Analysis of Factors Affecting Crime in Texas

Introduction Statement

In 2021, the FBI has reported 694,050 violent-crime incidents, and 817,020 offenses in the United States. Over the last decade there has been an explosion in the crime rate in many states of the USA, Texas being one of the states with the most significant rise. Identifying the underlying causes and preconditions of crime is crucial to crime reduction and prevention. Historically, the causes and origins of crime have been the subjects of investigation by many disciplines, some of the most cited factors have been correlated with crime are the socio-economic conditions; our team has decided to apply econometric models and validate existing research results with the data available for various factors for each county of Texas.

The overarching question is to determine the effects of socio-economic factors on crime rate. We plan to use techniques like Simple Linear Regression and Multiple Linear Regression to test if there is a connection between the factors and the crime rate with a t-test. Since Texas has 254 counties, we will use the 2021 cross-sectional data for our regression analysis. Once we have the results, we will validate them with existing research and gauge if the dynamics of Texas are similar to other countries. Our contribution to existing literature would be the localized study of Texas' counties and utilizing unconventional factors like number of churches, the divorce rate in the region, and time series analysis of some factors if time.

We attach importance to our study and our objective based on two dimensions. First, for the study of the dependent variable, which is important for the study of a region, a province or city or even a country to regulate the crime phenomenon. Second, for socio-economic factors among the dependent variables, it is beneficial for our application of the combination of economic and statistical knowledge throughout the course and for a broader understanding of macroeconomic knowledge.

Literature Review

Addressing the relationship between crime and other factors (age, income, poverty, unemployment, police, divorce rate, etc.) is extensive. The brief literature reviews below highlight various approaches from existing work linked to the factors we have chosen for our data analysis.

1. Average Age

In this line of research, Britt (1997) analyzes age-specific for the seven index offenses for the years 1958–1995 to test two propositions concerning the impact of unemployment on crime found in Greenberg's (1977, 1985) theoretical treatment of the age–crime relationship. Quoted verbatim from Britt's paper, the two propositions are as follows: (H1): ''The unemployment–crime relationship will vary by age group, where youth and young adults are expected to show a greater motivational (positive) effect of unemployment on criminal behavior'' (H2) ''The unemployment–crime relationship will vary overtime, especially for youth and young adults, where the motivational (positive) effect of unemployment is expected to increase over time.''

In another study, Gould et al. (2002) examined the impact of unemployment and wages on the criminal behavior of less-educated young males. It focused in detail on two specific elements, the relationship between crime and labor market opportunities and crime's economic returns from illegal opportunities. A panel of US counties finds significant effects for both wages and unemployment on property and violent crimes. There is a strong relationship between Youth Crime and Youth Unemployment. For this specific study, it indicates that the Youth are more likely to commit crime than any other age group if the economic returns from crime far outweigh the return from legal opportunities.

2. Dropout Rate

75% of America's state prison inmates are high school dropouts (Harlow, 2003).
High school dropouts are 3.5 times more likely than high school graduates to be arrested in their lifetime (Alliance for Excellent Education, 2003a)

Brandy et.al. (2014) examined the distribution of criminal behavior among dropouts derived from a nationally representative sample of 18–25-year-old (N = 19,312) emerging adults in the United States. Dropouts were more likely to be arrested for larceny, assault, drug possession or drug sales relative to their high school graduate counterparts.

3. Unemployment Rate

Next are some highlights of the effects of unemployment on crime rates by Guanlin Gao, Bo Liu, and Ingrid Kouassi. Their study used a fixed-effect model and employed balanced panel data from 23 sample counties in Indiana from 2006-2013. This study is strikingly different from our analysis where we looked at 254 counties in the State of Texas. From their analysis, they also used multiple linear regression. They controlled the population size, demographics, socioeconomic, and county-specific characteristics. The results from their study showed that

unemployment rate has a negative contemporaneous effect on the violent crime rates and null influence on property crime rate.

In another unemployment rate/crime study, David Cantor and Kenneth Land's (1985) analyzed the relationship between annual unemployment rates and crime rates in the United States. Their analysis has served as a model for subsequent criminological time series analyses. Cantor and Land's conclusion is that the impact of unemployment on crime rates is weak and inconsistent. The unemployed are expected to have greater motivation to violate the law, but we cannot conclude that this hypothesis goes for everyone. Those who are unemployed may be spending more time at home, completely avoiding running into the law. Cantor and Land also note that while unemployment could reduce the motivation to commit crime, it may also increase the motivation to do so. If both effects are instantaneous, a coefficient representing the net effect of unemployment on crime might be small and insignificant even though both effects are substantial. (Cantor, Land 1991)

4. Per Capita Income

Steven D. Levitt (1999) published a study looking at the relationship between crime and income. In his model, it employed nine potential explanatory variables that are correlated with violent crime. All variables are in logged form in order to avoid outliers. The model is to be estimated as follows: $\ln(y_{it}) = \alpha + \beta \ln(X_{it}) + \mu_i + \nu_{it}$ (3.1) using the Gini coefficient technique. Levitt's conclusion was that there was a significant positive relationship between high school dropout rates and violent crime during the Great Recession, but there doesn't seem to be any relationship during the recovery. This suggests that high school students were less likely to drop out during the Great Recession. High school dropouts have fewer opportunities in the labor market than in normal economic times. Those without a higher level of education will not have a great chance of surviving the labor market in normal economic times. Thus, these people will convert to criminal activities since it can bring high rewards. Based on this study, lower-income people may be more willing to commit crimes.

5. Racial Distribution

In a study published by Michael Light and Casey Harris in 2012, they analyzed the racial invariance hypothesis to community structure and crime. In their study, they used 1999-2001 county-level arrest data with different methods like regression models, lag models, and geographic weighted regression analyses. The data included the county-level white and black violent index which is the sum of arrests for homicide, aggravated assault, forcible rape, and robbery offending rates per 100,000. Light and Harris used three-year averages to reduce any potential bias, and to reduce the skewness of the dependent variables they log transformed the rates. The conclusion of their study indicates that white and black violent crime is indeed spatially patterned (not randomly distributed).

6. Divorce Rate

Pearl et. al. (2018) conducted a study on children whose parents were separated during childhood. Study included individuals born in Denmark, 1971–1997 (N=1,346,772). Child–parent separation status was ascertained each year from birth to 15th birthday, using residential addresses from the Danish register. Members were followed up from their 15th birthday until the date of first violent offense conviction, or December 31, 2012. Incidence rate ratios were estimated using survival analyses techniques.

Separation from a parent during childhood was associated with elevated risk for subsequent violent offending versus those who lived continuously with both parents. Associations were stronger for paternal than for maternal separation at least up until mid-childhood and rose with the number of separations.

Data Description

The data collected was for the year 2021 and each record is the some sort of aggregate per each county. Most of the source of our data has been the txcip.org; www.texascounties.net; ucr.fbi.gov; US Census data. Here is the summary of each of the variables collected

Variable	Label	Obs	Mean	Std Dev	Min	Max
lcrimerate	log of crimes reported per 100,000	254	-4.435536	0.8310079	-8.048788	-0.9985288
pop	Population	254	116251.7	412059.6	57	4728030
inc	Median Income	254	57291.25	12982.99	33513	106225
unem	Unemployment Rate	254	0.056126	0.0196225	0.009	0.162
age	Median Age	254	39.55827	5.745828	27.9	60
poverty	Poverty Rate	254	0.1453701	0.0442565	0.039	0.287
drop	High school Dropout rate(7-12 y.o)	254	1.008268	1.477864	0	13
church	No. of Churches per person	254	0.0027462	0.0017284	0.0002209	0.013986
police	No. of Police officers per person	254	0.0025575	0.0022118	0.0004624	0.0213816
div2010	Divorce rate in 2010	254	0.6546083	1.729677	0	26

We have applied a natural logarithm on the crime rate to make the affect of variables more explainable in terms of percentages. The variables 'church' and 'police' have been result of count of churches and police officers per each county divided by the population of that county. So, their values are smaller compared to other variables and we expect their coefficients to be relatively larger.

After some exploratory analysis we found that the county with the highest crime reported has the median salary of 61,906 and median age of 34.4 years which is quite not a surprise. It is

astonishing that the county with the highest unemployment rate is having the crimes reported are 797 and it has the highest number of white populations in the state which is 0.9856%. The county with the highest population has also have the highest crime rate which is Harris County.

The county with the highest median income Rockwall has a population just over 100,000 or more specifically, 116,38, with a median age of 37.5 and the police-to-population ratio is quite low. Now we observe the county with the highest median age which is 60 has a very few population that is 1949 and the crime reported is also very less that is 13. The Brooks county which has the highest poverty rate is 0.287% has the quite low median income that is 33,513 and contains the white and Hispanic population comes near to 2% combined.

Jefferson county has most of the black population is 48,808 which is quite low as compared to the market standards. La Selle county which is has the more white, black, and Hispanic population has 44 members crime reported. The highest dropout rate in the data set is 13% which is Bell County with the median age of 31.9 years and has 8,428 crimes reported. The county with the highest church to population ratio has 0 crime rate which is fascinating and has lower increasing crime rate. It's quite uncommon we have to see in the data set in which is Dallas County with the highest police to population ratio has the most recorded crimes and also with the highest population.

Empirical Method :

The overarching question is to determine the effects of factors on crime rate. We will use a Multiple Linear regression method to regress the crime rate in each county of Texas over the data of all the socio-economic factors that we have collected by county.

It is difficult to say exactly which socio-economic factors are most strongly correlated with crime rates, as this can vary depending on the specific context and location. However, some factors that have been commonly associated with higher crime rates across multiple research projects were poverty, unemployment, low levels of education, and social inequality. These factors may create conditions that are conducive to criminal behavior, such as a lack of opportunities for legitimate income, inadequate support systems, and a lack of social integration. So, we wanted to check what was the correlation between each of the variables in our Data across all the Texas Counties. We did this by running a correlate function in Stata.

Figure 1

As seen (Figure 1) above the factors that are highly correlated (>0.4) are poverty – income; divorce rate – high school drop our rate. Poverty is negatively correlated with income and that is expected since high median income in a county means there is lower poverty rate. It is interesting to note that divorce rate and dropout rate have a significant correlation.

Before running the regression analysis, we examined the data and checked if it satisfies Gauss-Markov assumptions.

Assumption MLR.1. Linear in Parameters

The model satisfies the first Gauss Markov assumption of linearity because all coefficients are constants multiplied by an explanatory variable.

Assumption MLR.2. Random Sampling

As most of the data we collected came from Census, we can assume that it was randomly sampled or with proper sampling techniques.

Assumption MLR.3. No Perfect Collinearity

According to figure 1, none of our explanatory variables are perfectly correlated with one another, thus meeting the third assumption of non-collinearity.

Assumption MLR.4. Zero Conditional Mean

This means that the expected value of the error term, u will be 0 for any independent variable values. There are likely other factors not included in the dataset impacting the correlation between income and

cost of living. As a result, we will assume that this is true, although it is a difficult assumption to make.

Assumption MLR.5. Homoskedasticity

This assumption is that our error term has a constant variance. While there may be outliers for the data which impact the error term, we will assume that it has a constant variance

The regression results are as follows

Figure 2

Below are the plots for crime rate against the factors that are highly significant.
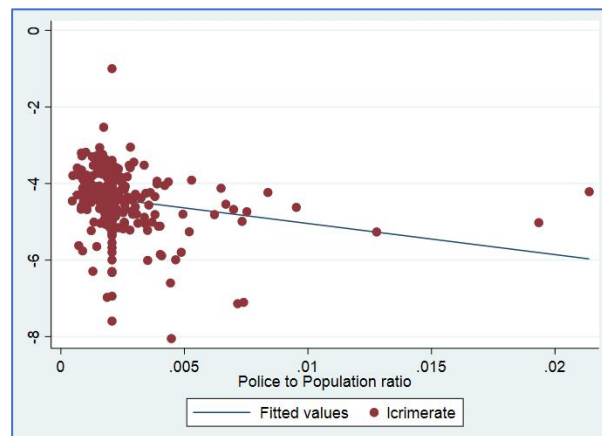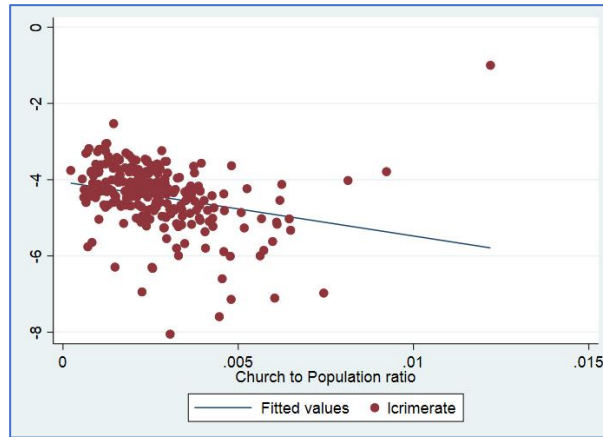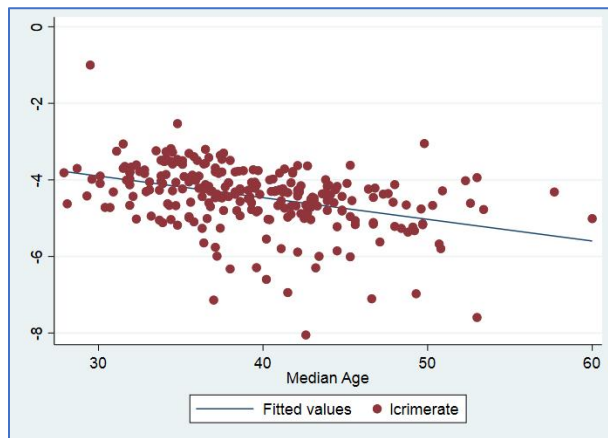

Figure 3

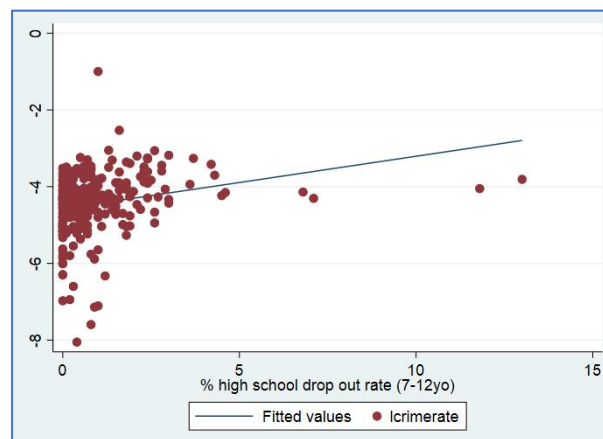*Figure 4*

Figure 5

Figure 6

Conclusion:

In summary, we can not reject the null hypothesis that church and Divorce rate in 2010 has no effect on crime rate in Texas in 2021 as evidenced by the corresponding p-values produced by our analysis. However, we can conclude that Median Age,No. of Police officers per person, High school Dropout rate(7-12 y.o) do have an impact on a crime rate in Texas in 2021 as these variables had p-values reflecting statistical significance.

Throughout our analysis of crime rate in Texas and unemployment rate, we can see that there is a slight positive correlation among the two variables of interest. This supports our original hypothesis stating that high unemployment rate will be correlated with increasing crime rate in Texas. The unemployment rate by itself, however, does not accurately predict the crime rate levels accurately.The research also suggests that in areas with a higher church and higher No. of Police officers per person, crime rate will likely be lower for the other areas from the respective area while holding other factors.

As such, our results concur with the existing body of literature as we failed to show that Divorce rate in 2010 has a significant impact on crime rate in Texas in 2021 . One possible explanation for this phenomenon could be related to the fact that divorce has impacts on children growth for younger children in relatively lower income level family , with the impact overall social atmosphere, the increase of income level and school education support, and the children become more self-sufficient,the relationship become not so significant.

Overall, as a result of the research and analysis conducted, we can infer that there may be more significant factors that influence crime rate level in Texas in 2021. The observations stated in this paper may be of use to persons hoping to find the potential factors effected the crime rate and hoping to properly reduce the crime rate by taking efficient measures,such increasing the amount of police officers and employment rate.

Bibliography:

(Adams, A. V., and Mangum, G. L. (1978). The Lingering Crisis of Youth Unemployment, W.E. Upjohn Institute for Employment Research, Kalamazoo, MI.)

(Phillips, L., Votey, H. L., and Maxwell, D. (1972). Crime, youth, and the labor market. J. Pol. Econ. 80: 491–504.)

Draca, Mirko, and Stephen Machin. "Crime and Economic Incentives." *Annual Review of Economics*, vol. 7, no. 1, 2015, pp. 389–408., https://doi.org/10.1146/annurev-economics-080614-115808.

The Lag Structure of Unemployment (Time Series Analysis of Crime Rates, Article in Journal of Quantitative Criminology · December 2001 DOI: 10.1023/A:1012507119569 David Greenberg New York University) https://link.springer.com/content/pdf/10.1023/A:1012507119569.pdf

Report Link: https://swer.wtamu.edu/sites/default/files/Data/Gao.pdf

Report link:
http://pricetheory.uchicago.edu/levitt/Papers/LevittTheChangingRelationship1999.pdf

Mok, Pearl L H et al. "Experience of Child-Parent Separation and Later Risk of Violent Criminality." American journal of preventive medicine vol. 55,2 (2018): 178-186. doi:10.1016/j.amepre.2018.04.008

Light, Michael T., and Casey T. Harris. "Race, Space, and Violence: Exploring Spatial Dependence in Structural Covariates of White and Black Violent Crime in US Counties." Journal of Quantitative Criminology, vol. 28, no. 4, 2012, pp. 559–586., https://doi.org/10.1007/s10940-011-9162-6