Economic Discussion Papers

Efficiency Series Paper 1/2016

Modelling the effect of crime on economic activity: The case of Mexican states

Antonio Álvarez, Rafael Garduño, Héctor Núñez

Departamento de Economía

Universidad de Oviedo

Available online at: http://economia.uniovi.es/investigacion/papers

Modelling the Effect of Crime on Economic Activity: The Case of Mexican States[✧]

Antonio Álvarez[a], Rafael Garduño[b] and Héctor Núñez[b *]

Abstract
We estimate the technical efficiency of Mexican states using stochastic production frontier models. In particular, we study the effect of crime on efficiency. The empirical section uses panel data over the period 1988-2008. A distinctive feature of the paper is the use of socioeconomic and location data in order to control for the heterogeneity of the states. The main contribution of this paper is to test for the existence of a threshold effect of crime. We find that crime rate negatively affects the efficiency of the states and that its effect is only significant after a certain level.

JEL Classification: D24, O18, R11

Keywords: Regional efficiency, stochastic frontier, Mexico, panel data, crime.

1. Introduction

The study of regional efficiency provides valuable information about the factors that explain the observed differences in productivity across regions in a country. This is important in Mexico where regional inequalities are substantial, with empirical evidence showing that most of the economic growth has been concentrated in regions near the U.S. Mexican border (Baylis, Garduño-Rivera and Piras, 2012).

One of the main drivers of economic growth is the improvement in technical efficiency. Nowadays, it is common to use production frontiers to measure (in)efficiency in terms of distance from the technological frontier.[1] In particular, some papers have estimated the technical efficiency of Mexican states using both parametric (e.g., Chávez and Fonseca,

[✧] The authors wish to thank helpful comments by Carlos Arias, Alberto Gude, Luis Orea, David Roibás and Alan Wall. This research was partially funded by the Government of the Principality of Asturias and the European Regional Development Fund (ERDF).

[a] Oviedo Efficiency Group, School of Economics and Business, University of Oviedo.
[b] Centro de Investigación y Docencia Económicas (CIDE), Aguascalientes, México.

[*]Corresponding author.
E-mail address: alvarez@uniovi.es (A. Álvarez).

[1] See Alvarez and Arias (2014) for a recent survey.

2012) and non-parametric approaches (e.g., Bannister and Stolp, 1995) to determine the frontier.

In this paper we follow a parametric approach and estimate two stochastic frontier models that allow us not only to calculate the level of technical efficiency of each state but also to study the variables that best explain the differences in technical efficiency across states. We use a panel dataset of the 31 Mexican states and Mexico City. The data come from the economic censuses undertaken by the Mexican Statistical Institute (INEGI) every five years. A distinctive feature of our paper is the use of a broad set of state characteristics that are expected to pick up most of the observed differences between the states. These variables reflect differences in human capital, public capital, productive specialization, location and business environment.

The recent economic literature has lent prominence to the role played by institutional variables in explaining economic growth. The work of North (1990) and, more recently, Acemoglu and Robinson (2008) show that the quality of the institutional framework is a key determinant of economic growth at the country level.

Institutions within a country are more homogeneous than across countries, but there are important differences in some relevant variables such as the business environment. One of the variables that affect this environment is the number of crimes. In this paper we will study the effect of violent crime (i.e., homicides) on the productive efficiency of the Mexican states. An important contribution of the paper is to develop a model specification that allows testing for the possible existence of a threshold effect of crime. That is, we want to check if the effect of crime is monotonous or there is a minimum level (threshold), below which, increases in crime do not affect economic activity.

The structure of the paper is as follows. In section 2, we review the relatively small literature that looks at the effects of crime on regional performance. Section 3 presents the stochastic frontier models. In section 4 we review the literature that uses stochastic frontiers to estimate regional efficiency in Mexico. Section 5 describes the data and the empirical models. In section 6 we present the estimation results. Section 7 discusses the estimated efficiency of the states. Section 8 contains some conclusions.

2. Crime and regional economic performance

The macroeconomic literature is increasingly assigning a key role to institutional variables when it comes to explain differences in economic growth. Most studies that analyze the role of institutional variables are based on cross-country data. At the regional level, it is less common to use institutional-type variables, since regions within a country generally share many institutions. One institutional variable that differs widely across regions in some countries is crime. This variable, and in particular violent crime, is expected to deteriorate the business environment by inhibiting investment and shifting economic activity from more conflictive regions to less conflictive ones. The empirical literature, while not very large, has originated in countries where organized crime has a pervasive presence such as Italy (with its mafia) and Mexico, where drug-related violence has become a real problem (Vilalta, 2014).

However, the empirical results concerning the negative role of violent crime are not conclusive. Ascari and Di Cosmo (2005) include the crime rate as an explanatory variable of the differences in Total Factor Productivity across Italian regions, finding that crime is not significant. However, when splitting the sample into northern and southern regions they find that crime is significant, with a negative sign in northern regions but a positive sign in southern regions, which is rather counterintuitive due to the widespread presence of mafia in southern Italy. Rincke (2014) shows that the crime rate negatively affects the growth of per capita income in US Metropolitan Statistical Areas. Interestingly, when separating crime into crime against property and violent crime, the findings show that only property crime reduces per capita income growth. Partridge and Rickman (1999) do not find that crime is significant in explaining differences of labor productivity across US states, although it seems to have a significant effect in the regional industry mix.

Even though there has been a broad literature that analyzes the social and quality-of-life consequences of drug-related crime in Mexico (Vilalta, 2014), little has been studied about the effect of crime on the regional efficiency. Weiss and Rosenblatt (2010) look at the role played by corruption and crime in state performance over the period 2001-2005. For corruption they use an index built by the non-governmental organization *Transparencia Mexicana* and they proxy crime by the number of offences reported to the police. Their findings indicate that corruption does not affect per capita GDP growth while crime is significant but with a surprising positive sign.

Pan Widner and Enomoto (2012) study the relationship between per capita GDP growth in Mexican states and crime using a spatial auto-regressive model that takes into account the crime rate of neighboring states. As in the paper by Weiss and Rosenblatt, they find that the total number of crimes in a state (lagged one year) positively and significantly affects the annual rate of state growth. However, if they use the homicide rate instead of the total crime rate they find that it is not significant.

Enamorado López-Calva and Rodríguez-Castelán (2014) look at the β-convergence of per capita income across Mexican municipalities during the period 2005-2010. They find that the initial crime rate is not significant but when they break down the homicide rate into drug and non-drug related crimes, results show evidence of a negative impact of drug-related crimes on income growth.

Finally, Cabral Varella Mollick and Saucedo (2016) examine the evolution of labor productivity across Mexican states during the period 2003-2013 finding that the total number of crimes in each state has a negative effect on labor productivity, but only during the “war on drugs” period (i.e., after 2007).

This literature review shows that most of the studies about the effect of crime on economic activity have either found that crime is non-significant or that crime rate has a positive effect. This last result was obviously not only unexpected but it is contrary to economic logic. Since crime data, especially that pertaining to violent crime, can be considered accurate, we think that this result may be driven by modelling and specification choices. The first one is the absence of enough control variables to account for the strong heterogeneity that exists across regions. For this reason, we include in our model a rather large set of regional characteristics, such as regional dummies, state geographical location, education level, infrastructures and productive specialization.

The second one is related to the specification of the crime variable in the production function. In the previous literature, the crime rate has been included in the estimated models as just another regressor. However, one can think that crime has a threshold level below which the variable has no effect on economic activity. That is, we would expect that at very low levels of crime its effect is negligible but after some point (threshold level) the economic agents start reacting. In the empirical section, we investigate the possible existence of this threshold effect.

3. Modeling technical efficiency

Technical efficiency is defined as the ratio between observed production and potential production, i.e. production on the frontier, given a set of inputs. We follow the stochastic frontier approach (Aigner, Lovell and Schmidt, 1977) in order to estimate the technical efficiency of Mexican states. Our basic model is a stochastic production frontier, which can be written as:

$$y_{it} = x_{it}\beta + v_{it} - u_{it} \tag{1}$$

where y is output, x are the inputs, v is random noise and u is a non-negative stochastic term that is assumed to be independent from v and capture distance from the frontier, i.e. technical inefficiency. When $u=0$, the observation lies on the technological frontier and is therefore efficient. When $u>0$, the observation is below the frontier, indicating that it is technically inefficient.

Since we are interested in finding which variables explain the efficiency of states, we estimate several models that modify equation (1) by allowing the inefficiency term u to be a function of some exogenous variables z. The general form of this type of models is:

$$y_{it} = x_{it}\beta + v_{it} - u_{it}(z_{it}) \tag{2}$$

There are two possible alternative specifications of $u_{it}(z_{it})$, depending on the way that the variables z affects the distribution of u. In particular, they can affect the mean or the variance of u. In this paper we use two models in order to explore these possibilities.

a) Modelling the mean of the inefficiency term

Kumbhakar, Ghosh and McGuckin (1991), and Huang and Liu (1994) were the first papers that attempted to model the inefficiency term in a stochastic frontier framework. Their approach consists of making the mean of the distribution of the inefficiency term depend on a set of exogenous variables. While these models were originally developed for cross-sectional data, Battese and Coelli (1995) (from now on referred as BC95) extended this approach to accommodate panel data. In the BC95 model the inefficiency term is assumed to follow a truncated normal distribution where the mean of the pre-truncated distribution of u_{it} depends on a set of exogenous variables, z. That is,

$$u_{it} \sim N^{+}(\mu_{it}, \sigma_u^2) \tag{3}$$

where superscript "+" indicates truncation of a distribution from the left at zero and $\mu_{it} = \delta \cdot z_{it}$. In this way, we ensure that $u_{it} \geq 0$.

b) Modeling the variance of the inefficiency term

Reifschneider and Stevenson (1991) was the first paper to incorporate heteroscedasticity in the stochastic frontier model. Caudill, Ford and Gropper (1995) (from now on referred as CFG95) assumed that u exhibits multiplicative heteroscedasticity, a choice that we will use in this paper. In particular, the CFG95 model suggests an exponential function:

$$u_{it} \sim N^{+}(0, \sigma_{it}^2) \tag{4}$$

where $\sigma_{it}^2 = g(z_{it}, \delta) = \sigma_u \cdot \exp(\delta z_{it})$. Modeling the variance of the one-sided error term is very important since the presence of heteroscedasticity in u_{it} will yield biased estimates of both the frontier parameters and the efficiency scores. This result differs markedly from the typical effect of heteroscedasticity in the two-sided error term v_{it}, which causes the variances of the parameter estimates to be biased.

Given the importance of controlling for possible heteroscedasticity in the variance of u_{it}, our preferred model will be the CFG95 model. However, we will also estimate the BC95 model for comparison purposes.

4. Regional efficiency in Mexico

The use of stochastic frontiers has been quite common to study regional efficiency after early work by Beeson and Husted (1989). In the case of Mexico, very few papers have explicitly studied regional efficiency following a parametric approach, but none of them has referred to the role of crime. Becerril et al. (2009) estimate a stochastic frontier using data from the federal entities. They use the BC95 model in order to analyze the effect of infrastructures on state efficiency and find that Nuevo León, Mexico City and the State of Mexico are on the efficient frontier. A sigma convergence analysis shows that the disparities among the states have been declining over time. They divide their period of analysis according to the two main economic policies of each period, namely Import-substitution Industrialization (1970-1985) and Export-oriented Industrialization (1988-2003). They conclude that convergence in technical efficiency and the effect of the infrastructure variable were more pronounced during the import substitution period.

Aguilar (2011) estimates the BC95 model for a sample of 21 municipalities with data for five economic sectors during the years 2006-2008. Interestingly, the author includes a trend in the inefficiency term and obtains a positive sign for all sectors, which implies that the technical inefficiency of Mexican municipalities increased over the sample period, although it is significant for only two of them.

Braun and Cullmann (2011) use a panel dataset of regional production data from the manufacturing sector at the municipality level. They estimate both the BC95 and the True Random Effects (Greene, 2004) models. They report significant disparities in the efficiency scores across municipalities, finding that northern states operate more efficiently than the southern ones.

Chávez and Fonseca (2012) estimate a stochastic frontier model for the manufacturing sector using data at the state level in order to analyze the regional disparities that exist in Mexico. Covering the years 1988, 1993, 1998, 2003 and 2008, they estimate a translogarithmic production frontier using the model proposed by Battese and Coelli (1992) and analyze beta and sigma convergence. They find a steady increase in the levels of technical efficiency from 53.7% to 76.4%, as well as the existence of beta and sigma convergence, i.e. efficiency gaps are closing as more inefficient states are becoming more productive. However, there are still marked differences between Central and Northern states and those in the South.

5. Data and empirical models

We use a balanced panel dataset of the 32 Mexican states, including Mexico City, during the period 1988-2008. The basic data (output, labor and capital) come from the economic censuses carried out by INEGI (National Institute of Statistics and Geography) every five years. We have information for five censuses, where the first one reports information of year 1988 from all Mexican formal economic units, excluding those in the agricultural sector, followed by censuses in years 1993, 1998, 2003 and 2008. As a result, the panel dataset consists of 160 observations (32 entities times 5 years).

Output (Y) is gross value added, without the mining sector[2]. Private capital (K) is measured as the total stock of fixed assets. Both variables are deflated using the producer price index reported by INEGI with December 2010 equal to 100. Labor (L) is the total number of workers.

In order to account for cross-state heterogeneity, we include several control variables. Human Capital is typically included in regional production functions in order to account for labor heterogeneity. As a proxy, the average years of education of labor in each state (*EDUCATION*) is used. The information was gathered from population censuses of 1989,

2 When analyzing the data, we noticed a sharp change in the mining' gross value added (GVA) from one census to the next. In talks with the INEGI, they mentioned that the data collection method for the mining sector had changed over the years: first, they assigned the mining' GVA to the states where mining offices were located (i.e. PEMEX in Mexico City) and at some point it was assigned to the state where the extraction was carried out. Therefore, we decided to take out the mining sector entirely.

1994, 1999, 2004 and 2009 by INEGI. Public capital has also been considered a relevant variable to explain differences in regional economic performance (e.g., Puig-Junoy, 2001). We have accounted for differences in state infrastructures by including the length of the roads network divided by the state area (*INFRASTRUCTURE*).

Additionally, we account for the different productive specialization of the regions by means of a Specialization Index (SI) following Alvarez (2007), which is computed as follows:

$$SI_i = \sum_{j=1}^{3} \left(\frac{VA_{ji}}{VA_i} - \frac{VA_{jN}}{VA_N} \right)^2 \tag{5}$$

where VA is Value Added, subscript *j* denotes sector (Commerce, Manufacturing and Services), *i* represents state and *N* indicates that the value refers to the national average. This index is zero when the regional productive structure is equal to the national average and increases with the level of specialization.

The model includes also a time trend (*TREND*), and its square (*TREND2*), which are expected to control for technical change. Finally, we include a set of regional dummy variables to account for time-invariant unobserved regional heterogeneity.

Following the Mexican Central Bank (Chiquiar, 2008), we divide the Mexican territory into seven groups:

- U.S. **Border** (Baja California, Coahuila, Chihuahua, Nuevo León, Sonora and Tamaulipas)
- **Capital** (Mexico City and the State of Mexico)
- **Center** (Colima, Guanajuato, Hidalgo, Jalisco, Michoacán, Morelos, Puebla, Querétaro, Tlaxcala, and Veracruz)
- **North** (Aguascalientes, Baja California Sur, Durango, Nayarit, San Luis Potosí, Sinaloa and Zacatecas)
- **Oil** (Campeche and Tabasco)
- **Peninsula** (Quintana Roo and Yucatán)
- **South** (Chiapas, Guerrero and Oaxaca)

Chávez and Fonseca (2012) used four regional areas (north, north-central, south and central). But in order to better capture the heterogeneity of the states we consider, first, the capital and the surrounded State of Mexico as a separate region. Second, we divide the north in two parts (border and north). Finally, we divide the South region in three parts (Oil, Peninsula, and South) to isolate the fast growing oil-producing area from the most touristic area in the Peninsula and the poor and slow growing area of the south.

In the inefficiency term we have included four exogenous variables, apart from a constant term. First, to capture the effect of the North American Free Trade Agreement we include a dummy variable (*NAFTA*) that equals 0 for periods before NAFTA (1988 &1993) and 1 for periods after NAFTA (1998, 2003 and 2008). Second, the average years of education of labor in each state (*EDUCATION*) is used. Third, the location of each region is expected to affect inefficiency. In particular, in Mexico there is a notorious difference between North and South, with the regions located in the southern part of the country being less developed than those located in the North. To account for this fact, we include the distance to the US border, which we reflect with a variable (*DISTANCE_US*) that is measured as the road distance in kilometers between the state capital and the US border. Finally, to account for the economic environment in which business is carried out, we include the crime rate of each state (*CRIME*), measured as total number of homicides per 100 thousand inhabitants, and gathered from INEGI (2013). Some summary statistics of the variables used in the empirical analysis are displayed in Table 2.

Given our interest in analyzing the effect of violent crime on state performance, we describe with some detail the evolution of this variable over time as well as the differences across states. Figure 1 shows the evolution of the crime rate in México. After achieving a reduction in the crime rate to 8 homicides per 100,000 inhabitants in 2007, it skyrocketed in the following years to 24 homicides per 100,000 inhabitants. This tremendous increase was due to President Felipe Calderon´s Drug War. As soon as he took office (December 1st, 2006), he declared "war" against all the Mexican drug cartels, with the result being an increase in the crime rate year by year until he left office (November 30th, 2013) when the death toll reached 121,000 people killed. This is in stark contrast to the terms of the previous three presidents before Felipe Calderon when Mexico was curiously peaceful, managing to decrease violent crime from 1990 to 2006 by coming to agreements with the different cartels and encouraging an atmosphere of peace. The dramatic increase in the crime rate from 2007 was due to increased participation in drug trafficking rings (Vilalta, 2014). This growing violence has sparked the interest to analyze the effects of crime on economic activity.

Crime differs significantly across Mexican states, following a clear spatial pattern in which main clusters of crime have been traditionally located in the north of the country and along the Pacific coast. Table 1 and Figure 2 show that in 2008 the states with the highest crime rate were concentrated in the Pacific Ocean (Baja California, Guerrero, Michoacán, Nayarit, Oaxaca, Sinaloa, and Sonora) and near the U.S. border, Durango and Chihuahua. The states located along the Pacific Ocean, with good communication to the Asian countries and Colombia, make them an ideal destination for the landing of different drugs and therefore the starting point for distribution within Mexico and the United States. It is important to highlight that the ranking of states in terms of crime rate is relatively stable. Of the ten states with the highest (lowest) crime rate in 1988, seven (eight) were located in that group in 2008.

The final model to be estimated is a restricted translog[3] stochastic production frontier of the following form:

$$\begin{aligned}\ln y_{it} &= \beta_0 + \beta_1 \ln K_{it} + \beta_2 \ln L_{it} + \beta_3 \ln K_{it} * \ln L_{it} + \frac{1}{2}\beta_4 \ln K_{it}^2 + \frac{1}{2}\beta_5 \ln L_{it}^2 \\ &+ \sum_{j=6} \beta_j \ln x_{jit} + \gamma_t t + \gamma_{tt} t^2 + \theta_r REG_r + v_{it} - u_{it}(z_{it})\end{aligned} \tag{6}$$

where subscript *i* indexes states, *t* indexes time, *r* indexes region, *K* is private capital, *L* is total number of workers, *x* is a set of control variables (*EDUCATION, INFRASTRUCTURE* and *SPECIALIZATION*), *t* is a time trend that takes values 1-5, and *REG* are regional dummies. The noise component v_{it} is assumed to be normally distributed with mean zero and constant variance, while the inefficiency component u_{it} will follow a truncated normal distribution for the model BC95 and a half normal for the model CFG95.

6. Estimation and results

Table 3 presents the maximum likelihood estimates of the stochastic production frontier models described in section 2. First of all, it is important to highlight that the results are very similar across the two models.

Frontier

In the frontier part, all explanatory variables (except infrastructure) are significant and carry the expected signs. Since basic inputs were divided by its geometric mean, the first order coefficients can be interpreted as the output-elasticities at the geometric mean of the sample. As expected with macro data, the elasticity of labor (0.81) is much larger than the elasticity of capital (0.34), giving a value of 1.16 for the scale elasticity, which allows to reject

[3] In order to mitigate multicollinearity problems, we restrict the coefficients of the squared and cross products of the control variables to be zero.

the hypothesis of constant returns to scale in capital and labor at the geometric mean of the sample, as shown in the last row of Table 3.

The estimated coefficients of the control variables are significant (except infrastructures) and carry the expected sign. Education and the specialization index are positive, indicating that states with more educated labor force or where the economic activity is more specialized, produce more output.

The linear term of the trend variables is positive while the squared term is negative (and smaller). This implies that neutral technical change is decreasing over time. In fact, it is 1.1% in the first period but it becomes negative already in the third period. Finding negative technical change is not appealing but it is not uncommon, especially when using stochastic frontiers with time-varying technical inefficiency. Our interpretation is that these models have some difficulties separating two different but similar effects: an outward shift of the technological frontier and a catching up effect of most regions towards the frontier. In the case of México, other researchers have also found that neutral technical change is negative (e.g., Chávez and Fonseca, 2012).

The regional dummies are significant and negative, indicating that the (time-invariant) unobserved characteristics of the regions make them different from the oil region, which is the excluded category. The south region effect is the closest one to the oil region, but it is not significant.

Inefficiency term

The value obtained for λ (0.828 in CFG95 and 1.518 in BC95), which is equal to the ratio between the standard deviations of inefficiency and statistical noise (i.e. σ_u/σ_v), indicates that in both models inefficiency explains part of the difference in production across regions which is not accounted for by the explanatory variables.

The NAFTA dummy variable in the inefficiency term is positive and significant in both models (but only at 10% in BC95) indicating that after the NAFTA agreement some unobserved factors (common to all states) are causing state inefficiency to increase. This result gives support to the claim of Krugman and Livas-Elizondo (1996) and Rodríguez-Pose and Sánchez-Reaza (2005) that trade has decreased efficiency in Mexico. They find that trade liberalization has not been homogeneous across regions because of economies of scale and transportation costs: states along the U.S. border have benefited from trade, but not the rest. However, this result has to be taken with caution, since the effect of NAFTA is probably confounded with other important events of the Mexican economy common to all states, such as the tequila crisis.

The *EDUCATION* variable carries the expected negative sign, indicating that investment in human capital reduces inefficiency in production (although it is not significant in both models).

As expected, the effect of the distance to the U.S. border (*DISTANCE_US*) is positive and significant. This implies that states near the U.S. border are more efficient than those states further away. This result is in line with Rodríguez-Pose and Sánchez-Reaza (2005), who found that states closer to the U.S. border grew faster than others.

Finally, our main variable of interest, the crime rate, is positive and significant in both models. The positive sign indicates that the higher the crime rate, the more inefficient the state is. This result is consistent with the commonly held view that crime does not provide a good economic environment to carry out business.

We now proceed to investigate if there is a threshold effect of crime. The intuition is that at low levels of crime rate, the marginal effect of an increase in crime is not significant. For this purpose, we first sort the crime rate variable following an increasing order and then compute the deciles of the sorted variable. Next we create dummy variables for some of these deciles and interact the crime variable with them. For example, we create a dummy variable DECILE_2 that takes the value one if the state is in the 2nd decile (zero, otherwise) and interact this dummy with the crime rate. The new variable will therefore have zeroes for all states with a crime rate larger than the state that marks the upper limit of the interval for the 2^{nd} decile and the crime rate for the rest of the states.

Our empirical specification includes the interaction of the decile dummy with the crime rate together with the interaction of the dummy for the complementary decile with the crime rate. Our intention with this approach is to split the effect of the crime rate into two variables, one for the states with the lowest crime rate (as defined by the Q_j decile) and another variable for the rest of the states. We expect that the estimated coefficient of the variable associated with lower crime rates will have a smaller (possibly non-significant) coefficient than the variable with the larger crime rates.

In Table 4 we show the results of estimating models using two different deciles: 10% and 60%.[4] The estimation is carried out only for our preferred model, the CFG95 model. We find that in the two cases the coefficient of the variable with the highest values of crime rate (which additionally is always significant) is always larger than the variable with the lower values. This result seems to suggest that individuals react to crime only at high levels of the crime rate. When the crime rates are small, they seem to have no effect on individuals. Therefore, the results are consistent with our hypothesis that there is a threshold effect of crime.

7. Evaluating the technical efficiency of Mexican states

We now proceed to analyze the estimated efficiency of each state. Maximum likelihood only provides an estimate of the composed error term. However, using the conditional expectation of u_{it} on v_{it} - u_{it} , we can recover an estimate of u_{it}.

Since the output variable is in logs, the output-based Farrell technical efficiency index can be calculated as:

$$TE_{it} = \exp(-\hat{u}_{it}) \tag{7}$$

Since the inefficiency term u_{it} varies across states and over time, there is an estimate of u for each state in each year. To better summarize this information, in Table 5 we show the estimated initial (i.e. 1988) and final (i.e. 2008) efficiency indexes for the CFG95 model as well as the efficiency change and the state ranking according to this variable. A positive change in technical efficiency implies a movement towards the technological frontier and can therefore be interpreted as evidence that the state is "catching-up" with the best practice frontier.

Overall, the average technical efficiency of the country was reduced by 5.3% between the two periods. Only three states increase their technical efficiency index during the sample period in the CFG95. For the initial and final year, the most efficient state is Baja California, which shares border with California. On the other hand, the least efficient state in 1988 was Michoacán, which also had the highest growth in efficiency. Not surprisingly, Guerrero became the least efficient state in 2008 after having the largest decrease in technical efficiency over the sample period. This state depends mainly on commercial and tourist activities. Acapulco, which is the largest city and main tourist destination in Guerrero, has

[4] We have done the regression for the ten deciles, but we only present those where the split was significant, either the lower or upper bound. The other decile-regressions are available upon request.

reported a record economic decline due to a large decrease in tourist activities, mainly caused by predominance of illegal activities in the region.

To better illustrate the results, we show a quantile map for the estimated technical efficiency index of each state in 2008 in Figure 3 for the CFG95 model. The map shows that the US border effect is a key driver of efficiency since those states on the border report the highest technical efficiency index, while two out of the three poorest states, namely Oaxaca and Guerrero, located in the south of the country, report the lowest technical efficient indexes.

8. Conclusions

This paper studies the regional efficiency of the 31 Mexican states and Mexico City by estimating two production frontier models in order to measure inefficiency as distance from the technological frontier. Besides labor and capital, we incorporate a number of state characteristics that are expected to pick up most of the observed differences between the states across time. Likewise, we model the possible heteroskedasticity in the inefficiency term in order to reduce the risk of bias of the estimators.

We follow the parametric approaches proposed by Battese and Coelli (1995) and Caudill, Ford and Gropper (1995) to estimate the level of technical efficiency as well as the variables that explain best the differences in technical efficiency across states.

We present several important findings. First, we confirm that the Mexican economy (excluding the agricultural and mining sectors) reports constant returns to scale in labor and capital. Second, we find that characteristics such as education and productive specialization contribute to increase output. Third, we find that inefficiency in both models explains part of the difference in production across regions, which is not accounted for by the explanatory variables. Last, our study determines the efficiency of each state with respect to the frontier. Comparing the initial and final year in our sample, we find that Tabasco is the state that has reduced inefficiency the most.

With regards to the determinants of state inefficiency, we find that the distance to the US border and the crime rate increase the level of inefficiency. From a policy point of view, one of the most interesting findings of our paper is that after the NAFTA agreement the efficiency of the states has declined. While our model does not provide additional information of the reasons behind this finding, the results seems to agree with some previous empirical findings.

According to our expectations, we have found the existence of a threshold effect of crime. In fact, when splitting the crime rate into two - one for the states with high crime rate and another for the states with low crime rate - we find that the coefficient of the variable for the high crime rate is always larger than the other one. This seems to indicate that the economic performance of the state in terms of technical efficiency only worsen at high levels of crime rate. This result needs to be further tested in order to assess its robustness.

9. References

Acemoglu, D. and J. Robinson, (2008), “The role of institutions in growth and development. Commission on growth and development”, *Working Paper No. 10, The World Bank*.

Aguilar Gutiérrez, G., (2011), “Eficiencia industrial en las regiones de México”, *EconoQuantum*, 7(2), 93-113.

Aigner, D., C.A.K. Lovell and P. Schmidt, (1977), “Formulation and estimation of stochastic frontier production function models”, *Journal of Econometrics*, 6(1), 21-37.

Alvarez, A. and C. Arias, (2014), “A selection of relevant issues in applied stochastic frontier analysis”, *Economics and Business Letters*, 3(1), 3-11.

Alvarez, A., (2007), “Decomposing regional productivity growth using an aggregate production frontier”, *Annals of Regional Science*, 41(2), 431-441.

Ascari, G. and V. Di Cosmo, (2005), “Determinants of total factor productivity in the Italian regions”, *Scienze Regionali/Italian Journal of Regional Science*, 4(2), 27-49.

Bannister, G.J., and C. Stolp, (1995), “Regional concentration and efficiency in Mexican manufacturing”, *European Journal of Operational Research*, 80(3), 672-690.

Battese, G.E. and T.J. Coelli, (1992), “Frontier production functions, technical efficiency and panel data: With applications to paddy farmers in India”, *Journal of Productivity Analysis*, 3(1-2), 153-169.

Battese, G.E. and T.J. Coelli, (1995), “A model for technical inefficiency effects in a stochastic frontier production function for panel data”, *Empirical Economics*, 20(2), 325–332.

Baylis, K., R. Garduño-Rivera, and G. Piras, (2012), “The distributional effects of NAFTA in Mexico: Evidence from a panel of municipalities”, *Regional Science and Urban Economics*, 42(1-2), 286-302.

Becerril-Torres, O.U., I.C. Álvarez-Ayuso, and L.E. del Moral Barrera, (2010), "Disparidades en eficiencia técnica e influencia de las infraestructuras sobre la convergencia en eficiencia en México", *Investigaciones Regionales*, (17), 51-69.

Beeson, P.E. and S. Husted, (1989), “Patterns and determinants of productive efficiency in state manufacturing”, *Journal of Regional Science*, 29(1), 15-28.

Braun, F.G., and A. Cullmann, (2011), “Regional differences of production and efficiency of Mexican manufacturing: An application of nested and stochastic frontier panel models”, *Journal of Developing Areas*, 45(1), 291-311.

Cabral, R., A. Varella Mollick and E. Saucedo, (2016), “Violence in Mexico and its effects on labor productivity”, *Annals of Regional Science*, 56(2), 317-339.

Caudill, S.B., J.M. Ford and D.M. Gropper, (1995), “Frontier estimation and firm-specific inefficiency measures in the presence of heteroscedasticity”, *Journal of Business and Economic Statistics*, 13(1), 105-111.

Chávez, J.C., and F.J. Fonseca, (2012), “Technical and structural efficiency in Mexican manufacturing: A regional approach”, *Documentos de Investigación No. 2012-03, Banco de México*.

Chiquiar, D., (2008), “Globalization, regional wage differentials and the Stolper–Samuelson theorem: Evidence from Mexico”, *Journal of International Economics*, 74(1), 70-93.

Enamorado, T., L.F. López-Calva and C. Rodríguez-Castelán, (2014), “Crime and growth convergence: Evidence from Mexico”, *Economics Letters*, 125(1), 9-13.

Greene, W., (2004), “Distinguishing between heterogeneity and inefficiency: Stochastic frontier analysis of the World Health Organization's panel data on national health care systems”, *Health Economics*, 13(10), 959-980.

Huang, C.J. and J.T. Liu, (1994), “Estimation of a non-neutral stochastic frontier production function", *Journal of Productivity Analysis*, 5(2), 171-180.

INEGI. (2015). “Population Census 1989, 1994, 1999, 2004 and 2009,” *Instituto Nacional de Estadística y Geografía*. [Online] Available at: http://www.inegi.org.mx/est/contenidos/proyectos/ccpv/default.aspx

INEGI. 2008. “Economic Census 1993, 1998, 2003 and 2008,” *Instituto Nacional de Estadística y Geografía*. [Online] Available at: http://www.inegi.org.mx/est/contenidos/Proyectos/ce/Default.aspx

INEGI. 2013. “Anuario estadístico y geográfico por entidad federativa 2013,” *Instituto Nacional de Estadística y Geografía*. [Online] Available at: http://www3.inegi.org.mx/sistemas/productos/default.aspx?c=265&s=inegi&upc=702825054014&pf=Prod&ef=&f=2&cl=0&tg=7&pg=0&ct=106010000

Krugman, P. and R. Livas Elizondo, (1996), “Trade policy and the Third World metropolis,” *Journal of Development Economics*, 49(1), 137-150.

Kumbhakar, S.C., S. Ghosh and J.T. McGuckin, (1991), “A generalized production frontier approach for estimating determinants of inefficiency in U.S. dairy farms”, *Journal of Business and Economic Statistics*, 9(3), 279-286.

North, D.C., (1990), *Institutions, institutional change and economic performance*, Cambridge University Press.

Pan, M., B. Widner and C.E. Enomoto, (2012), “Growth and crime in contiguous states of Mexico”, *Review of Urban and Regional Development Studies*, 24(1-2), 51-64.

Partridge, M.D. and D.S. Rickman, (1999), “Static and dynamic externalities, industry composition, and state labor productivity: A panel study of states”, *Southern Economic Journal*, 66(2), 319-335.

Puig-Junoy, J., (2001), “Technical inefficiency and public capital in U.S. states: A stochastic frontier approach”, *Journal of Regional Science*, 41(1), 75-96.

Reifschneider, D. and R. Stevenson, (1991), “Systematic departures from the frontier: A framework for the analysis of firm inefficiency”, *International Economic Review*, 32(3), 715-723.

Rincke, J., (2014), “The effect of crime on regional economic performance: Evidence from U.S. metropolitan areas”, *Working Paper, University of Erlangen-Nuremberg*.

Rodríguez-Pose, A. and J. Sánchez-Reaza, (2005), “Economic polarization through trade: Trade liberalization and regional growth in Mexico”, in R. Kanbur and A.J. Venables (eds.), *Spatial Inequality and Development* (237-259) Oxford: Oxford University Press.

Vilalta, C., (2014), “How did things get so bad so quickly? An assessment of the initial conditions of the war against organized crime in Mexico”, *European Journal on Criminal Policy and Research*, 20(1), 137-161.

Weiss, E. and D. Rosenblatt, (2010), “Regional economic growth in Mexico. Recent evolution and the role of governance”, *Working Paper Series No. 5369, World Bank Policy Research*.

Table 1 - Crime rate (homicides per 100,000 inhabitants)

Entity	1988	Entity	2008
Oaxaca	40.1	Chihuahua	76.4
Mexico	35.1	Baja California	32.7
Michoacán	33.0	Sinaloa	29.6
Morelos	31.7	Guerrero	29.5
Nayarit	29.5	Durango	26.1
Guerrero	29.4	Sonora	16.3
Durango	21.4	Oaxaca	16.2
Sinaloa	20.2	Michoacán	15.1
Colima	18.2	Nayarit	14.3
Jalisco	15.7	Morelos	12.1
Baja California	15.7	Quintana Roo	10.9
Mexico City	15.4	Mexico City	10.5
Campeche	15.3	Mexico	10.4
Puebla	12.7	Colima	8.6
Chihuahua	12.5	Tamaulipas	8.1
Veracruz	12.5	San Luis Potosi	7.7
Tamaulipas	11.8	Jalisco	7.4
San Luis Potosi	11.7	Tabasco	6.9
Quintana Roo	11.6	Zacatecas	6.6
Zacatecas	11.4	Campeche	6.6
Hidalgo	10.7	Coahuila	6.4
Coahuila	9.2	Puebla	6.1
Sonora	9.1	Baja California Sur	6.0
Guanajuato	9.1	Chiapas	5.5
Tabasco	8.8	Guanajuato	5.4
Chiapas	8.5	Nuevo Leon	5.2
Queretaro	7.8	Aguascalientes	5.0
Tlaxcala	6.2	Tlaxcala	4.6
Aguascalientes	6.0	Veracruz	4.5
Yucatan	4.8	Queretaro	4.0
Baja California Sur	3.8	Hidalgo	2.8
Nuevo Leon	2.3	Yucatan	2.5

Source: INEGI (2009)

Table 2 - Summary Statistics of Variables in Equation (6)

Variable	Mean	Std. Dev.	Min	Max
Total Value Added (million MXN, 2010 Prices)	72,525.7	114,999.9	2,505.7	845,678.4
Total number of workers (LABOR)	371,386.8	429,011.7	29,024	2,849,557
Total stock of fixed assets (CAPITAL – million MXN, 2010 Prices)	90,815.7	115,831.5	2,628.0	780,460.6
Distance of the region to the U.S. border in kilometers (DISTANCE_US)	968.0	493.4	1.0	2,004.4
Average years of education of labor in each state (EDUCATION)	7.6	1.1	4.7	10.5
Specialization Index (SI)	0.07	0.08	0.001	0.40
Km of main roads per km^2 (INFRASTRUCTURE)	0.04	0.04	0.01	0.47
Crime Rate per 100,000 Inhabitants (CRIME)	13.1	10.1	2.3	76.4

Table 3 - Stochastic Production Frontier Estimation

		CFG95 Eq. (4)		BC95 Eq. (3)	
Variable	Par.	Estimate	t-ratio	Estimate	t-ratio
INTERCEPT	β_0	-1.647	[-3.06]	-2.014	[-4.06]
CAPITAL	β_1	0.346	[8.42]	0.341	[8.13]
LABOR	β_2	0.815	[15.52]	0.813	[15.52]
CAPITAL× LABOR	β_3	0.388	[2.64]	0.384	[2.67]
$CAPITAL^2$	β_4	-0.378	[-2.82]	-0.372	[-2.83]
$LABOR^2$	β_5	-0.441	[-2.59]	-0.447	[-2.63]
EDUCATION	β_6	1.053	[4.17]	1.189	[5.25]
SPECIALIZATION	β_7	0.040	[2.94]	0.036	[2.56]
INFRASTRUCTURE	β_8	0.001	[0.06]	-0.002	[-0.09]
TREND	γ_1	0.199	[3.74]	0.207	[3.50]
$TREND^2$	γ_2	-0.087	[-5.07]	-0.089	[-4.88]
REG_BORDER	θ_1	-0.401	[-4.33]	-0.368	[-3.76]
REG_CAPITAL	θ_2	-0.383	[-2.85]	-0.331	[-2.28]
REG_CENTER	θ_3	-0.309	[-4.28]	-0.295	[-3.99]
REG_NORTH	θ_4	-0.392	[-5.47]	-0.365	[-5.18]
REG_PENINSULA	θ_5	-0.378	[-4.51]	-0.350	[-4.19]
REG_SOUTH	θ_6	-0.058	[-0.55]	-0.110	[-0.99]
Inefficiency Model					
CONSTANT	δ_0	-21.415	[-1.87]	-5.128	[-1.46]
NAFTA	δ_1	1.987	[2.92]	0.628	[1.28]
EDUCATION	δ_2	-2.881	[-1.11]	0.164	[0.36]
DISTANCE_US	δ_3	2.574	[2.19]	0.467	[1.57]
CRIME	δ_4	1.529	[2.52]	0.376	[1.39]
sigma_u2	σ_u^2	0.121		0.221	
sigma_v2	σ_v^2	0.146		0.146	
Lambda	λ	0.828		1.518	
Observations	NxT	160		160	
Log-likelihood		59.5668		59.6575	
H0: Constant Returns to Scale*	$\beta_1+\beta_2=1$	30.93	0.000	25.70	0.000

*p-value reported instead of t-ratio

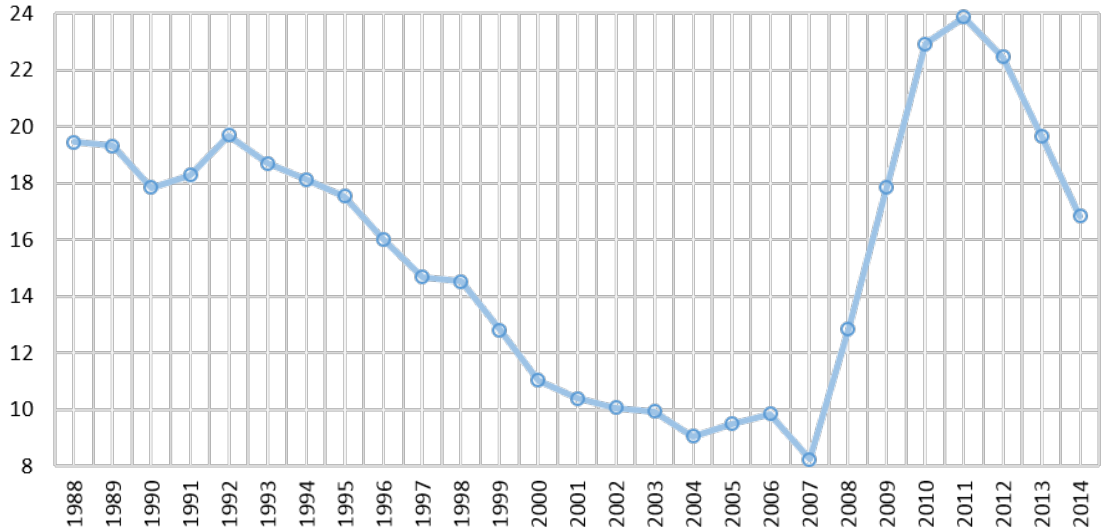
Table 4 - Stochastic Production Frontier Estimation using Crime in Deciles

		CFG95 1st Decile		CFG95 1st to 6th Decile	
Variable	Par.	Estimate	t-ratio	Estimate	t-ratio
INTERCEPT	β_0	8.634	[20.57]	8.591	[20.87]
CAPITAL	β_1	0.346	[8.36]	0.339	[8.19]
LABOR	β_2	0.813	[15.33]	0.822	[15.45]
CAPITAL× LABOR	β_3	0.399	[2.72]	0.383	[2.61]
$CAPITAL^2$	β_4	0.389	[-2.9]	0.371	[-2.77]
$LABOR^2$	β_5	0.452	[-2.65]	0.443	[-2.6]
EDUCATION	β_6	1.162	[5.44]	1.186	[5.62]
SPECIALIZATION	β_5	0.039	[2.83]	0.039	[2.84]
TREND	γ_1	0.188	[3.57]	0.184	[3.5]
$TREND^2$	γ_2	0.085	[-4.9]	0.084	[-4.9]
REG_ BORDER	θ_1	0.410	[-4.3]	0.413	[-4.49]
REG_CAPITAL	θ_2	0.370	[-2.72]	0.365	[-2.73]
REG_CENTER	θ_3	0.311	[-4.31]	0.307	[-4.33]
REG_NORTH	θ_4	0.391	[-5.39]	0.386	[-5.47]
REG_PENINSULA	θ_5	0.392	[-4.4]	0.363	[-4.29]
REG_SOUTH	θ_6	0.086	[-0.86]	0.058	[-0.57]
Inefficiency Model					
CONSTANT	δ_0	28.129	[-3]	30.141	[-2.83]
NAFTA	δ_1	1.838	[2.39]	1.719	[2.44]
DISTANCE_US	δ_2	2.690	[2.33]	3.084	[2.27]
CRIME_1 decile	δ_3	-0.297	[-0.07]		
CRIME_2 to10 decile	δ_4	1.592	[2.13]		
CRIME_1 to 6 decile	δ_3			1.095	[1.23]
CRIME_7 to 10decile	δ_4			1.430	[2.05]
sigma_u2	σ_u^2	0.110		0.112	
sigma_v2	σ_v^2	0.149		0.149	
lambda	λ	0.734		0.749	
Observations	NxT	160		160	
Log-likelihood		59.3334		59.4403	

Table 5 - Initial and Final Technical Efficiency Index using CFG95

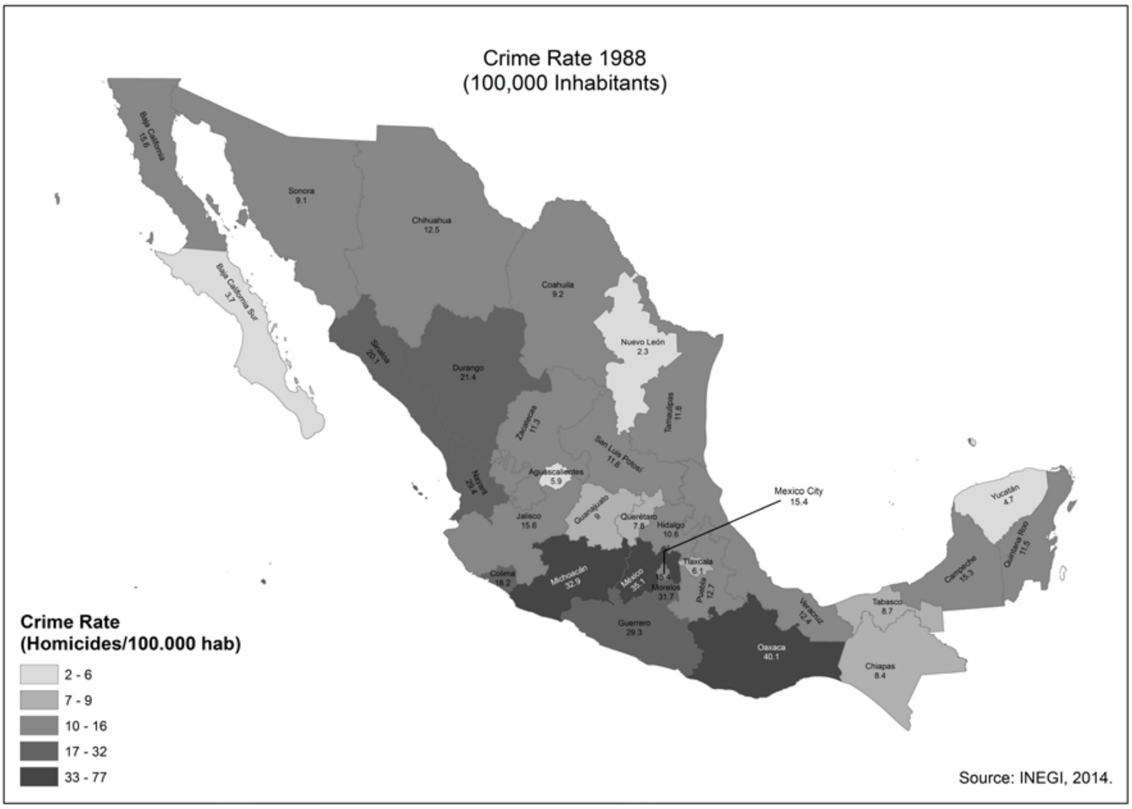
State	Initial TE	Final TE	TE change (x100%)	Rank TE change
Michoacán	0.75	0.78	0.03	1
Veracruz	0.94	0.96	0.03	2
Tabasco	0.91	0.93	0.01	3
Distrito Federal	0.95	0.95	0.00	4
Baja California	1.00	1.00	0.00	5
Hidalgo	0.96	0.96	0.00	6
Querétaro	0.97	0.97	0.00	7
Nuevo León	1.00	1.00	0.00	8
México	0.92	0.92	0.00	9
Coahuila	0.99	0.99	0.00	10
Tamaulipas	0.99	0.98	-0.01	11
Aguascalientes	0.98	0.97	-0.01	12
Zacatecas	0.97	0.96	-0.01	13
San Luis Potosí	0.97	0.96	-0.01	14
Guanajuato	0.97	0.95	-0.01	15
Sonora	0.99	0.98	-0.01	16
Jalisco	0.93	0.92	-0.02	17
Tlaxcala	0.97	0.94	-0.03	18
Durango	0.94	0.91	-0.03	19
Puebla	0.92	0.90	-0.03	20
Chiapas	0.92	0.89	-0.03	21
Campeche	0.90	0.87	-0.03	22
Yucatán	0.94	0.90	-0.05	23
Chihuahua	0.99	0.94	-0.05	24
Colima	0.91	0.85	-0.06	25
Baja California Sur	0.97	0.91	-0.06	26
Morelos	0.97	0.87	-0.11	27
Sinaloa	0.94	0.77	-0.18	28
Nayarit	0.94	0.75	-0.21	29
Quintana Roo	0.90	0.71	-0.22	30
Oaxaca	0.90	0.62	-0.31	31
Guerrero	0.78	0.50	-0.36	32
National Mean	**0.94**	**0.89**	**-0.05**	

Figure 1 – Evolution of Crime Rate in Mexico (homicides per 100,000 inhabitants)



Source: INEGI (2015)

Figure 2 – Crime rate Maps 1988 & 2008



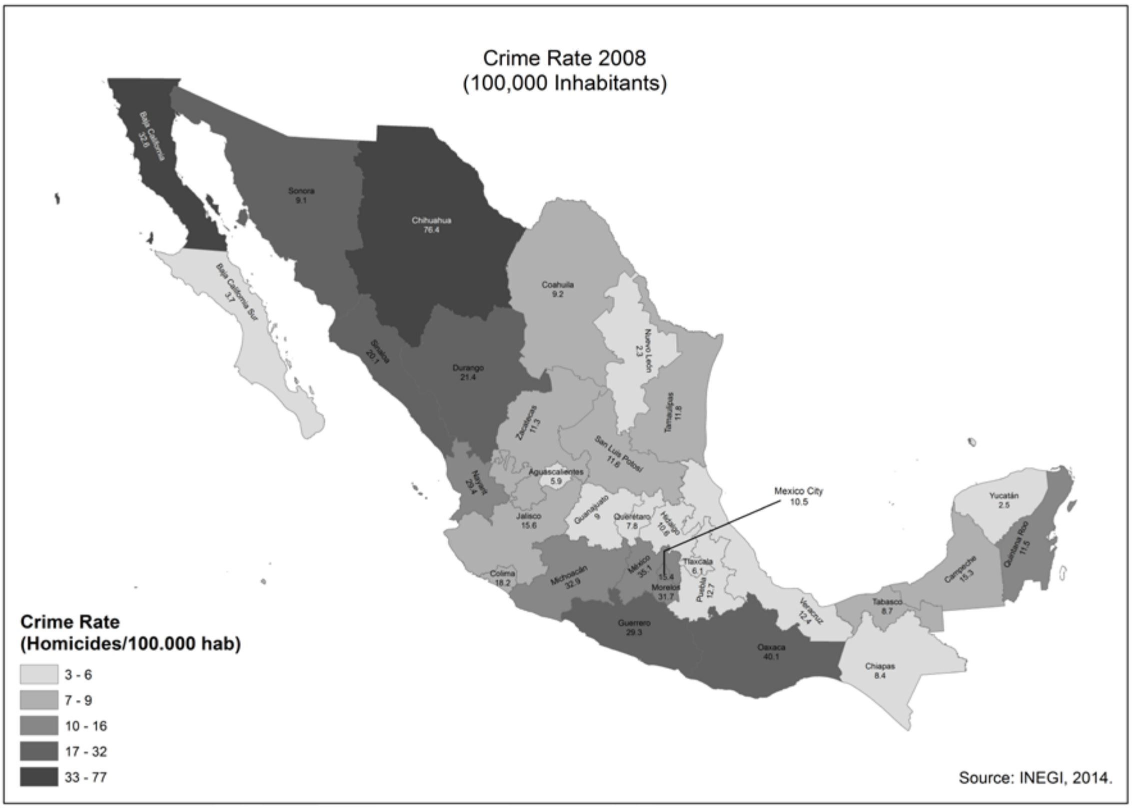


Figure 3 - Map Efficiency using CFG95 coefficients