for the 32 states of Mexico during 2005-2019. This is carried out by including growth variables such as average real labor income (or average real wage), average real hourly labor income (or average real hourly wage), life expectancy, survival rate, infant survival rate, and level of education of workers. While it is true that there is extensive literature analyzing convergence for Mexico, studies on social convergence on a regional level are scarce.

In addition to updating the database and creating three variables (average real labor income (or average real wage), average real hourly labor income (or average real hourly wage), and level of education of workers), an attempt was made to find robust results regarding the existence of absolute convergence in socioeconomic variables, using a wide range of techniques available in this analysis. Likewise, issues of spatial distribution, particularly on a regional level, have attracted particular attention to spatial statistical methods and spatial econometrics.

This paper analyzes social convergence in an unprecedented way for Mexico using various socioeconomic variables. The findings are consistent with case studies for developing economies (Royuela and Garcia 2015) and suggest the existence of absolute convergence in economic variables (average real labor income and average real hourly labor income (or average real hourly wage)) and social variables. It was also found that the spatial correlation presented by the economic variables is high, particularly those related to labor income. In contrast, only some social variables are spatially correlated: life expectancy and workers with more than a high school diploma.
2. REVISION OF LITERATURE

The starting point of neoclassical theories on economic growth is the article published by Solow (1956). Despite its numerous criticisms and amendments (Lengyel, 2004; Mankiw et al., 1992; Romer, 1986), the theory describes an existing phenomenon and one that is still widely used as a framework for convergence and territorial catch-up analyses (Bucur and Stangaciu, 2015; Konya and Guisan, 2008; Viegas and Antunes, 2013; Vojinović et al., 2009).

The pioneering authors include Barro and Sala-i-Martin (1992), whose research assumes the existence of a long-term steady-state equilibrium in the different economies and, simultaneously, a convergence process. The empirical strategy consists of regressing the growth rate in relation to a broad set of variables, considering how each affects both the convergence process and the long-term steady state.

This theory was questioned regarding the relevance of the forecasts and prescriptions derived from the neoclassical model. Since the second half of the 1980s, various efforts to formalize growth models have gained momentum. Since the papers by Romer (1986) and Lucas (1988), some assumptions, hypotheses, and methodologies different from those previously considered began to be put forward. Academic contributions multiplied, including many others, in addition to the later research by Romer and Lucas, Rebelo, Barro, Aghion and Howitt, and Grossan and Helpman (Arrous, 1999; Guellec and Raile, 1995; Sala-i-Martin, 1994).

The most cited and recognized works regarding the combination of theory and empirical testing include those of Mankiw et al. (1992), Barro and Sala-i-Martin (1992), Aghion and Howitt (1992), Barro and Sala-i-Martin (1995), De la Fuente Moreno (1995) and Durlauf and Quah (1998), who discuss some of the main implications of growth models concerning the existence of convergence or divergence between different economies.

Along the same lines, Barro (1996) conducted a convergence study among 100 countries and concluded that there is convergence when considering the level of education, life expectancy, low birth rates, and little government intervention. Rodríguez-Pose (1998) points out that each region depends on its geographic situation and that economies rely not so much on their foundations as on their capacity to respond and adapt to social and economic changes.

Previous works focus solely on economic convergence (GDP and income), and it is only recently that studies have begun to emerge that consider convergence in relation to social and not merely economic indicators (Barro, 2016; Royuela and García, 2015; Egri and Tánctzos, 2018; Lafuente et al., 2020).

Research on the subject includes that of Kuc-Czerep (2017) with a regional analysis of Nordic countries, which finds that there is no social convergence in living standards in most of the territory. Meanwhile, Egri and Tánctzos (2018) confirm the absolute convergence theory for social convergence and catch-up effects of economic and social performance in Central and Eastern European regions, while Rodríguez-Pose and Tselios (2015) point to a negative relationship between the growth rate of social welfare and the initial welfare level.

Even fewer studies have been conducted in Latin America. However, an analysis applied to the regions of Colombia stands out, where convergence is identified, as well as the fact that spatial autocorrelation reinforces convergence processes through the deepening of social factors (Royuela and García, 2015).

Most existing studies concerning Mexico find evidence of some convergence among the country's regions. Relevant works on the subject include those of Díaz-Bautista (2003, 2000); Esquivel (1999); Messmacher (2000). The results do not show a consensus in economic convergence despite the diverse number of economic growth studies conducted for different economies. Still, more importantly, there are no empirical studies that analyze social convergence in Mexico.

3. METHODOLOGY

Data

According to the precepts of Sen (1999), an annual panel of regional data was used for the period 2005-2019 for the 32 states of Mexico. This database is constructed using information from various statistical sources. For the calculation of GDP per capita, GDP data was used from the National Institute of Statistics and Geography (INEGI), which are at constant 2013 values, as well as the population by state from the "Demographic Conciliation of Mexico, 1950-2015" and the "Population Projections for Mexico and the Federal States, 2016-2050" of the National Population Council (CONAPO). The average real wage and the average real hourly wage were calculated using the National Occupation and Employment Survey (ENOE), adding up the wages of the workers and then dividing by the number of workers in each quarter for all the surveys for the years in question, after which the annual average was calculated.

Life expectancy was obtained from CONAPO estimates based on the "Demographic conciliation, 1990-2010" and "Population projections, 2010-2030". The survival rate was calculated based on information from the Executive Secretariat of the National Public Security System (SESNSP). The infant survival rate was calculated based on information on births and infant mortality published by INEGI.

Finally, the percentage of workers with more than a high school diploma was calculated using the ENOE, adding up the total number of workers with more than a high school diploma and dividing by the total number of workers each quarter for all surveys in the years in question. The annual average was then obtained. Table 1 shows the descriptive statistics of the socioeconomic variables.
Meanwhile, Figure 1 shows the evolution of the average of each of the socioeconomic variables of interest. Figure 1 establishes that from 2005 to 2019, the wage or labor income decreased, as did the survival rate. While life expectancy increased, the infant survival rate and the percentage of workers with more than a high school diploma were maintained over time.

Figure 1. Evolution of socioeconomic variables

This paper aims to test whether absolute social convergence exists using a fixed effects panel model proposed by Cuadrado-Roura et al. (1999) and Raymond and Garcia (1994). The advantage of this model is that it permits the exploitation of both the individual and temporal dimensions of the data, providing a more exhaustive analysis. The model proposed by the authors is based on the following beta convergence equation:

$$\Delta \ln y_{it} = \ln y_{it} - \ln y_{it-1} = a_i - \beta \ln y_{i,t-1} + \gamma_t + \epsilon_{it}$$  \hspace{1cm} (1)$$

Where $\ln y_{it}$ represents the economic or social variable of interest for state $i$ at time $t$, $a_i$ is the specific individual effect that can capture other components that affect these variables and are common to the states. $\ln y_{i,t-1}$ is the variable of interest from the previous year, $\gamma_t$ is the specific time effect, and $\epsilon_{it}$ is the random effect that captures the influence of some omitted variables (Cuadrado-Roura et al., 1999; Raymond and Garcia, 1994).
The convergence equation was transformed by calculating the weighted mean for all $i$ at each $t$:

$$\Delta \ln y_i = \bar{y} - \beta \ln y_{i,t-1} + \gamma_t + \epsilon_{i,t}$$  \hspace{1cm} (2)$$

Where the dependent variable is the variable of interest on a national level, and the explanatory variable is the level of the variable of interest in the initial period. The difference between equation (1) and equation (2) results in:

$$\Delta \ln y_i - \Delta \ln y = (a_i - \bar{a}) - \beta (\ln y_{i,t-1} - \ln y_{i,t-1}) + \epsilon_{i,t}$$  \hspace{1cm} (3)$$

Where $(a_i - \bar{a})$ is the autonomous component of the state growth rate of differences concerning the national average and $\epsilon_{i,t} = \epsilon_i - \epsilon_t$ the new random disturbance element.

It can also be expressed as follows:

$$\Delta \ln y_i - \Delta \ln y = (a_i - \bar{a}) + (1 - \beta)[\ln y_{i,t-1} - \ln y_{i,t-1}] + \epsilon_{i,t}$$  \hspace{1cm} (4)$$

Thus, economic and social convergence was estimated using the following equation:

$$\ln \left(\frac{y_i}{y_t}\right) = a + \beta \ln y_{i,t-1} + \mu_i + \eta_t + \epsilon_{i,t}$$

$$\ln y_i - \ln y_{i,t-1} = a + \beta \ln y_{i,t-1} + \mu_i + \eta_t + \epsilon_{i,t}$$

$$\Delta \ln y_i = a + \beta \ln y_{i,t-1} + \mu_i + \eta_t + \epsilon_{i,t}$$  \hspace{1cm} (5)$$

Where $y$ is the socioeconomic variable of interest, $\mu_i$ are the individual fixed effects (by federal state), $\eta_t$ are the temporal fixed effects, and $a$ and $\epsilon_{i,t}$ are the constant and the error term, respectively.

A panel data regression equation with fixed effects is established as the basis of model (5) so that, in this case, $\mu_i$ captures the structural differences between states using $\mu - 1$ additional independent terms. Furthermore, the Hausman and Wald F statistical tests indicate that fixed and temporal effects should be estimated, except for the GDP per capita, which should be calculated with random effects and temporal effects, as shown in Table A1.

Following the Wooldridge and Wald tests, auto-correlation and heteroscedasticity problems were detected, which should be solved using Generalized Least Squares (GLS) estimators or Panel Corrected Standard Errors (PCSE) in the case of fixed effects. In this case, both methods were estimated.

Once the $\beta$ coefficient was obtained, the convergence speed was calculated as follows:

$$r = \frac{\ln(1 - \beta)}{\bar{y}}$$  \hspace{1cm} (6)$$

In addition, a spatial analysis was performed using global and local spatial indicators. The international and local Moran’s I was calculated for each variable of interest and, based on the latter, the LISA maps for the years 2005 and 2018 were presented, as prepared by Arbia et al. (2005a, 2005b); Arbia and Piras (2005); Elhorst (2001a, 2001b, 2003, 2005).

The inclusion of spatial relationships in the construction of taxonomic measures of development is justified since regions are not isolated in space and may be affected by other units, which is why the presence of spatial autocorrelation was tested using Moran’s I-statistics.

Exploratory spatial data analysis was used because events recorded in a specific geographic location affect their direct neighbors and others. For this purpose, specialized literature has considered the use of spatial data analysis as an appropriate tool for processing spatial data, and the use of spatial data analysis as a proper tool for processing, describing, and analyzing Moran’s I information (Moran, 1948). The parameters of this indicator establish that the higher the value of this statistic, the stronger the degree of spatial dependence in the variable, and vice versa. In other words, to what extent does a change in the levels of a variable in one location in space affect the levels of that variable in neighboring areas.

The formula for calculating the overall Moran’s I is as follows:

$$I_i = \frac{N}{\sum_{k=1}^{N} \omega_{ik}} \sum_{j=1}^{N} \omega_{ij} \left(\frac{\bar{y}_j - \ln y_j}{\bar{y} - \ln y_i}\right) \left(\frac{\bar{y}_i - \ln y_j}{\bar{y} - \ln y_i}\right)$$  \hspace{1cm} (7)$$

Where $N$ is the total number of states indexed by $i$ and $j$, and $\ln y_i$ is the variable of interest. $\bar{y}_j$ is the mean of $\ln y_j$ and $\omega_{ij}$ is each element of the spatial weight matrix, which in this case was assumed to be a queen-type spatial weight matrix. Meanwhile, the calculation of the local Moran’s I is as follows:

$$I_i = \frac{\sum_{j=1}^{N} \omega_{ij} \left(\ln y_j - \ln y_i\right)}{\sum_{k=1}^{N} \sum_{j=1}^{N} \omega_{ij}} \left(\frac{\bar{y}_j - \ln y_j}{\bar{y} - \ln y_i}\right)$$  \hspace{1cm} (8)$$

4. RESULTS

Table 2 provides the main results of the estimations carried out in which the various estimates of the panel of fixed and temporal effects using the PCSE and GLS estimation methods can be observed.
According to the results shown in Table 2, convergence was found in GDP per capita, as well as in socioeconomic variables such as labor income, hourly labor income, life expectancy, survival rate, infant survival rate, and workers with more than a high school degree, since it is observed that the variables are significant and with a negative sign, i.e., the socioeconomic differences between the states are becoming smaller and smaller.

<table>
<thead>
<tr>
<th>Table 2. Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GDP per capita</strong></td>
</tr>
<tr>
<td>PCE Model</td>
</tr>
<tr>
<td>( \log_{(+1)} )</td>
</tr>
<tr>
<td>Gh2</td>
</tr>
<tr>
<td>Convergence speed</td>
</tr>
<tr>
<td>Q5 Model</td>
</tr>
<tr>
<td>( \log_{(+1)} )</td>
</tr>
<tr>
<td>Gh2</td>
</tr>
<tr>
<td>Convergence speed</td>
</tr>
<tr>
<td>Notes: standard error values in parentheses. Significance at *** 1%, ** 5% and * 10% levels. Source: compiled by authors.</td>
</tr>
</tbody>
</table>

As for the speed of convergence, the higher the percentage, the greater the speed of convergence, which implies that socioeconomic differences between the states are decreasing for the period in question.

Furthermore, it is observed that the speed of convergence of socioeconomic variables is greater than that of GDP per capita, which leads to the assumption that the convergence process can be considered a national phenomenon, probably based on the general economic growth of the country, which has been slow (Calderon and Sanchez, 2012) around 2.2% (Moreno-Brid, 2013).

As for education, the percentage of workers with more than a high school diploma was analyzed. This variable, in particular, is a clear example of convergence, and it was found that it was increasing at a much higher speed than other specifications of the convergence panel; i.e., the direction of convergence of education is a slightly worse stationary state, which, to some extent, may be explained by the influence of neighboring states. Therefore, states with a higher percentage decrease and those with a lower percentage increase. In this respect, authors such as Ros (2008) point out that the slow process of human capital formation is not responsible for the slowdown in growth. Instead, the fact that human capital training has not been faster can be attributed in part to the slowdown in growth. Levy (2007) mentions that the PROGRESA-Oportunidades social program was aimed at families living in extreme poverty, leading to the poorest states coming ever closer to the plateau of the national average.

One possible explanation is that implementing social programs such as PROGRESA may have provided social services that permitted absolute convergence in life expectancy, survival rate, and infant survival rate. In this respect, PROGRESA was the first conditional cash transfer program linked to the receipt of social benefits with investment in human capital, health, and education. It consisted of offering money to low-income families. They would only receive money if their children regularly attended school, the family sought health services, and the children studied a higher grade (Banerjee and Duflo, 2012). The goal was to make not sending children to school costly, regardless of how parents felt about education.

The pilot program showed that PROGRESA increased school enrollment, especially at the secondary level, with girls rising from 67% to 75% and boys from 73% to 77% (Schultz, 2004). However, it has been questioned whether this type of public intervention is feasible (Easterly, 2002).

The purpose of this paper is not to analyze the PROGRESA social program, although the fact (as already mentioned by Banerjee and Duflo (2012), Schultz (2004), Attanasio et al. (2012), among others) that the program served as a determinant so that social variables on a sub-national level would converge needs to be underlined.

The traditional neoclassical growth model assumes that each regional economy is geographically independent, without considering the possibility of spatial interactions (Tselios, 2009). Using spatial econometric techniques, social variables were introduced into the convergence model by exploring substantive and modest spatial dependence.

Table 3 presents Moran’s I for each variable of interest for each year.

Table 3 shows a spatial correlation in GDP per capita, labor income, and hourly labor income during the entire period of analysis. Meanwhile, the infant survival rate shows no spatial correlation during the whole period of the study. Concerning the life expectancy variable, it only begins to show spatial correlation at the end of the period. The survival rate is spatially correlated mainly at the beginning of the analyzed period. Finally, the percentage of workers with more than a high school diploma is spatially correlated at the beginning and end of the analyzed period.
It should be noted that the variables related to labor income present a high spatial correlation, which increases throughout the analyzed period. In other words, the influence of neighboring states may be growing.

Regarding the local Moran’s I, Figure 2 illustrates the local Moran’s I for 2005 and 2018 for the socioeconomic variables of interest.

**Figure 2. Local Moran’s I**
5. CONCLUSIONS

This paper analyzed social convergence in Mexico, considering not only the traditional variable of GDP per capita but also labor income and social indicators such as life expectancy, survival rate, infant survival rate, and the percentage of workers with more than a high school education. Furthermore, it is shown that the convergence speed of these socioeconomic variables is greater than that of GDP per capita. The spatial distribution of the variables was also analyzed by inspecting autocorrelation statistics such as Moran's I. The spatial distribution of the variables was also studied.

The results confirm the theory of absolute social and economic convergence. Although the panel estimates show absolute convergence, they are an expression of short-term convergence with the constant of each state region. Convergence in economic and social variables may have resulted from the PROGRESA conditional monetary transfer program implemented between 1994 and 2000.

Sen (1999) points out that health and education are among the basic capabilities that give value to human life and contribute to social and economic prosperity. From this perspective, policies, activities, and institutions are needed to support better education and health services among the population, especially among the poorest sectors of the population.

Talking about economic growth is not unrelated to talking about development issues, and this is indicated in new studies on social convergence, which is why we need to bear in mind the fact that social programs, such as PROGRESA, a program focused on a specific social group, will not improve...
economic development on their own; other initiatives that form part of an integral and coherent strategy, as well as a context of sustained economic growth, will be necessary (Levy and Rodriguez, 2005).

The study’s findings suggest that the Mexican government should focus its policies on increasing nominal wages and public investment to improve the country’s access to quality education and health services. The to-do list is long, but hope is required to create a good economy.

**ANNEX**

<table>
<thead>
<tr>
<th>Table A1. Statistical tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>GDP per capita</strong></td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>Breusch-Pagan test (multiplier for random effects)</td>
</tr>
<tr>
<td>P-value</td>
</tr>
<tr>
<td>Hausman test (fixed effects vs. random effects)</td>
</tr>
<tr>
<td>P-value</td>
</tr>
<tr>
<td>Individual effects test (Wild F-statistics)</td>
</tr>
<tr>
<td>P-value</td>
</tr>
<tr>
<td>Temporaries effects test (Wild F-statistics)</td>
</tr>
<tr>
<td>P-value</td>
</tr>
<tr>
<td>Wooldridge Test (Auto-correlation)</td>
</tr>
<tr>
<td>P-value</td>
</tr>
<tr>
<td>Modified Wild test (Heteroscedasticity)</td>
</tr>
<tr>
<td>P-value</td>
</tr>
<tr>
<td>Pesaran test (LM conditional)</td>
</tr>
<tr>
<td>P-value</td>
</tr>
</tbody>
</table>

**BIBLIOGRAPHY**


According to Sen (1999), a "good life" is composed of four key elements: material well-being, health and survival, education and personal development, and social inclusion/participation.