MAKÜFEBED

ISSN Online: 1309-2243
*http://dergipark.org.tr/makufebed*
https://doi.org/10.29048/makufebed.744870

Mehmet Akif Ersoy Üniversitesi Fen Bilimleri Enstitüsü Dergisi 11(Ek Sayı 1): 346-363 (2020)
*The Journal of Graduate School of Natural and Applied Sciences of Mehmet Akif Ersoy University 11(Supplementary Issue 1): 346-363 (2020)*

**Araştırma Makalesi / Research Paper**

# Investigation of the Factors Affecting the Homicide Counts in the USA by Quantile Regression

Oğuzhan DEMİREL[1]

[1]Hacettepe University, Faculty of Science, Ankara, Turkey

*Geliş Tarihi (Received): 29.05.2020, Kabul Tarihi (Accepted): 11.12.2020*
*Sorumlu Yazar (Corresponding author): oguzhandemirel@hacettepe.edu.tr*
*+90 539 5487003*

## ABSTRACT

This study has conducted in the United States of America (USA) in the 1960s, 1970s, 1980s, and 1990s to determine whether divorce rates, unemployment rates, and the population had an impact on homicide numbers. Initially, all variables were examined and interpreted geographically on the map by districts. Subsequently, the stationary circumstance of variables has been tested with the Augmented Dickey-Fuller (ADF) Test, which is one of the unit root tests. After it has found that the variables did not need to stabilize, regression analysis has performed by the Least Squares (LS) method. Quantile regression, which is an alternative method to the LS method, has been used since all the resulting models do not have a normal distribution. These models have been created with 3 diverse quantile values for each period. Among these models, the ones with the highest correlation coefficient are the models having the 0.75 quantile value. Therefore, the results have been obtained from models with the 0.75 quantile value. Hence, for the homicide counts in the USA, those have found that the country population had a positive effect in the 1960s, the country population and the divorce rates had positive effects in the 1970s, the country population had a positive effect in the 1980s, and the country population and the unemployment rates had positive effects in the 1990s. Furthermore, the unemployment rates in the 1970s and 1980s had a negative effect on the homicide counts in the USA.

**Keywords:** Homicide counts, geoda, least squares method, quantile regression, the USA homicide

# ABD'de Cinayet Sayısını Etkileyen Faktörlerin Kantil Regresyon ile İncelenmesi

## ÖZ

Bu çalışma, 1960, 1970, 1980 ve 1990'larda ABD'deki boşanma oranlarının, işsizlik oranlarının ve ülke nüfusunun cinayet sayılarını etkileyip etkilemediğini belirlemek amacıyla yapılmıştır. İlk olarak, tüm değişkenler harita üzerinde bölgeler tarafından coğrafi olarak incelenmiş ve yorumlanmıştır. Ardından değişkenlerin durağanlık durumu, birim kök testlerinden biri olan ADF testi ile test edilmiştir. Değişkenlerin durağanlaştırılmasına gerek olmadığı tespit edildikten sonra, en küçük kareler (EKK) yöntemi ile regresyon analizi yapılmıştır. Elde edilen tüm modellerin normal dağılım göstermemesi nedeniyle, EKK yöntemine alternatif bir yöntem olan kantil regresyon kullanılmıştır. Oluşturulan bu modellerde, her bir dönem için 3 farklı kantil değeri kullanılmıştır. Bu modeller arasında 0,75 kantil değerine sahip modeller, en yüksek korelasyon katsayına sahiptir. Bu nedenle 0,75 kantil değerine sahip modeller kullanılarak sonuçlar elde edilmiştir. Elde edilen bu sonuçlara göre, ABD'deki cinayet sayıları için; 1960'larda yalnızca ülke nüfusunun olumlu bir etkisi olduğu, 1970'lerde ülke nüfusunun ve boşanma oranlarının olumlu etkileri olduğu, 1980'lerde ülke nüfusunun olumlu bir etkisi olduğu ve son olarak ülke nüfusunun ve işsizlik oranlarının 1990'larda olumlu etkilerinin olduğu bulunmuştur. Ayrıca, işsizlik oranlarının 1970 ve 1980'lerde cinayet sayıları üzerinde olumsuz bir etkisinin olduğu sonucu da elde edilmiştir.

Oğuzhan DEMİREL, https://orcid.org/0000-0003-4352-6531

**Anahtar Kelimeler:** Cinayet sayıları, geoda, en küçük kareler yöntemi, kantil regresyon, ABD cinayet

## INTRODUCTION

Since the existence of humanity, many needs have emerged. These needs have evolved over time and sometimes changed. As the world's population naturally grows constantly, these needs are shaped and changed. People started to work after a certain time to earn money and live on. Besides, people felt the need to marry and multiply by naturally. But as the years progressed, these needs were replaced by problems. Nowadays, since the intolerance of people increases, marriages last for a short time. This led to an increase in divorce rates. Since this intolerance increases not only among spouses but also among all people in the world, there is a significant increase in the number of murders. In addition, the world's population is increased rapidly. But this increasing population has caused some problems and existing problems grew in time. Despite the law and legal regulations, homicide and security have become one of the biggest problems of many countries.

### Literature Review

Kposowa et al. (1995), in this study with a large set of the USA counties, measures for the subculture of violence theory, economic deprivation, economic inequality, social integration, and other structural variables were tested on the property and violent crime indices and homicide rates. Support was found for economic deprivation in the case of homicide and social integration across every dependent variable. Urbanity was the main determinant of property crime, urbanity, and population density were important factors in violent crime, and poverty, divorce, and density figure strongly in homicide. Poverty and divorce were continued to be the strongest determinants of homicide in rural counties, while population mobility and urbanity were the strongest factors in both rural violence and property crime. Unemployment also played a strong role in rural property crime.

Sen et al. (2012), the purpose of this study is to explore whether, in the USA, there are associations between state-level variations in mortality among young children and state abortion restriction policies - such as parental-consent requirements, parental-notification requirements, mandatory delay laws, and restrictions on Medicaid funding for abortion. To investigated this, were used the National Center for Health Statistics (NCHS) multiple cause of deaths public-use data files for the period 1983-2002 and were compiled data on children ages 0-4 identified as having died as a result of assault/homicide in each state and year. Medicaid funding of abortion, mandatory delay laws, and parental involvement laws for minors seeking abortions were included as the main predictor variables of interest. Multivariate count data models using pooled state-year-age cohort data, with state and time fixed effects and other state-level controls, were estimated. In the count data models, parental-consent laws were associated with a 13% increase in child homicide deaths; parental-notification laws were associated with an 8% increase in child homicide deaths though the results were less robust to alternative model specifications; mandatory delay requirements were associated with a 13% increase in child homicide deaths. While these data do not allow to discern precise pathways via which state abortion-restrictions can lead to more child homicide deaths, were speculated that state restrictions on abortion may result in a disproportionate increase in children born into relatively high-risk environments.

Ousey and Kubrin (2014), in the current study, were addressed this issue by investigating whether within-city changes in immigration are related to temporal variations in rates of overall and circumstance-specific homicide for a sample of large USA cities during the period between 1980 and 2010. Fixed-effects negative binomial and two-stage least squares instrumental variable regression models were used to analyze data from 156 large USA cities observed during the 1980-2010 period. Findings from the analyses suggest that temporal change in overall homicide and drug homicide rates were significantly related to changes in immigration. Specifically, increases in immigration were associated with declining rates for each of the preceding outcome measures. Moreover, for several of the homicide types, findings suggest that the effects of changes in immigration vary across places, with the largest negative associations appearing in cities that had relatively high initial (i.e., 1970) immigration levels.

Humphreys et al. (2017), the aim of this study to estimate the impact of Florida's stand your ground law on rates of homicide and homicide by firearm. Using an interrupted time series design were analyzed monthly rates of homicide and homicide by firearm in Florida between 1999 and 2014. Data were collected from the Wide-ranging Online Data for Epidemiologic Research (WONDER) web portal at the Centers for Disease Control and Prevention. Were used seasonally adjusted segmented Poisson regression models to assess whether the onset of the law was associated with changes in the underlying trends for homicide and homicide by firearm in Florida, October 1, 2005, the effective date of the law, was used to define homicides before and after the change. Prior to the stand your ground law, the mean monthly homicide rate in Florida was 0.49 deaths per

100,000 (mean monthly count, 81.93), and the rate of homicide by firearm was 0.29 deaths per 100,000 (mean monthly count, 49.06). Both rates had an underlying trend of 0.1% decrease per month. After accounting for underlying trends, these results were estimated that after the law took effect there was an abrupt and sustained increase in the monthly homicide rate of 24.4% (relative risk [RR], 1.24; 95%CI, 1.16-1.33) and in the rate of homicide by firearm of 31.6% (RR, 1.32; 95%CI, 1.21-1.44). No evidence of change was found in the analyses of comparison states for either homicide (RR, 1.06; 95%CI, 0.98-1.13) or homicide by firearm (RR, 1.08; 95%CI, 0.99-1.17). Furthermore, no changes were observed in control outcomes such as suicide (RR, 0.99; 95%CI, 0.94-1.05) and suicide by firearm (RR, 0.98; 95%CI, 0.91-1.06) in Florida between 2005 and 2014. The implementation of Florida's stand your ground self-defense law was associated with a significant increase in homicides and homicides by firearm but no changed in rates of suicide or suicide by firearm.

Sipsma et al. (2017), the aim of this study to examine whether state-level spending on social and public health services is associated with lower rates of homicide in the USA Participants were selected from all states in the USA and the district of Columbia for which data were available (n=42). Consequently, after adjusting for potential confounding variables, were found that every $10,000 increase in spending per person living in poverty was associated with 0.87 fewer homicides per 100,000 population or approximately a 16% decrease in the average homicide rate (estimate='0.87, SE=0.15, p<0.001). Furthermore, there was no significant effect in the quartile of states with the highest percentages of individuals living in poverty but significant effects in the quartiles of states with lower percentages of individuals living in poverty. Spending on social and public health services were associated with significantly lower homicide rates at the state level.

## MATERIAL AND METHODS

The aim of this study is to determine the factors that affect the number of murders in the USA periodically. The homicide data used in studies (Messner et al., 2000; Baller et al., 2001) were taken from the GeoDa Center web-site (GeoDa Data and Lab, 2003). Homicide counts (HC), unemployment rates (UE), divorce rates (DV), and country population (CO) variables were selected and a dataset was created. These variables include periodic data for the 1960s, 1970s, 1980s and 1990s. The GeoDa and the EViews 9 programs were used in all analyses and statistics. In the study, an abbreviation was used for all variables. The meanings of the abbreviations of the variables used in Table 1 are given.

**Table 1.** Description of variables

| Variable | Description |
|---|---|
| HC | Homicide counts, three-year average centered on the 1960s, 1970s, 1980s, 1990s |
| PO | County population, 1960s, 1970, 1980s, 1990s |
| UE | Unemployment rates 1960s, 1970s, 1980s, 1990s |
| DV | Divorce rates 1960s, 1970s, 1980s, 1990s (% males over 14 divorced) |

### Geographic Statistics of Variables

The values of the HC variable were divided into groups and the Quantile Map (QM) method was applied in the GeoDa program. Data for the 1960s, 1970s, 1980s and 1990s were examined separately. The density and percentages of the data analyzed on districts are visualized on the USA map.

#### *Change of Homicide Counts by Years*

The distributions of HC in the 1960s were shown in Figure 1. There is no district where the murder was not observed, and most of the murders are in group 3 (0.353: 1). The distributions of HC in the 1970s were shown in Figure 2. Again, there are no districts without murder. The distribution of most of the murders in the second group (0: 0.333) is seen. The distributions of HC in the 1980s were shown in Figure 3. In this distribution, there are 556 districts with no murders. Further, the distribution of the murders was more evenly distributed than in previous years. Finally, the distributions of HC in the 1990s were shown in Figure 4. There is no district where the murder was not observed. Also, most of the murders are distributed over the 2nd and 3rd groups. According to these 4 figures, most of the murders in the United States are observed in the Southwest, Southeast, and Eastern regions. The middle regions of the USA are calmer compared to these regions.

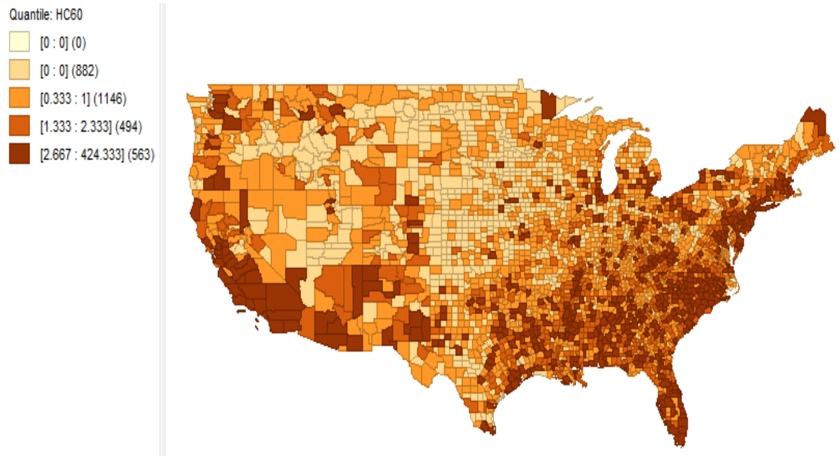


**Figure 1.** Homicide Counts in the 1960s (HC60)

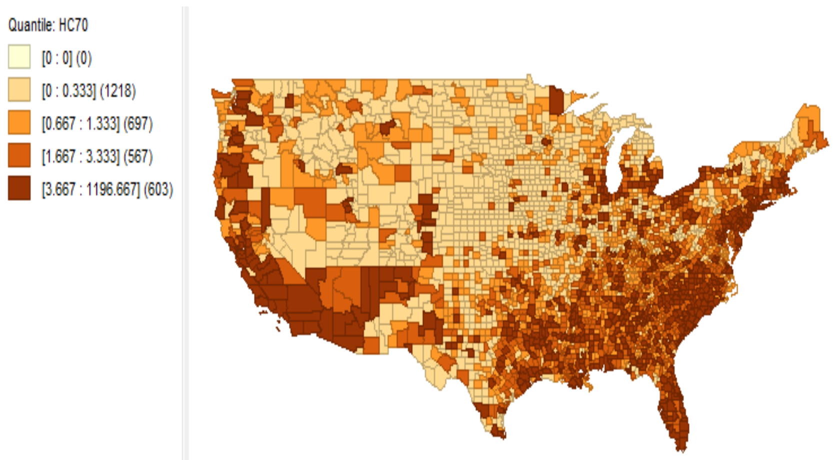


**Figure 2.** Homicide counts in the 1970s (HC70)

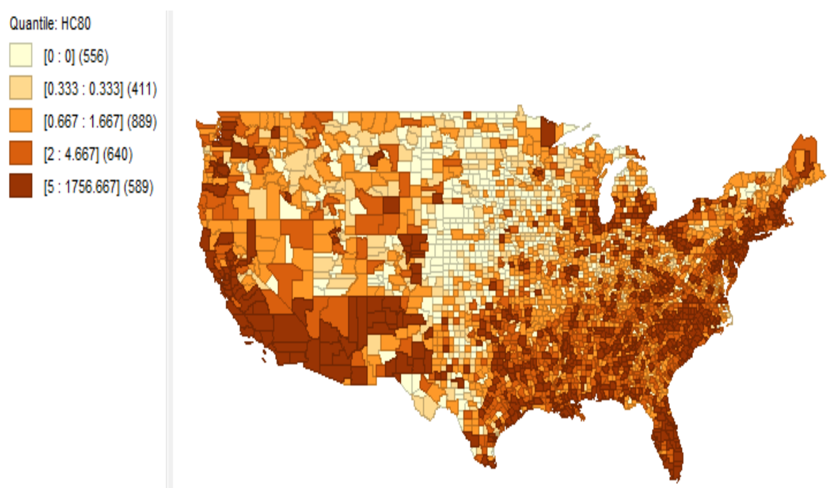


**Figure 3.** Homicide counts in the 1980s (HC80)

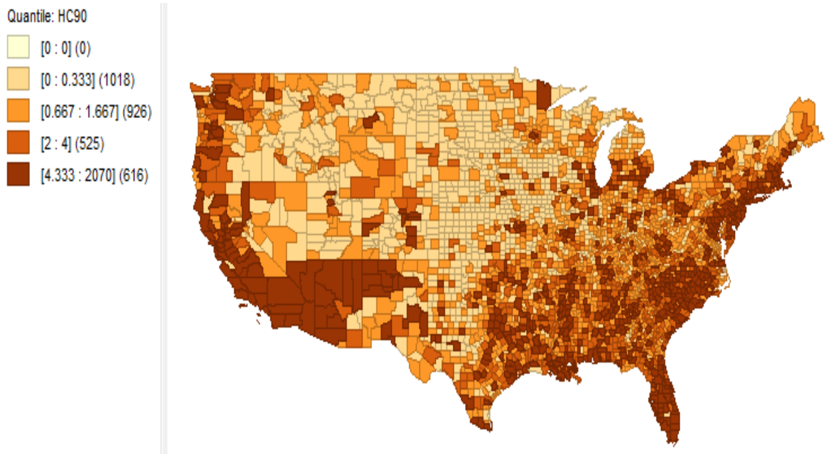


**Figure 4.** Homicide counts in 1990s (HC90)

### *Change of Country Population by Years*

The PO variable shows the distribution of the population by districts. The PO was visualized for the 1960s, 1970s, 1980s, and 1990s using the Percentile Map method. Populations, according to population density; <1%, 1%-10%, 10%-50%, 50%-90%, 90%-99% and >99%. The visuals of the distributions for the 1960s, 1970s, 1980s and 1990s are shown on the USA map in Figure 5, Figure 6, Figure 7 and Figure 8, respectively. Population density has not changed in almost any region. Although the density is the same, only the number of populations increased as the years progressed.

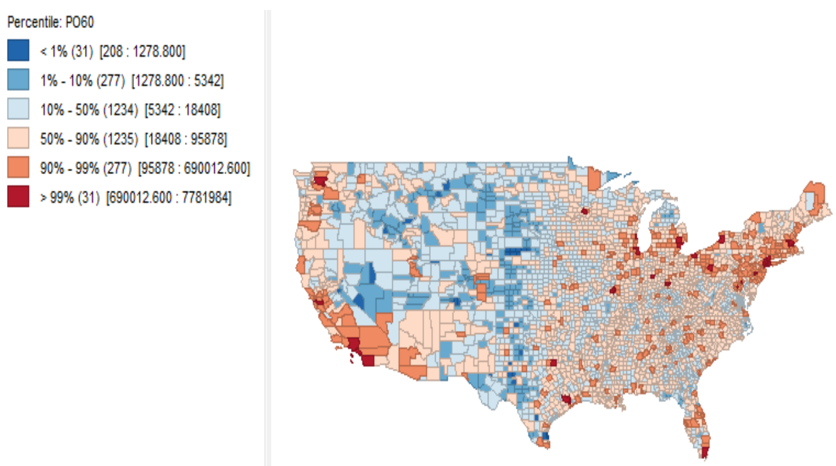


**Figure 5.** Country population in the 1960s (PO60)

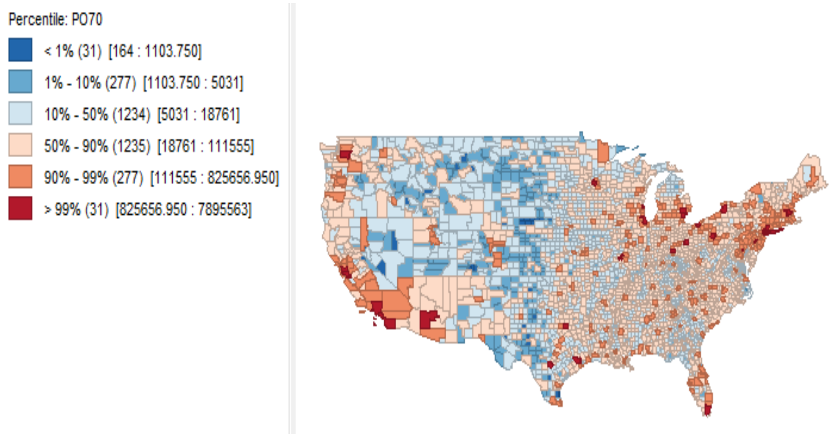


**Figure 6.** Country population in the 1970s (PO70)

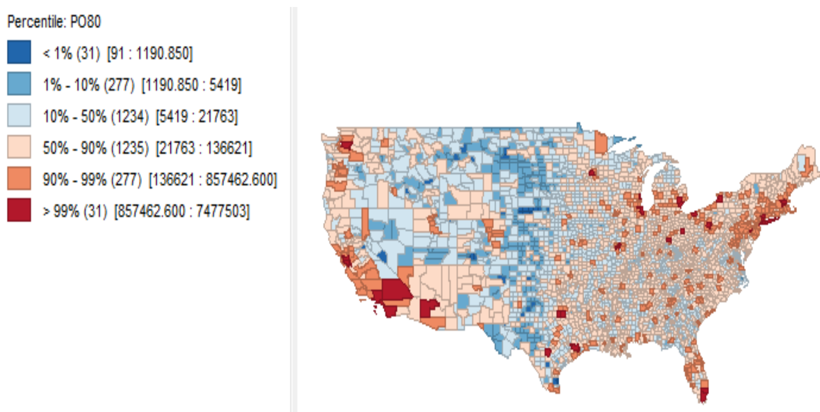


**Figure 7.** Country population in the 1980s (PO80)

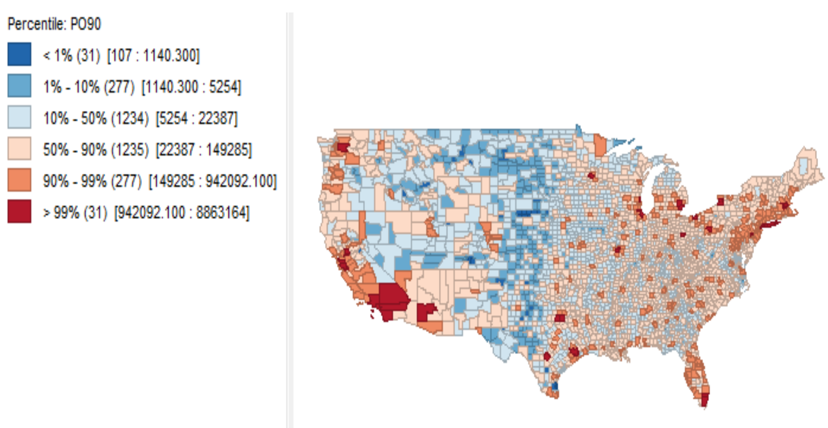


**Figure 8.** Country population in the 1990s (PO90)

### Change of Unemployment Rate by Years

Unemployment is undoubtedly one of the major problems today. The UE was used in this study considering that there is an effect on the number of murders. The UE for the 1960s, 1970s, 1980s and 1990s were classified using the QM method. It is seen that unemployment is concentrated in certain regions. In Figure 10 for the 1970s, UE increased, especially in the Western region. In Figure 12 for the 1990s, unemployment in Western regions; It is seen that it has made progress towards the

South West region. In general, it is seen that the intensity of UE decreases with years.

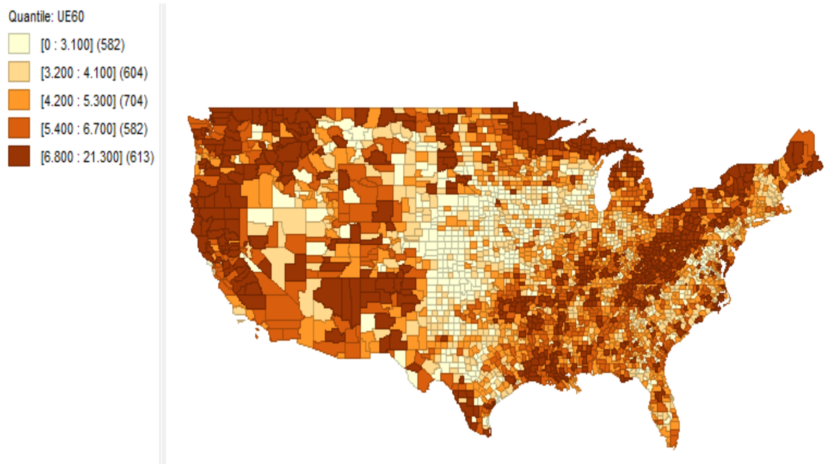


**Figure 9.** Unemployment rates in the 1960s (UE60)

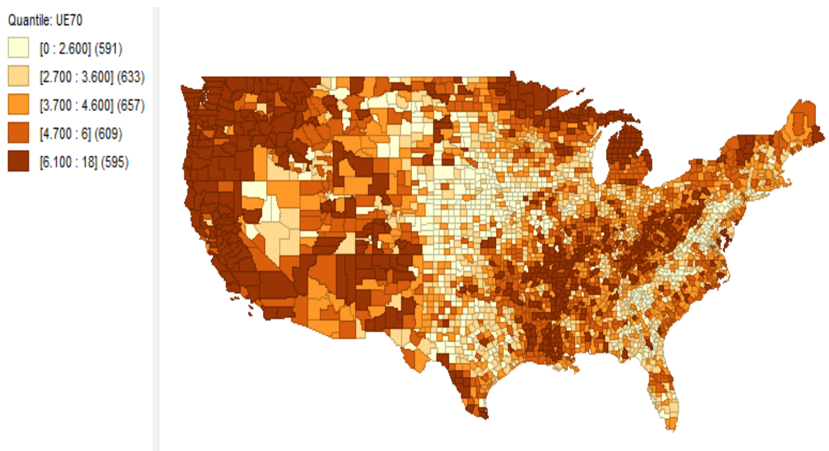


**Figure 10.** Unemployment Rates in the 1970s (UE70)

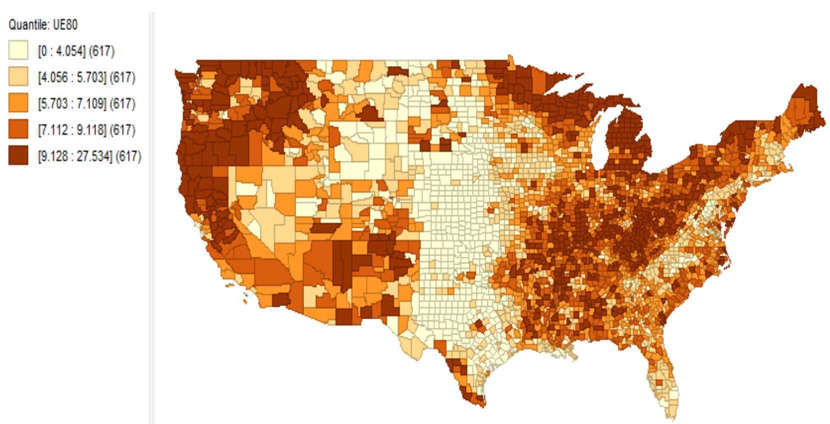


**Figure 11.** Unemployment rates in the 1980s (UE80)

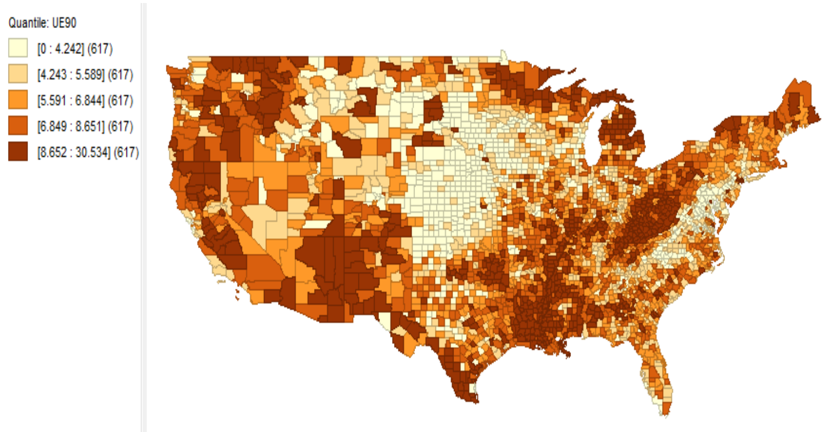


**Figure 12.** Unemployment rates in the 1990s (UE90)

### Change of Divorce Rates by Years

Rates of DV; 1960s, 1970s, 1980s, 1990s are given in the following figures. Data of these ratios were divided into 6 groups using <1%, 1%-10%, 10%-50%, 50%-90%, 90%-99% and >99% using Percentile Map (PM) method. The DV was found to be high, especially in the counties of the state of Nevada. It has been determined that the number of divorces has increased continuously by years.

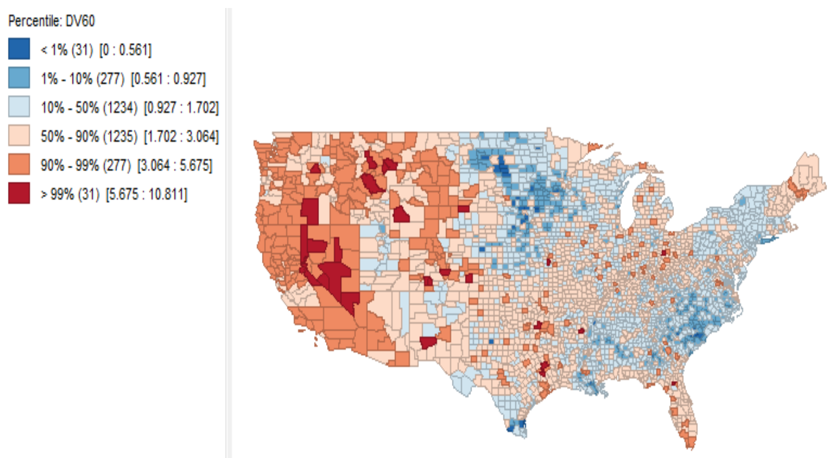


**Figure 13.** Divorce rates in the 1960s (DV60)

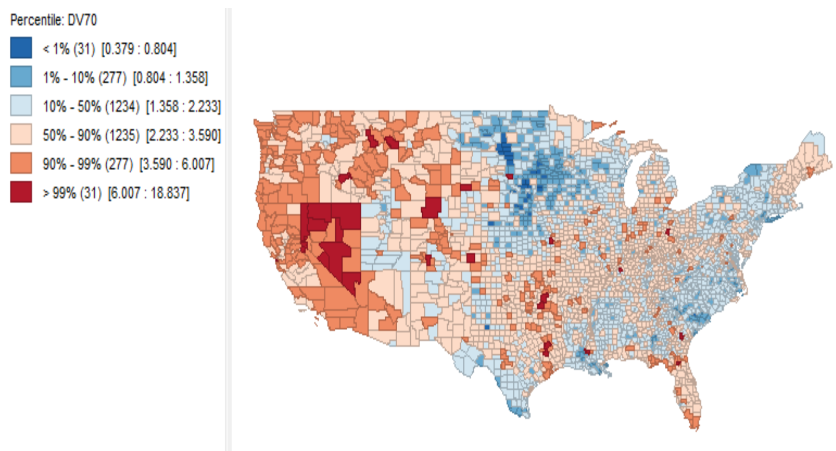


**Figure 14.** Divorce rates in the 1970s (DV70)

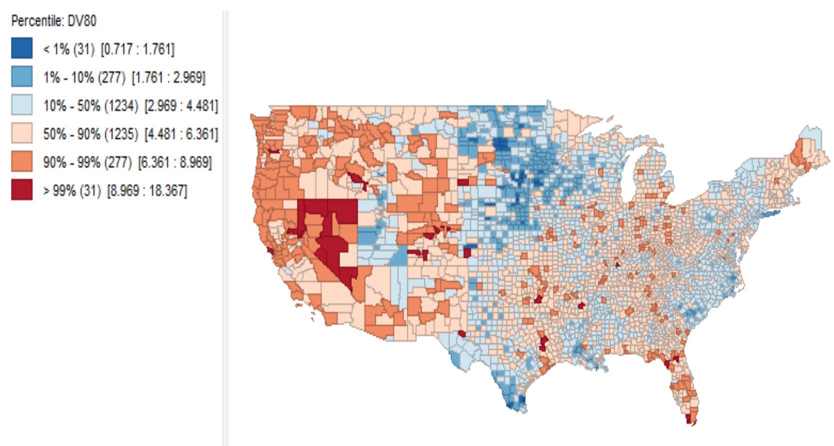


**Figure 15.** Divorce rates in the 1980s (DV80)

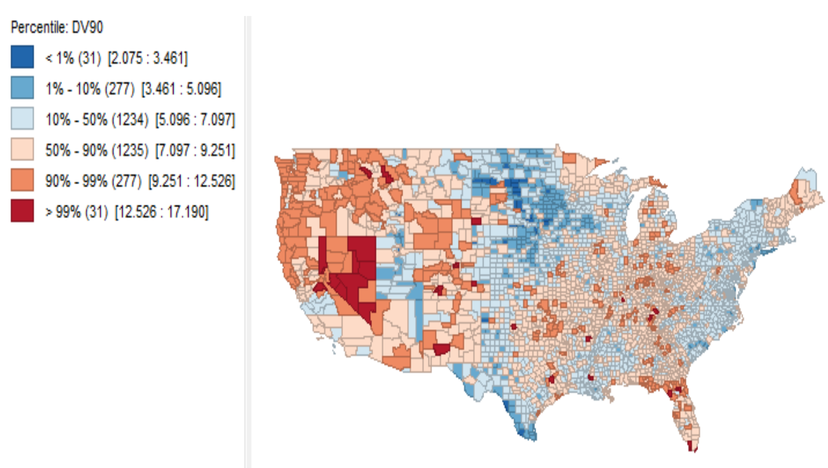


**Figure 16.** Divorce rates in the 1990s (DV90)

### The Augmented Dickey-Fuller Test

The ADF test is the extended version of the simple Dickey-Fuller (DF) Test. DF test was extended by including extra lagged in terms of the dependent variables in order to eliminate the problem of autocorrelation (Mushtaq, 2012). One of the common methods to find the order of integration of variables is the unit root test. One of the most popular among them is the ADF test. The equation for the ADF test is shown in Equation 1 (Dickey, 1979; Dickey, 1981).

$$\Delta Y_t = a + \beta T + pY_{t-1} + \sum_{i=1}^{k} \gamma_i \Delta Y_{t-i} + e_t \qquad (1)$$

Where is "$Y_t$" the variable in period t, "$T$" denotes a time trend, "$\Delta$" is the difference operator, "$e_t$" is an error term disturbance with mean zero and variance $\sigma^2$, and "$k$" represents the number of lags of the differences in the ADF equation. The ADF test is restricted by its number of lags. It decreases the power of the test to reject the null of a unit root because the increased number of lags necessitates the estimation of additional parameters and a loss of a degree of freedom (Hosseini, 2011).

### The Least Squares Regression Method

The LS method is a statistical method of analysis that estimates the relationship between one or more independent variables and the dependent variable by minimizing the sum of squares of the difference between the observed and predicted values of the dependent variable (What-when-how, 2019). A simple LS regression model involving only one independent variable "$X$" predicting a dependent variable "$Y$" is expressed by Equation 2.

$$Y = a + bX + \varepsilon \qquad (2)$$

In Equation 2, "$a$" is the intercept that indicates where the straight line intersects Y-axis; "$b$" is the slope that indicates the degree of steepness of the straight line and "$\varepsilon$" represents the error. The best line or relationship would be the one with the least sum of squared errors (Darity, 2008).

$$SSE = \Sigma \left(Yi - \hat{Y}\right)^2 \qquad (3)$$

Where, $Y_i$ = Dependent variable for $i$ and where $i$ = 1,......n

$$\hat{Y} = (\Sigma\, Yi)/n \qquad (4)$$

The LS regression is based upon the data having a normal distribution and this does not always occur. If the data is truly normal, then the mean and median would be the same. The other major disadvantage of the LS method is that it is sensitive to the outlier that is due to the squaring of the error term. Since the LS method is not robust, this might have a tremendous impact on the predicted cost (Foussier, 2010).

### Quantile Regression

Quantile regression is the estimation of quantiles of the conditional distribution. It models the relationship between a set of predictor variables and the specific quantiles of the response variable (Koenker, 2005). Quantile regression is another method to estimate a median and confidence interval (CI) or a difference in median charges between groups of interest. In contrast to LS methods, quantile regression makes no distributional assumptions and can result in estimates of either additive or multiplicative differences between groups (Koenker, 2001). Therefore, it can be considered as a more flexible approach than the LS regression. The LS regression looks for a model for the conditional expected value of the dependent variable, while the quantile regression determines the model for the selected quantiles in the conditional distribution of the dependent variable (Yavuz, 2017). In this regression method, any quantile can be used to model the predetermined position of the distribution (Davino, 2013). Quantile regression considers the effects of the dependent variable on each point of the conditional distribution; therefore, the asymmetric tail of the distribution is also included in the analysis (Abadie et al., 2002). Also, the LS method is dependent on the conditional average; the quantile regression method is dependent on conditional quantile function (Koenker, 2005).

This method can be used to measure the effect of explanatory variables, not just in the center of the distribution, but also on the right or left tail of the distribution. This is very useful in the application, particularly for estimating extreme values (Djuraidah, 2011). For a random variable $Y$ given distribution function as follows in Equation 5:

$$F(y) = P(Y \leq y) \qquad (5)$$

The quantile to τ. from Y to 0 <τ <1, denoted by:

$$Q(\tau) = inf\{y{:}F(y) \geq \tau\} \qquad (6)$$

## RESULTS

### Augmented Dickey-Fuller Test

Due to the variables used in this study are time-dependent variables, a unit root test has applied. The ADF test was select among the unit root tests. The probability values obtained from this test have given in Table 2. Probability values for all variables in the table are below the 0.05 level. Therefore, the $H_{1a}$ hypothesis has been rejecting for all variables. In other words, all variables in the study are statistically significant and do not have unit-roots. Thus, there is no need for fixing for variables.

$H_{1a}$ null hypothesis: p=0 (The variable is non-stationary and has a unit root).

$H_{1b}$ alternative hypothesis: p<0 (The variable is stationary and hasn't a unit root).

**Table 2.** Augmented Dickey-Fuller Test

| Variable | t-Statistic | Probability |
|---|---|---|
| HC60 | -6.161878 | 0.0001* |
| HC70 | -6.184898 | 0.0001* |
| HC80 | -6.985473 | 0.0001* |
| HC90 | -54.92391 | 0.0001* |
| DV60 | -32.29889 | 0.0001* |
| DV70 | -33.11428 | 0.0001* |
| DV80 | -32.94403 | 0.0001* |
| DV90 | -31.68651 | 0.0001* |
| PO60 | -5.892215 | 0.0001* |
| PO70 | -6.330890 | 0.0001* |
| PO80 | -53.60768 | 0.0001* |
| PO90 | -53.48813 | 0.0001* |
| UE60 | -20.78529 | 0.0001* |
| UE70 | -18.25447 | 0.0001* |
| UE80 | -19.25076 | 0.0001* |
| UE90 | -18.32790 | 0.0001* |

Note: Probability values marked with * are statistically significant at the 0.05 significance level.

### The Least Squares Regression Method

While HC was the dependent variable, the LS method has applied separately every 10 years. The numbers 60, 70, 80, and 90 indicate periods in data. As a result of this method, 4 different regression models had obtained. Determination coefficients ($R^2$), probability values, and F-statistics of these models have given in Table 3. Whist the HC60 variable is the dependent variable in the 1st model; DV60, PO60 variables, and constant (C) coefficient have been found as significant. Besides, the $R^2$ of the 1st model is 86.81%. In other words, the correlation coefficient among the variables in the 1st model is 86.81%. The same results have been found for the 2nd model, which has established in the 1970s. In this model, HC70 is the dependent variable, and only the coefficients are different. The $R^2$ value for this model is 84.01%. While the HC80 variable is the dependent variable in the 3rd model; DV80, PO80, UE80 variables, and C coefficient have been found as significant. Besides, the $R^2$ value of the model is 85.43%. While HC90 is the dependent variable for the 4th model, PO90, UE90 variables, and C coefficient are significant. Since the probability values of all models are less than 0.05 level, the $H_{2a}$ hypothesis has been rejecting for all models. In other words, established all regression models are statistically significant.

$H_{2a}$ null hypothesis: Correlation of observed points with the regression line is negligible.

$H_{2b}$ alternative hypothesis: Correlation of observed points with the regression line is not negligible.

To use these models established by the LS method, they have to provide some assumptions. One of the main assumptions is to distribute in accordance with the normal distribution. In Table 3, four Jarque-Bera probability values obtained for each model are given. All of these values are less than 0.05 significance level. Therefore, the $H_{3a}$ hypothesis was rejected. It is said that the data in the 4 models established are not distributed according to the normal distribution at a 95% confidence level. Since the data are not normally distributed, it would be more accurate to use the quantile regression method, which is an alternative to the LS method.

$H_{3a}$ null hypothesis: The data correspond to the normal distribution at a 95% confidence level.

$H_{3b}$ alternative hypothesis: The data don't correspond to the normal distribution at a 95% confidence level.

**Table 3.** Regression Models with LS Method

| **Dependent Variable**: HC60 | | **Independent Variables**: DV60, PO60, UE60, C | |
|---|---|---|---|
| **1st Model**: HC60 = DV60*0.239111+PO60*5.31E-05-1.061573 | | | |
| **Model $R^2$** | 86.81% | **Normality Test** | |
| **Model Prob.** | 0.0001 | **Jarque-Bera** | 959363.8 |
| **Model F-statistic** | 6760.968 | **Probability** | 0.0001 |
| **Dependent Variable**: HC70 | | **Independent Variables**: DV70, PO70, UE70, C | |
| **2nd Model**: HC70 = DV70*0.590382+PO70*0.000126-4.337648 | | | |
| **Model $R^2$** | 84.01% | **Normality Test** | |
| **Model Prob.** | 0.0001 | Jarque-Bera | 1504273 |
| **Model F-statistic** | 5397.158 | Probability | 0.0001 |
| **Dependent Variable**: HC80 | | **Independent Variables**: DV80, PO80, UE80, C | |
| **3rd Model**: HC80 = DV80*-0.909979+PO80*0.000193+UE80*0.309593-4.530943 | | | |
| **Model $R^2$** | 85.43% | **Normality Test** | |
| **Model Prob.** | 0.0001 | **Jarque-Bera** | 1185321 |
| **Model F-statistic** | 6024.415 | **Probability** | 0.0001 |
| **Dependent Variable**: HC90 | | **Independent Variables**: DV90, PO90, UE90, C | |
| **4th Model**: HC90 = PO90*0.000191+UE90*0.956572-10.64183 | | | |
| **Model $R^2$** | 82.80% | **Normality Test** | |
| **Model Prob.** | 0.0001 | **Jarque-Bera** | 5351930 |
| **Model F-statistic** | 4944.192 | **Probability** | 0.0001 |

**The Quantile Regression**

3 different quantile regression models have created using 0.25, 0.50, and 0.75 quantile values. These quantile values have examined the distribution in three parts. Variables with a probability value of less than 0.05 for the independent variables in Table 4 are include in the models created. Due to the $H_{4a}$ hypothesis has been rejecting, these variables are statistically significant.

$H_{4a}$ null hypothesis: The variable is not significant.
$H_{4b}$ alternative hypothesis: The variable is significant.
For the first dependent variable HC60 in Table 4, the $R^2$ of the 1st model established with 0.25 quantile is 20.9%. Significant independent variables in this model have founded to be DV60, PO60, and C coefficient. The $R^2$ is 32.9% in the 2nd model, which has been establishing with 0.50 quantile. Significant independent variables in the model have founded to be DV60 and PO60. The $R^2$ is 53.7% in the 3rd model, which has been establishing with 0.75 quantile value. Only the PO60 variable is significant in this model. In models where 3 quantile values have been used, the 3rd model has the highest $R^2$ value. Therefore, the interpretations have been creating on the 3rd model. As a result, the PO60 variable positively affects the dependent variable HC60. This effect is a positive and has smallish value of 6.55E-05. Besides, the Quasi-LR probability values of the models installed with

all quantiles are less than 0.05 level. Therefore, the $H_{5a}$ hypothesis has been rejecting, and all models are statistically significant.

$H_{5a}$: Correlation of observed points with the regression line is negligible. (The model is not significant).

$H_{5b}$: Correlation of observed points with the regression line isn't negligible. (The model is significant).

**Table 4.** Quantile Regression for HC60

| **Dependent Variable =** HC60 | | **Quantile (Tau) =** 0.25 | |
|---|---|---|---|
| **Variable** | **Coefficient** | **t-Statistic** | **Probability** |
| **DV60** | 0.036515 | 4.165415 | 0.0001* |
| **PO60** | 1.88E-05 | 14.99944 | 0.0001* |
| **UE60** | 0.005234 | 1.535359 | 0.1248 |
| **C** | -0.245256 | -7.993588 | 0.0001* |
| **1st Model:** HC60 = DV60*(0.036515)+PO60*(1.88E-05)-0.245256 | | | |
| **Pseudo $R^2$ =** 20.9% | | **Prob.(Quasi-LR) =** 0.0001* | |
| **Dependent Variable =** HC60 | | **Quantile (Tau) =** 0.50 | |
| **Variable** | **Coefficient** | **t-Statistic** | **Probability** |
| **DV60** | 0.034899 | 2.964337 | 0.0031* |
| **PO60** | 3.95E-05 | 2.332133 | 0.0198* |
| **UE60** | 0.011009 | 0.692142 | 0.4889 |
| **C** | -0.267985 | -1.822900 | 0.0684 |
| **2nd Model:** HC60 = DV60*(0.034899)+PO60*(3.95E-05) | | | |
| **Pseudo $R^2$ =** 32.9% | | **Prob.(Quasi-LR) =** 0.0001* | |
| **Dependent Variable =** HC60 | | **Quantile (Tau) =** 0.75 | |
| **Variable** | **Coefficient** | **t-Statistic** | **Probability** |
| **DV60** | -0.010944 | -0.366748 | 0.7138 |
| **PO60** | 6.55E-05 | 4.048032 | 0.0001* |
| **UE60** | 0.020490 | 0.951978 | 0.3412 |
| **C** | -0.027869 | -0.155852 | 0.8762 |
| **3rd Model:** HC60 = PO60*(6.55E-05) | | | |
| **Pseudo $R^2$ =** 53.7% | | **Prob.(Quasi-LR) =** 0.0001* | |

Note: Probability values marked with * are statistically significant at the 0.05 significance level.

For the first dependent variable HC70 in Table 5, the $R^2$ of the 1st model established with 0.25 quantile is 20.1%. Significant independent variables in this model have founded to be DV70, PO70, and C coefficient. The $R^2$ is 32.3% in the 2nd model, which has been establishing with 0.50 quantile. Significant independent variables in this model have founded to be DV70, PO70, and C coefficient. The $R^2$ is 54.6% in the 3rd model, which has been establishing with 0.75 quantile value. Significant independent variables in this model have founded to be DV70, PO70, UE70 and C coefficient. In models where 3 quantile values have been used, the 3rd model has the highest $R^2$ value. Therefore, the interpretations have been creating on the 3rd model. As a result, DV70, PO70, UE70, and C coefficient positively affects the dependent variable HC70. Besides, the Quasi-LR probability values of the models installed with all quantiles are less than 0.05 level. Therefore, the $H_{5a}$ hypothesis has been rejecting, and all models are statistically significant.

**Table 5.** Quantile Regression for HC70

| **Dependent Variable =** HC70 | | **Quantile (Tau) =** 0.25 | |
|---|---|---|---|
| **Variable** | **Coefficient** | **t-Statistic** | **Probability** |
| **DV70** | 0.081608 | 5.621572 | 0.0001* |
| **PO70** | 3.33E-05 | 7.267548 | 0.0001* |
| **UE70** | -0.005730 | -1.083652 | 0.2786 |
| **C** | -0.451057 | -5.246897 | 0.0001* |
| **1.Model:** HC70 = DV70*(0.081608)+PO70*(3.33E-05)-0.245256 | | | |
| **Pseudo $R^2$ =** 20.1% | | **Prob.(Quasi-LR) =** 0.0001* | |
| **Dependent Variable =** HC70 | | **Quantile (Tau) =** 0.50 | |
| **Variable** | **Coefficient** | **t-Statistic** | **Probability** |
| **DV70** | 0.067540 | 4.469420 | 0.0001* |
| **PO70** | 6.22E-05 | 29.11586 | 0.0001* |
| **UE70** | -0.007082 | -1.148390 | 0.2509 |
| **C** | -0.374598 | -7.932910 | 0.0001* |
| **2.Model:** HC70 = DV70*(0.067540)+PO70*(6.22E-05)-0.374598 | | | |
| **Pseudo $R^2$ =** 32.3% | | **Prob.(Quasi-LR) =** 0.0001* | |
| **Dependent Variable =** HC70 | | **Quantile (Tau) =** 0.75 | |
| **Variable** | **Coefficient** | **t-Statistic** | **Probability** |
| **DV70** | 0.079696 | 3.655579 | 0.0003* |
| **PO70** | 0.000143 | 19.23782 | 0.0001* |
| **UE70** | -0.025680 | -2.886389 | 0.0039* |
| **C** | -0.585001 | -5.632792 | 0.0001* |
| **3.Model:** HC70 = DV70*(0.079696)+PO70*(0.000143)+UE70*(-0.025680)-0.585001 | | | |
| **Pseudo $R^2$ =** 54.6% | | **Prob.(Quasi-LR) =** 0.0001* | |

Note: Probability values marked with * are statistically significant at the 0.05 significance level.

For the first dependent variable HC80 in Table 6, the $R^2$ of the 1st model established with 0.25 quantile is 22.7%. Significant independent variables in this model have founded to be DV80, PO80, UE80, and C coefficient. The $R^2$ is 35.5% in the 2nd model, which has been establishing with 0.50 quantile. Significant independent variables in this model have founded to be DV80, PO80, UE80, and C coefficient. The $R^2$ is 55.1% in the 3rd model, which has been establishing with 0.75 quantile value. Significant independent variables in this model have founded to be PO80, UE80, and C coefficient. In models where 3 quantile values have been used, the 3rd model has the highest $R^2$ value. Therefore, the interpretations have been creating on the 3rd model. As a result, PO80, UE80, and C coefficient positively affects the dependent variable HC80. Besides, the Quasi-LR probability values of the models installed with all quantiles are less than 0.05 level. Therefore, the $H_{5a}$ hypothesis has been rejecting, and all models are statistically significant.

**Table 6.** Quantile Regression for HC80

| **Dependent Variable =** HC80 | | **Quantile (Tau) =** 0.25 | |
|---|---|---|---|
| **Variable** | **Coefficient** | **t-Statistic** | **Probability** |
| **DV80** | 0.108299 | 5.740182 | 0.0001* |
| **PO80** | 4.32E-05 | 9.358983 | 0.0001* |
| **UE80** | -0.022339 | -3.010243 | 0.0026* |
| **C** | -0.663429 | -11.15546 | 0.0001* |
| **1st Model:** HC80 = DV80*(0.108299)+PO80*(4.32E-05)+UE80*(-0.022339)-0.663429 | | | |
| **Pseudo $R^2$ =** 22.7% | | **Prob.(Quasi-LR) =** 0.0001* | |
| **Dependent Variable =** HC80 | | **Quantile (Tau) =** 0.50 | |
| **Variable** | **Coefficient** | **t-Statistic** | **Probability** |
| **DV80** | 0.089863 | 3.382167 | 0.0007* |
| **PO80** | 8.01E-05 | 7.924656 | 0.0001* |
| **UE80** | -0.024293 | -3.082437 | 0.0021* |
| **C** | -0.596491 | -8.454475 | 0.0001* |
| **2nd Model:** HC80 = DV80*(0.089863)+PO80*(8.01E-05)+UE80*(-0.024293)-0.596491 | | | |
| **Pseudo $R^2$ =** 35.5% | | **Prob.(Quasi-LR) =** 0.0001* | |
| **Dependent Variable =** HC80 | | **Quantile (Tau) =** 0.75 | |
| **Variable** | **Coefficient** | **t-Statistic** | **Probability** |
| **DV80** | 0.040317 | 1.724240 | 0.0848 |
| **PO80** | 0.000160 | 16.75676 | 0.0001* |
| **UE80** | -0.034280 | -4.490418 | 0.0001* |
| **C** | -0.646749 | -6.803574 | 0.0001* |
| **3rd Model:** HC80 = PO80*(0.000160)+UE80*(-0.034280)-0.646749 | | | |
| **Pseudo $R^2$ =** 55.1% | | **Prob.(Quasi-LR) =** 0.0001* | |

Note: Probability values marked with * are statistically significant at the 0.05 significance level.

For the first dependent variable HC90 in Table 7, the $R^2$ of the 1st model established with 0.25 quantile is 21.8%. Significant independent variables in this model have founded to be DV90, PO90, UE90, and C coefficient. The $R^2$ is 34.5% in the 2nd model, which has been establishing with 0.50 quantile. Significant independent variables in this model have founded to be PO90, UE90, and C coefficient. The $R^2$ is 52.5% in the 3rd model, which has been establishing with 0.75 quantile value. Significant independent variables in this model have founded to be PO90, UE90, and C coefficient. In models where 3 quantile values have been used, the 3rd model has the highest $R^2$ value. Therefore, the interpretations have been creating on the 3rd model. As a result, PO90, UE90, and C coefficient positively affects the dependent variable HC90. Besides, the Quasi-LR probability values of the models installed with all quantiles are less than 0.05 level. Therefore, the $H_{5a}$ hypothesis has been rejecting, and all models are statistically significant.

**Table 7.** Quantile Regression for HC90

| **Dependent Variable =** HC90 | | **Quantile (Tau) =** 0.25 | |
|---|---|---|---|
| **Variable** | **Coefficient** | **t-Statistic** | **Probability** |
| **DV90** | 0.043989 | 3.979241 | 0.0001* |
| **PO90** | 4.17E-05 | 9.384289 | 0.0001* |
| **UE90** | 0.044659 | 5.092985 | 0.0001* |
| **C** | -1.024431 | -5.909827 | 0.0001* |
| **1st Model:** HC90 = DV90*(0.043989)+PO90*(4.17E-05)+UE90*(0.044659)-1.024431 | | | |
| **Pseudo $R^2$ =** 21.8% | | **Prob.(Quasi-LR) =** 0.0001* | |
| **Dependent Variable =** HC90 | | **Quantile (Tau) =** 0.50 | |
| **Variable** | **Coefficient** | **t-Statistic** | **Probability** |
| **DV90** | 0.007338 | 0.576202 | 0.5645 |
| **PO90** | 7.49E-05 | 13.98986 | 0.0001* |
| **UE90** | 0.046036 | 6.473973 | 0.0001* |
| **C** | -0.764783 | -8.300535 | 0.0001* |
| **2nd Model:** HC90 = PO90*(7.49E-05)+UE90*(0.046036)-0.764783 | | | |
| **Pseudo $R^2$ =** 34.5% | | **Prob.(Quasi-LR) =** 0.0001* | |
| **Dependent Variable =** HC90 | | **Quantile (Tau) =** 0.75 | |
| **Variable** | **Coefficient** | **t-Statistic** | **Probability** |
| **DV90** | -0.021876 | -1.373682 | 0.1696 |
| **PO90** | 0.000142 | 66.83652 | 0.0001* |
| **UE90** | 0.044648 | 4.907233 | 0.0001* |
| **C** | -0.704655 | -7.100558 | 0.0001* |
| **3rd Model:** HC90 = PO90*(0.000142)+UE90*(0.044648)-7.100558 | | | |
| **Pseudo $R^2$ =** 52.5% | | **Prob.(Quasi-LR) =** 0.0001* | |

Note: Probability values marked with * are statistically significant at the 0.05 significance level.

## CONCLUSIONS

This study was carried out to determine some factors affecting the number of deaths in the 1960s, 1970s, 1980s, and 1990s throughout the USA. In the data set; divorce rates, country population, and unemployment rates were added as variables to determine these factors. Initially, the densities of all variables including the HC variable were examined. Such densities were shown by grouping the USA. map into districts. In this demonstration, the GeoDa program is used and the results were shown with the help of a figure. Some of the results are as follows: Most of the murders in the United States are observed in the Southwest, Southeast, and Eastern regions. The middle regions of the USA are calmer compared to these regions. Population density has not changed in almost any region. Although the density is the same, only the number of populations increased as the years progressed. In the 1990s unemployment in Western regions; It is seen that it has made progress towards the South West region. In general, it is seen that the intensity of unemployment rates decreases with years. Divorce rates were found to be high, especially in the counties of the state of Nevada. It has been determined that the number of divorces has increased continuously by years.

Since the variables used in the study were time-dependent, the unit root test was performed in the EViews 9 program. For this purpose, the ADF test was applied. All variables in the test results are stationary. Thus, regression analysis was started without any stationarizing. With the LS method, 4 different regression models were established for 4 periods. However, all models don't correspond to the normal distribution for this regression method. Therefore, it was decided to apply quantile regression, a more flexible regression type in terms of assumptions.

Quantile regression is more flexible for assumptions such as normal distribution and extreme values, so the results are determined through quantile regression. Since the data were analyzed by dividing in quantile regression, the data were divided into 3 parts in this study. 0.25 (1st quarter), 0.50 (Median), and 0.75 (3rd quarter) were used as a quantile value for separation. Three regression equations were established for each of the 1960s, 1970s, 1980s, and 1990s. The correlation coefficient of the model, which was established with a value of 0.75 for all periods, was found to be the model with the largest and most accurate results. The results of the model equations visualized in Figure 17 are as follows (The effect of the constant-coefficient was considered insignificant in the comments.): For the 1960s, 1-unit increase in-country population increases homicide counts by 6.55E-05 units. In the 1970s, 1-unit increases in the country population increase the homicide counts by 0.000143 units. 1-unit increases in divorce rates also increase the homicide counts by 0.079696 units. In contrast, 1-unit decreases in the unemployment rates decrease the homicide counts by 0.025680 units. It can be said that the energy crises of the 1970s caused all the variables to have an effect in the 1970s (Time, 1979). Owing to these crises have affected the American economy seriously. The deterioration of the economy may have had an impact on the factors in this study. In the 1980s, 1-unit increases in the country population increase the homicide counts by 0.000160 units. In contrast, 1-unit decreases in the unemployment rates decrease the homicide counts by 0.034280 units. And finally, in the 1990s, 1-unit increases in the country population increase the homicide counts by 0.000142 units. 1-unit increases in unemployment rates also increase the homicide counts by 0.044648 units. As a general result; It can be said that the country population, which has a common effect on the 1960s, 1970s, 1980s, and 1990s, is a directly proportional factor for these four periods. The increase in the population of the USA increases the homicide counts.

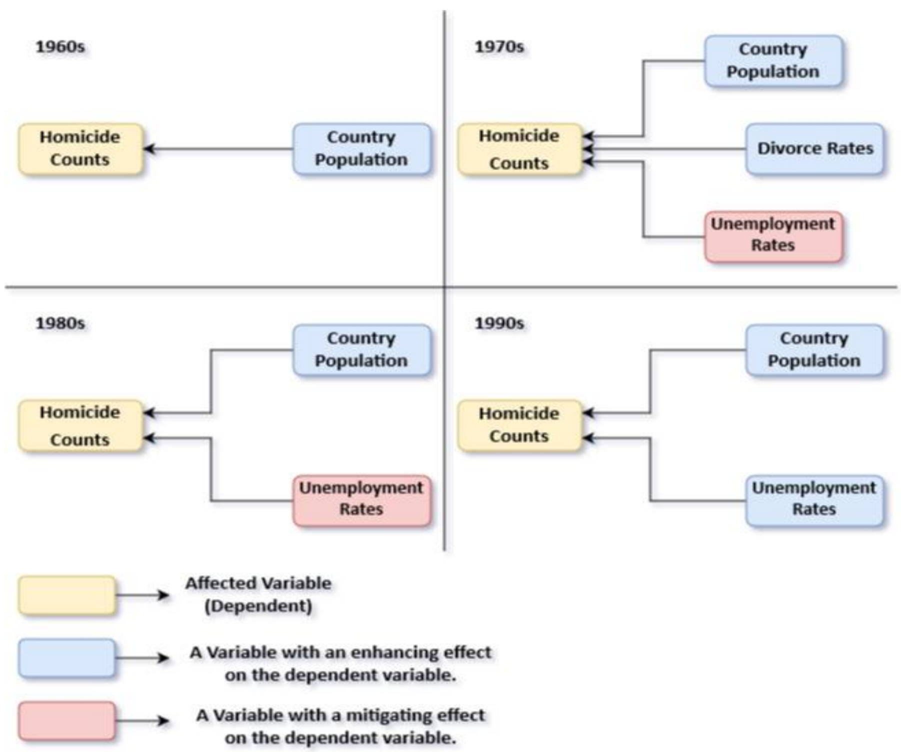


**Figure 17.** Summary of 0.75 quantile models

## REFERENCES

Abadie, A., Angrist, J., Imbens, G. (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. *Econometrica* 70(1): 91-117.

Baller, R., Anselin, L., Messner, S., Deane, G., Hawkins, D. (2001). Structural covariates of us county homicide rates, incorporating spatial effects. *Criminology* 39: 561-590.

Davino, C., Furno, M., Vistocco, D. (2013). Quantile Regression, Theory and Applications. John Wiley and Sons.

Dickey, D.A., Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74(366): 427-431.

Dickey, D.A., Fuller, W.A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica, Journal of the Econometric Society* 49(4): 1057-1072.

Djuraidah, A., Wigena, A.H. (2011). Regression quantile for exploration rainfall patterns in indramayu. *A Basic Sciences* 12(1): 50-56.

Foussier, P. (2010). Improving CER building, basing a cer on the median. *Journal of Cost Analysis and Parametrics* 3(2): 1-12.

GeoDa Data and Lab (2003). Homicides + socio-economics (1960-90). https://geodacenter.github.io/data-and-lab/ncovr/. (Access Date: 12.19.2019).

Hosseini, S.M., Ahmad, Z., Lai, Y.W. (2011). The role of macroeconomic variables on stock market index in China and India. *International Journal of Economics and Finance* 3(6): 233-143.

Humphreys, D., Gasparrini, A., Wiebe, D. (2017). Evaluating the impact of Florida's "stand your ground" self-defense lawon homicide and suicide by firearm, an interrupted time series study. *JAMA Internal Medicine* 177(1): 44-50.

Darity, W.A. (2008). Ordinary least squares regression. *International Encyclopedia of the Social Sciences* 2(6): 57-61.

Koenker, R., Hallock, K.F. (2001). Quantile regression. *Journal of Economic Perspectives* 15: 143-56.

Koenker, R. (2005). *Quantile Regression*. London, Cambridge University Press, 1-30.

Kposowa, A., Breault, K., Harrison, B. (1995). Reassessing the structural covariates of violent and property crimes in the USA, a county level analysis. *The British Journal of Sociology* 46(1): 79.

Messner, S.F., Anselin, L., Hawkins, D., Deane G., Tolnay S., Baller, R. (2000). An atlas of the spatial patterning of county-level homicide, 1960-1990. Pittsburgh, National Consortium on Violence Resarch, Carnegie-Mellon University.

Mushtaq, R. (2012). Augmented Dickey Fuller Test. SSRN Electronic Journal, 1-19.

Ousey, G.C., Kubrin, C.E. (2014). Immigration and the changing nature of homicide in us cities, 1980-2010. *Journal of Quantitative Criminology* 30: 453-483.

Sen, B., Wingate, M., Kirby, R. (2012). The relationship between state abortion-restrictions and homicide deaths among children under 5 years of age, a longitudinal study. Social Science and Medicine 75(1): 156-164.

Sipsma, H., Canavan, M., Rogan, E., Taylor, L.A., Talbert-Slagle, K.M., Bradley, E.H. (2017). Spending on social and public health services and its association with homicide in the USA, an ecological study. *BMJ Open* 7(10): 1-7.

Time (1979). Business: Oil Squeeze. http://content.time.com/time/magazine/article/0,9171,946222,00.html (Access Date: 29.12.2019).

What-when-how (2019). Ordinary least squares regression (social science). http://what-when-how.com/social-sciences/ordinary-least-squares-regression-social-science/ (Access Date: 12.22.2019).

Yavuz, A.A., Aşik, E.G. (2017). Quantile regression. International Journal of Engineering Research and Development 9(2): 137-146.