THE INFLUENCE OF FEMICIDE ON CRIMINAL BEHAVIOR: AN EMPIRICAL APPROACH USING ECONOMIC COMPLEXITY FOR CRIME PREVENTION IN MEXICO

LA INFLUENCIA DEL FEMINICIDIO EN EL COMPORTAMIENTO CRIMINAL: UN ENFOQUE EMPÍRICO BASADO EN LA COMPLEJIDAD ECONÓMICA PARA LA PREVENCIÓN DEL DELITO EN MÉXICO

Max Lugo Delgadillo

Entropia AI

Resumen: Bajo la metodología de complejidad económica propuesta por Hausmann et al. (2013), esta investigación estableció un ordenamiento de la evolución del crimen en México. Este ordenamiento se basa en la complejidad de los delitos, determinada por las capacidades requeridas para cometerlos. De acuerdo con esta investigación, los delitos con menor complejidad en México incluyen el robo en transporte público colectivo, robo en transporte individual, falsificación y robo de autopartes. Por otro lado, los delitos con mayor complejidad están relacionados con el crimen organizado, el homicidio intencional e involuntario y el tráfico de menores. Además, el análisis reveló que el feminicidio y la violación exhiben la centralidad o influencia más significativa dentro de la red criminal. Según el modelo, los delitos asociados con la violencia contra las mujeres son los que más aumentan la probabilidad de cometer otro delito. Por lo tanto, los esfuerzos de prevención dirigidos a reducir el feminicidio y la violación pueden tener un impacto sustancial en los niveles generales de criminalidad en México. Este estudio también destaca la importancia de abordar la violencia contra las mujeres en el diseño de políticas de prevención del delito en México. Además, la metodología adoptada en este estudio puede ser reinterpretada como un algoritmo de agrupamiento espectral y, por lo tanto, también contribuye a la literatura sobre aplicaciones de aprendizaje automático en el diseño de políticas públicas.

Abstract: Following the economic complexity methodology introduced by Hausmann et al. (2013), this study establishes an order of crime evolution in Mexico. This ordering is based on the complexity of crimes, as determined by the capabilities required to commit them. According to the study, the least complex crimes in Mexico include robbery in collective public transport, robbery in individual transport, counterfeiting, and auto parts theft. On the other hand, the crimes characterized by the highest complexity involve organized crime, intentional and unintentional homicide, and the trafficking of minors. Furthermore, the analysis reveals that femicide and rape exhibit the most significant centrality or influence within the criminal network. According to the model, crimes associated with violence against women are those that most increase the probability of committing another crime. Therefore, targeted prevention efforts aimed at reducing femicide and rape may substantially impact overall levels of criminality in Mexico. This study also highlights the importance of addressing violence against women in designing crime prevention policies in Mexico. Moreover, the methodology adopted in this study can be reinterpreted as a spectral clustering algorithm and thus also contributes to the literature on machine learning applications in public policy design.

1. Introduction

According to data from the Executive Secretariat of the National Public Security System in Mexico (*Secretariado Ejecutivo del Sistema Nacional de Seguridad Pública*, SESNSP), from 2015 to 2019, each year was marked as the most violent in Mexico's recent history. Throughout this period, 119 887 intentional homicides and 3 594 femicides were committed. These numbers equate to an average of 66 intentional homicides and two femicides per day during these five years. The data on homicide rates in Mexico, as reported by the World Bank, highlights that the country has alarmingly high levels of homicides compared to other countries. Specifically, in 2018, Mexico had a rate of 29 homicides per 100 000 inhabitants, which is higher than countries such as the United States, Japan, Spain, the United Kingdom, Sweden, Switzerland, Ireland, and Turkey, which do not exceed five homicides per 100 000 inhabitants. Furthermore, in comparison to other Latin American countries, Mexico has similar levels to Brazil (27) and Colombia (25) and is significantly higher than the levels of Argentina (5), Chile (4), Ecuador (6), and Peru (8). This information emphasizes the need for effective measures to address the high levels of violence in Mexico, especially when compared to other countries (World Bank, 2021).[1]

Furthermore, femicide stands out as a distinct component of crime in Mexico compared to other countries. Data from the World Bank reveals that the homicide rate for women in Mexico is 5.8 per 100 000 inhabitants, significantly higher than rates in developed countries such as the United States (2.2), Germany (0.9), France (0.7), Spain (0.5), Sweden (0.7), and Switzerland (0.7), among others. This data highlights the prevalence of violence against women in Mexico and emphasizes the urgent need for targeted interventions to address this specific form of violence.[2] Compared to other Latin American countries, Mexico exhibits comparable levels of female homicide to countries like Colombia (4.2) and Brazil (4.3). Additionally, studies indicate a close association between femicide and gender violence, sexual harassment, sexual abuse, rape, organized crime, and human

---

[1] The data presented for 2018 is based on the available information for the country. If such data is unavailable, the most current value at the time of analysis is presented.

[2] The data on homicides of women in 2018 were obtained from the World Bank and can be accessed through the following link: https://datos.bancomundial.org/indicador/VC.IHR.PSRC.FE.P5.

and drug trafficking. Prieto-Carrón *et al.* (2007) provide evidence that neighborhoods with high occurrences of organized crime, drug trafficking, and homicides are linked to a greater likelihood of femicide.

This research seeks to address crime in Mexico by providing an evidence-based analysis using a complexity model to characterize criminal behavior. This model identifies critical areas that significantly impact the criminal network by examining the evolution of crimes and their interrelationships. The study employs this model to offer policy recommendations for crime prevention. It highlights the importance of addressing violence against women in criminal activity, as it notably impacts overall crime rates. With a focus on Mexico, this study aims to contribute to the existing literature on crime prevention by providing a comprehensive understanding of the underlying factors contributing to criminal activity.

Moreover, the literature has established that criminal activity adversely affects economic activity in Mexico. Research indicates that elevated crime levels can harm investment levels, economic growth, and the welfare of individuals and communities.

Previous literature has shown that criminal activity can adversely affect Foreign Direct Investment (FDI). A study conducted by Torres *et al.* (2018) found that homicides and thefts have a statistically significant negative impact on FDI in Mexico. The same pattern has been observed in other nations, as demonstrated by Brown and Hibbert (2017), who examined information from 62 countries from 1997 to 2012 and found that FDI is similarly affected by violent crimes. Additionally, studies have indicated that criminal activity can adversely affect labor and productivity. Cabral *et al.* (2016) found that criminal activities have a negative effect on Mexican labor productivity from 2006 to 2016, which corresponded to the period known as the "drugs war period". Additionally, during the same timeframe, Benyishay and Pearlman (2013) identified a 1% to 2% decrease in hours worked. Benyishay and Pearlman (2014) also discovered that higher theft rates decrease the probability that micro-enterprises in Mexico will expand their operations, resulting in a decline in economic activity. Conversely, Pan *et al.* (2012) found that the crime rate of neighboring states had a negative correlation with economic growth. Finally, González (2016) established a negative correlation between economic growth and crime rates in Mexico.

From an alternative perspective, criminal activity has significant direct economic costs. One way to evaluate these costs is through the accountability method, which divides all crime-related expenditures

into three categories: social costs, private sector costs, and government costs. Social costs cover a range of expenses, such as lost income and victim expenses. Private sector costs include expenses incurred by households and businesses for crime prevention. Government costs comprise costs related to the justice system, police, and correctional facilities.

Using the aforementioned methodology, recent studies have measured the direct cost of crime as a percentage of Gross Domestic Product (GDP) for several countries. Chalfin (2015) estimates that the cost of crime in the United States is approximately 2.5% of the GDP. Olavarria-Gambi (2007) estimates that Chile's total crime cost is approximately 6% of the GDP. The Inter-American Development Bank (IADB) estimates that for 17 Latin American countries, the average cost of crime is approximately 3.5% of the GDP, with an estimated 1.92% for Mexico (Jaitman *et al.*, 2017).

From another perspective, the economic aspects of criminal behavior have been the subject of extensive study. Becker (1968) established the economic incentives for individuals to engage in criminal activities. Becker characterizes criminal behavior as a choice between productive and criminal activities concerning the available labor opportunities and highlights how this choice is closely related to labor market policies, the effectiveness and quality of the judicial system, and other economic factors. Recent literature has expanded on the determinants of criminal behavior, including wage and income inequality, poverty, cultural and family background, social exclusion, education, and other economic and social factors. A comprehensive review of 20 years of literature on crime from an economic perspective by Buonanno (2003) concluded that criminal activity is a complex phenomenon affected by various socioeconomic determinants. He highlights that understanding these determinants and their relationships is crucial for effective policy implementation.

The policy intervention methodology proposed in this study is based on the economic complexity approach. As outlined by Hausmann *et al.* (2013), the approach involves using a network procedure to examine the relationships between production outputs by region and to understand the evolution of production capabilities. The analysis examines each economic region and assumes that regions producing similar products have comparable production conditions. For instance, the analysis might indicate that the production of shirts is more closely connected to the production of pants than to the production of turbines. The intricacy of each product is established as a measure of the essential capabilities needed for its production. A

product ordering by complexity is then created to implement policies. Similarly, this study uses a similar approach to construct a network of crimes and comprehend their progression. An ordering of criminal activities by their complexity is subsequently introduced to guide policy interventions. To summarize, this research aims to describe the crime network, the evolution of criminal activities (order), and identify crime characteristics for policy intervention using the economic complexity approach.

This study employed crime incidence data from state and federal jurisdictions at the state level from 2015 to 2019. The research outcomes reveal that the evolution of crime in Mexico commences with public transport robberies and ultimately leads to organized crime.

It represents a notable addition to the extant literature on economic complexity, crime behavior analysis, network analysis, public policy analysis, and machine learning applications. Additionally, it presents a fresh approach to analyzing criminal behavior with associated policy recommendations. Specifically, the findings indicate that preventative measures targeting femicide and rape would be the most productive means of dismantling the crime network. By targeting these crimes and the offenses linked in their network, it appears that overall crime rates can be reduced while also preventing the diversification of criminal activity.

The model used in this study can be considered an unsupervised model. It employs the data to estimate relationships between variables (crime incidence). Such a model is typically called a clustering model and is frequently classified as a type of machine learning model. The application of network analysis has become widespread across a variety of academic disciplines, including public policy, computer science, economics, and epidemiology. In the field of epidemiology, scholars such as Li *et al.* (2014), Aihara and Kim (2009), and Verdasca *et al.* (2005) have employed this approach to examine the spread and control of diseases. Within the computational field, network analysis has been used to identify potential nodes, as well as weak and strong links within networks, and to develop algorithms that efficiently dismantle them, Ren *et al.* (2018), Braunstein *et al.* (2016), and Han *et al.* (2021) are among the authors who have contributed to this area of research. Additionally, Kertész and Wachs (2020) argue that complexity analysis plays a crucial role in understanding criminal behavior, particularly in fraud and corruption. The authors suggest that the complexity approach has become an important tool for describing, analyzing, and detecting crime over the last three decades.

Various methodologies have been employed in crime behavior and network analysis, including social network analysis and spatial analysis. Several scholars have contributed to this literature, including Papachristos *et al.* (2013), Sierra-Arévalo and Papachristos (2015), Scott and Carrington (2011), and Neto (2017), among others. In Mexico, there has been a growing interest in investigating the correlation between complexity and crime. For example, Ríos (2017) quantifies the impact of crime on economic diversity through the lens of complexity.

2. Economic complexity methodology

The proposed model draws on the seminal works of Hidalgo and Hausmann (2009) and Hausmann *et al.* (2013), exploring the foundational concepts of economic complexity. Hausmann *et al.* (2013) develop a methodology to examine the progression of products across different economies, classifying them by complexity level based on the capabilities, resources, and environmental factors required for production. Complexity is defined as the magnitude of the material resources, productive capacities, and contextual elements, such as institutional framework and educational attainment, necessary for creating a product. These authors further demonstrate the construction of a product complexity index, providing a means of characterizing products in terms of complexity.

Building on the argument put forth by Hausmann *et al.* (2013), it can be argued that, similar to products, the commission of crimes demands specific capabilities. These include access to resources, educational level, institutional environment, and level of organization. This perspective suggests that crimes that require limited education, an ineffective institutional framework, and minimal organizational structures and resources are likely to be considered less complex, while those requiring greater levels of these capabilities may be viewed as more complex.

The algorithm is premised on the idea that certain geographical locations possess a predisposition for committing specific crimes, reflecting the area's inherent capabilities. This predisposition can be formally defined as a competitive advantage for committing a crime in a particular location. By analyzing the crimes that have a competitive advantage in specific locations, it is feasible to comprehend their interplay with other crimes and develop a map of relationships, referred to as the proximity matrix. This matrix represents the conditional probability that two distinct crimes are committed in the same

location and serves as the basis for constructing the product space, a two-dimensional visualization of the interactions between all crimes. The analysis of the complexity and the graph will serve as tools for comprehending the evolution of crime and formulating a strategy for its disruption.

In accordance with Mealy *et al.* (2017), in this study, complexity is established in a matrix format where the M^o matrix represents crimes as columns and the 32 states in Mexico as rows. Each entry of the matrix depicts the average of every crime. To determine the competitive advantages, the matrix is transformed into a binary matrix with entries of 0 and 1, with 0 indicating that a state does not have a competitive advantage in committing a particular crime and 1 signifying the presence of such an advantage. This analysis aims to identify, for each state, those crimes more easily committed than the national average and to comprehend their relationships.

In further detail, let us define $M^o_{e,c}$ as the element of the matrix M^o for the state e and for the crime c. $S_{e,c}\left(M^o_{e,c}; M^o\right)$ is defined as the ratio of $M^o_{e,c}$ over the total crimes in state e (sum of row e in M^o); and finally, $S_c\left(M^o\right)$ is defined as the crime ratio of crime c (sum of column c) over the total sum of crimes (total sum of the matrix M^o). Based on international trade theory, the revealed comparative advantage (RCA) function of the state e with respect to crime c is defined as:

$$RCA\left(M^o_{e,c}; M^o\right) = \begin{cases} 1, \ if \ \frac{S_{e,c}}{S_c} \geq 1 \\ 0, \ in \ other \ case \end{cases}$$

The function defines a specific crime as a competitive advantage for a given state when the proportion of that crime in the state is greater than the proportion of the same crime across the entire nation.[3]

---

[3] It is worth noting that the approach could be modified to use crime rates instead without affecting the underlying algorithm. To ensure the comparability of rates, the RCA function could be adjusted as follows: $RCA\left(M^o_{e,c}; M^o\right) = \begin{cases} 1, \ if \ \frac{T_{e,c}}{T_c} \geq 1 \\ 0, \ in \ other \ case \end{cases}$. Where $T_{e,c}$ represents the rate per 100 000 inhabitants of crime c in the state e over the total crime rate in the state. T_c represents the rate per 100 000 inhabitants of the nation of crime c over the total crime rate. Therefore this calculation is equivalent to the original function. Another option would be to simply compare the rates with respect to the national average, which would have implications for the RCA function. In this case, we follow the definition

An alternative approach for conducting the analysis is to use the singular value decomposition (SVD) of the RCA matrix. This method provides metrics related to the proximity matrix and complexity, as exemplified by Che (2020) in the context of policy recommendation. Nonetheless, recent research by Hidalgo (2021) has demonstrated that the economic complexity index already encompasses the information derived from the SVD descomposition. It is worth noting that the SVD can be interpreted as a Cobb-Douglas production function, indicating that both metrics essentially measure the production of goods or crimes.

The matrix M is defined with the same size of M^o with the results of applying the RCA function to the corresponding element in M^o. Therefore, M is the transformation of the matrix M^o to 1 and 0 elements.

Following the above procedure, it is appropriate to introduce the concepts of diversity and ubiquity. Diversity is the collection of crimes that exhibit a competitive advantage in a given state, while ubiquity is the collection of states that engage in a particular crime with a competitive advantage. In a more formal sense, we define these concepts as follows:[4]

$$Diversity = M1_{(C)}$$

$$Ubiquity = M^T 1_{(E)}$$

The concept of complexity can be expressed in terms of two matrix equations obtained through the use of the operator $diag(.)$ that diagonalizes a vector. Let $U = diag(Ubiquity)$ and $D = diag(Diversity)$. Now, we can define a system of equations that characterize complexity. The first equation defines the complexity of a state as the average complexity of the crimes in which it has a competitive advantage, while the second equation defines the complexity of a crime as the average of the complexity of the states where it is committed. The system can be expressed as follows:[5]

of Hausmann *et al.* (2013).

[4] $1_{(C)}$ denotes a vector of ones of length C (i.e., the number of columns or distinct crimes), and $1_{(E)}$ denotes a vector of ones of length E (i.e., the number of rows or states).

[5] Recent developments in the field of complexity have led to the introduction of novel theoretical frameworks that build upon the more general formulations of the equations presented here. For instance, notable contributions have been

$$X = D^{-1}MY \tag{1}$$

$$Y = U^{-1}M^T X \tag{2}$$

In this formulation, the complexity of states and crimes is represented by vectors X and Y, respectively. Substituting the expression for X from the first equation into the second equation yields:

$$Y = U^{-1}M^T D^{-1}MY \tag{3}$$

The eigenvectors of the matrix $U^{-1}M^T D^{-1}M$ provide solutions to the equation presented above. The second eigenvector of this matrix characterizes the complexity, denoted by the vector Y. Note that the first eigenvector is constant by construction. Since an eigenvector may have multiple associated solutions, the vector is ordered in a manner that guarantees its correlation with centrality is positive. Generally, the greater the complexity, the stronger its relation to other crimes. It is important to bear in mind that Y represents a metric of complexity, which captures the required capabilities for committing a given crime.

In this final step, the proximity matrix is constructed from the network under analysis. The proximity matrix captures the conditional probability of a particular crime occurring given the occurrence of another crime. Formally, the proximity matrix Ω can be defined as follows:

$$\Omega = \min \left\{ U^{-1}M^T M, M^T M U^{-1} \right\} \tag{4}$$

The network to be analyzed, known as the proximity matrix, is constructed as the minimum element-wise operator (min) among the matrices. The proximity matrix is a symmetric matrix that represents the proximity between crimes through nodes and links. Its symmetry implies that for two offenses A and B, the conditional probability is $P(A|B = P(B|A)$, and therefore $P(A) = P(B)$.[6] The variance in the matrix is given by the joint probability of two crimes, and

---

made by Albeaik *et al.* (2017) and Salinas (2021), respectively. These scholars have advanced the field by proposing modifications and extensions to the existing models, which have expanded the range of applications and improved the accuracy of the results.

[6] For $P(A) \neq 0$, $P(B) \neq 0$, and $P(A\&B) \neq 0$.

since the matrix is not constant in all its elements (except the diagonal), it implies that the crimes are not independent of each other $(P(A\&B) \neq P(A) * P(B))$. A joint density function is assumed for any pair of offenses A and B defined as $f(A,B)$, which is increasing in both arguments. Hence, it is assumed that reducing the crime with the highest joint probability concerning the others would have a greater impact on the marginal probabilities. This argument is later resolved with the term centrality.

In other words, the proximity matrix measures the likelihood of producing different crimes and visualizes the potential pathways for diversifying criminal activities. Although it does not directly incorporate the concept of complexity, it will be highlighted later that this algorithm can also capture this aspect.

2.1 *Decommissioning of the network*

This section aims to explore the relationship between the centrality concept and proximity matrix in the context of policy intervention, which is one of the main objectives of this research.[7]

Reflecting on the network's construction, the proximity matrix is established as the conditional probability of a crime being committed, given that another specific crime has occurred. As described in equation (4), it is the minimum element-wise of two matrices. Let us consider two crimes, denoted as A and B, represented in two vectors within the matrix M. By applying the concept of ubiquity and equation (4), the conditional probability of crime B occurring given the occurrence of crime A can be calculated as:

$$P(B|A) = \frac{\textit{Locations where both crimes A and B were commited}}{\textit{Locations where crime A was committed}}$$

The model construction implies that $P(B|A) = P(A|B)$ in the proximity matrix due to the computation of the minimum of both probabilities, as described in equation (4). This characteristic also guarantees the symmetry of the matrix. The impact of proximity matrix definition on complexity metrics has been briefly discussed

---

[7] Centrality is a well-established concept used to measure the importance of elements in a network. Specifically, it allows us to identify nodes that play a critical role in the network structure. Nodes with high centrality are considered potential targets for dismantling a network as their removal can significantly impact the network's overall functionality.

by Hausmann and Klinger (2007). To verify the assumption that $P(B|A) = P(A|B)$, an analysis was conducted on the matrices $U^{-1}M^T M$ and $M^T M U^{-1}$, serving as inputs for the proximity matrix. To compare the two matrices, the square difference was computed element-wise, and the mean was calculated. The resulting mean difference was found to be less than $1e^{17}$, indicating that the matrices are almost identical. The cosine similarity between the flattened versions of the two matrices was then computed, yielding a similarity score of 0.91. The results suggest that the assumption of symmetric conditional probabilities holds true for the constructed proximity matrix.[8] The two metrics employed in this study demonstrate that the assumption of $P(B|A) = P(A|B)$ doesn't impose huge restrictions on the conditional probabilities.

The assumption that $P(B|A) = P(A|B)$ leads to the conclusion that $P(A) = P(B)$, indicating that all crimes have the same probability. As a result, the proximity matrix captures the joint probability among crimes, imposing a normalization in probability measures across crimes. It is important to note that the matrix is not constant, as observed in the data, which implies that $P(A \cap B) \neq P(A) P(B)$.

The interdependence among crimes implies that preventing one crime could affect the probability of committing another crime. We define a set S containing all possible events, and the events represented in the proximity matrix are a subset of S. It is worth noting that we define a probability space with a σ-algebra, a subsample of events S (discrete), and a probability function. The subset represents the events in which a crime has already occurred. Therefore, the marginal probability of event A can be computed as:

$$P(A) = \sum_{s \in S}^{f} (A, s)$$

Assuming that the function f is strictly increasing in both arguments, it follows that an increase in the value of an element $s \in S$ will result in an increase in the probability of event A.[9] The objective is to identify the element $s \in S$ that will minimize the probability of event A, representing the crime of interest. In other words, it aims

---

[8] In this case, a cosine similarity of 1 indicates complete similarity between the compared matrices, whereas a similarity of 0 indicates complete dissimilarity.

[9] It was assumed that the marginal probabilities have the same probability function for all crimes

to identify the element with the highest contribution to the probability of committing another crime. Preventing this particular event would lead to a greater reduction in the likelihood of the other crimes occurring.

We obtain the joint probabilities for all tuples of crimes (function f for some A) from the proximity matrix. Under this assumption, given that one crime has been committed, we compute the sum for every row in the matrix (excluding the diagonal) and divide it by the total number of crimes (rows). This operation can be expressed mathematically as a matrix operation:

$$V = \frac{\Omega' 1_{(P)}}{C}$$

Where $1_{(P)}$ is a vector of ones with the corresponding size, C is a constant value (number of crimes) and Ω' is equal to Ω but with the diagonal entries equal to 0. The entries of V provide information about the average contribution of each crime to the probability of another crime being committed, given that a crime has already occurred. A higher value of the j-th entry in V compared to the i-th entry indicates that the crime represented by the j-th row of the proximity matrix, on average, contributes more to an increase in the probability of committing another crime than the crime represented by the i-th row. As a result, preventing crime j will have a greater impact on reducing the probability of crime production compared to preventing crime i. Hence, we identified the crime with the highest average contribution as a potential target for crime prevention. The proposed metric is akin to the closeness centrality measure in network analysis, where probabilities correspond to edges. In this context, closeness centrality can be computed as the sum of joint probabilities. Drawing on this analogy, our probability-based approach is connected to network analysis, enabling the use of existing tools for centrality computations.

The closeness centrality does not fully account for the fact that a node, such as a crime, can be connected to other nodes with high centrality. We can use eigenvector centrality to capture this aspect of the network structure better. This measure assigns a centrality score to each node based on its proximity to other nodes while considering the centrality of those neighboring nodes. As a result, the eigenvector centrality provides a more comprehensive view of the network structure. Therefore, it is a more suitable metric for identifying crimes likely to have the greatest impact on reducing criminal activity.

3. Data

The analysis in this study employed data on crime incidence between 2015 and 2019 from both state and federal jurisdictions in Mexico. The data was gathered on a state-wise basis from the SESNSP. The year 2020 was excluded from the analysis to exclude the effect of the COVID-19 pandemic on crime rates.

It is pertinent to note several significant features of the data. The analysis was performed using aggregated data at the state level, which enabled the collection of crime data for state and federal jurisdictions and ensured comparability across all states in Mexico. Currently, it is not feasible to construct detailed crime data at the municipal level for the entire country (federal crime data is only available at the state level). Consequently, this leads to a sparsity issue in the model. The model thus has a trade-off between incorporating fine-grained local characteristics in geographical terms and ensuring the completeness of the data.

It should be emphasized that extreme cases arise when the data is characterized at a very fine-grained level of analysis (e.g., neighborhoods), as the resulting matrix M may have only one element equal to 1 for every row.[10] In such cases, equation (3) does not have a solution. Conversely, there is no solution if only one aggregation unit is used. Therefore, it is essential to consider the aggregation of units (states/municipalities, etc.) or crimes (higher or lower aggregations) that ensure variance and capture some of the local features implicit in the relations between crimes. As a result, the analysis does not incorporate local crime conditions, such as street conditions or neighborhood income levels. The model focuses on state-level characteristics and makes inferences based on the relationships between crimes.

An alternative approach to address this issue is to construct units that ensure similar characteristics and have a suitable level of aggregation for the analysis. In the present study, the state-level aggregation was chosen due to the completeness and comparability of the crime data and to ensure some homogeneity in key factors such as income, poverty, state of law, and institutional context, among others. However, it is acknowledged that this approach may not capture more localized effects, such as the conditions of specific neighborhoods. Thus, future research could explore the development of new aggregation units that capture different characteristics of interest.

---

[10] In a more common analysis, this can be interpreted as a fixed effect added to every unit of analysis in the data.

The complexity of the crime network was assessed using the average annual incidence of crime at the state level, considering both the type and subtype of crime defined by the classification methodology of the SESNSP. For the federal jurisdiction, only the concept of crime was considered based on the SESNSP classification catalog. A total of 9 907 975 crimes were classified into 59 different categories. To ensure the robustness of the analysis, certain categories of crimes were excluded from the analysis, namely: family violence, rapture for marital purposes, incest, failure to comply with family assistance obligations, and other crimes against the family, as they are known to be significantly underreported, which may skew the indicators. Moreover, the categories "Other crimes of the General Health Law (LGS)" and "Other crimes in Federal Laws and Codes" were also removed as they encompassed a broad range of crimes within a subset of classification.

Appendix A demonstrates that the inclusion or exclusion of underreported crimes in the analysis does not affect the model's solution, assuming that all crimes have the same level of underreporting. However, we removed certain crimes from the analysis because we sought to maintain a similar level of underreporting across the sample. The crimes of family violence, rapture for marital purposes, incest, failure to comply with family assistance obligations, and other crimes against the family have an exceptionally high underreporting rate of 99%, as reported by México Evalúa. In addition, we present a model in in Appendix A that accounts for the impact of underreporting on the crime data. Despite this adjustment, the overall results remain consistent. The final sample comprises 87% of all reported crimes from 2015 to 2019, encompassing a total of 50 distinct crime categories.[11]

It is important to acknowledge that under-registration affects all crimes, which could impact the results' validity. It is assumed that the SESNSP registry accurately reflects the percentage composition of crimes for each state and that the under-registration rate is roughly similar across all states. Thus, the relative measurement of one crime concerning another should not be biased. To verify the robustness of the findings, an analysis was conducted in Appendix B that considers the under-registration rate for each crime. However, the results of this analysis remained consistent with the general conclusions drawn

---

[11] Approximately 10% of the crimes reported from 2015 to 2019 were categorized as family violence and non-compliance with family assistance obligations. A distribution of the total reported crimes during this period, categorized by class, is presented in Appendix A. Burglary was aggregated with other crimes against property to ensure comparable levels of underreporting across the sample.

in the absence of this consideration.

In constructing our dataset, we decided to exclude the modality of crimes as a classification variable. This decision is because the SESNSP catalog does not consistently characterize all crimes as violent or non-violent. This could lead to subjective criteria being used to determine how to split the crime data, which would complicate the construction of the dataset. Therefore, we chose to use the classification system already established by the Mexican government to maintain consistency and avoid any unnecessary subjective interpretations.

4. Results: crime complexity ordering

During the calculations above, the complexity of each crime is determined, culminating in the establishment of a ranking among all crimes. Additionally, the evolution of crimes is characterized based on their complexity. The model classifies crimes into two broad categories, which are referred to as high-complexity (positive complexity) and low-complexity (negative complexity).[12] The low-complexity category encompasses most robberies, except for those targeting banking institutions or machinery. Conversely, the high-complexity crimes are primarily federal jurisdiction offenses, such as organized crime, crimes against public health, and "huachicoleo", among others, and state jurisdiction offenses, such as homicide, femicide, and kidnapping.[13]

The unsupervised model identifies organized crime as the one with the highest complexity, while robbery in collective public transport is the lowest. Previous research by Mealy *et al.* (2017) suggest that the economic complexity index is comparable to spectral clustering, which is widely applied in image recognition, web page ranking, information retrieval, and RNA motif classification, among other fields. These machine learning algorithms are commonly used in Artificial Intelligence (AI) applications. Thus, the model can be viewed as a machine learning model that learns about criminal behavior and makes inferences regarding its evolution.

Upon analyzing the four crimes with the highest complexity - namely organized crime, accidental homicide, trafficking of minors, and intentional homicide- it was found that organized crime and

---

[12] The second eigenvector, which provides solutions to equation (3), contains both positive and negative values.

[13] "Huachicoleo" in Mexico refers to the robbery of gasoline.

trafficking of minors demand significant resources and organization, which can explain their high-complexity. Additionally, accidental homicide and intentional homicide were relatively classified together, despite intentional homicide being considered less complex than accidental homicide, which may seem counter intuitive. The categorization of crimes may offer a potential explanation for the relative complexity of intentional and accidental homicide.

The available evidence suggests the possibility of fraudulent registration of homicides as accidental deaths. Accidental homicides should exhibit a random pattern due to their unintentional nature. However, recent data from the SESNSP indicate an increasing trend in accidental homicides, as well as a positive correlation with intentional homicides. The data suggests that certain states may be manipulating data for various reasons. A comprehensive review of crime registration conducted by México Evalúa (2020) supports this claim and documents the evidence of such practices.

The above findings highlight a significant concern regarding the classification of accidental homicide, which appears to be influenced by intentional homicide. Moreover, intentional homicide is closely linked with various other criminal activities, including organized crime, kidnapping, and crimes against health, among others. Therefore, accidental homicide may be capturing this complexity, reflecting a significant component of intentional homicide.

Regarding the low-complexity crimes, it is found that robbery in collective public transport, theft of auto parts, robbery in individual transport, and forgery occupy the top positions. Their prevalence is likely the main determinant for their classification as low-complexity crimes. However, the data does not support this explanation, as these crimes only represent around 3% of all crimes. Moreover, the argument that prevalence alone determines high-complexity crimes does not hold either for the same reason.

The results obtained from the analysis provide insights into the progression of criminal activities in Mexico, indicating an incremental increase in criminal complexity that appears to correlate with the severity of the crime. Hausmann *et al.* (2013) mention that the complexity index represents an estimation of a region's productive capacity. In the present case, this crime transition indicates unobserved capabilities. Further research could delve into the underlying factors that influence the complexity index, including but not limited to demography, economic conditions, law enforcement, and institutional environment. Table 1 displays the crimes ranked from high-complexity to low-complexity; the complexity transitions from

positive to negative values, highlighting two complexity groups.

Table 1

Complexity and centrality by crime in Mexico

Crime Type / Subtype	Complexity	Centrality
1 Federal Law Against Organized Crime (L.F.C.D.O...	1.282244	0.116838
2 Homicide / Accidental Homicide	1.114279	0.169203
3 Trafficking of minors / Trafficking of minors	1.109835	0.066315
4 Homicide / Intentional Homicide	1.072569	0.1705
5 Robbery / Bank robbery	0.967762	0.170021
6 Kidnapping / Kidnapping	0.966362	0.162753
7 Other Crimes / Other Crimes	0.960398	0.172158
8 Rape / Rape	0.954729	0.178588
9 Femicide / Femicide	0.873179	0.176932
10 Against Health / Against Health	0.871286	0.165363
11 Drug dealing / Drug dealing	0.724743	0.151585
12 Hindering prosecution or apprehension / Hinder...	0.702957	0.164484
13 Human trafficking / Human trafficking	0.591491	0.154977
14 Corruption of minors / Corruption of minors	0.585163	0.165804
15 Abortion / Abortion	0.544504	0.14767
16 Robbery / Motor vehicle theft	0.542611	0.150154
17 Robbery / Robbery of machinery	0.517433	0.153746
18 Dispossession / Dispossession	0.495476	0.169062
19 Other crimes that violate sexual freedom and s...	0.474758	0.15733
20 Falsehood / Falsehood	0.421231	0.176206
21 Crimes committed by public servants / Crimes c...	0.406796	0.150411
22 Other crimes that threaten life and physical i...	0.340641	0.163397
23 Statutory rape / Statutory rape	0.340377	0.161657
24 Sexual coercion / Sexual coercion	0.274787	0.154193
25 Sexual abuse / Sexual abuse	0.220053	0.17231
26 Breach of trust / Breach of trust	0.219709	0.163593

Table 1
(Continued)

Crime Type / Subtype	Complexity	Centrality
27 Property damage / Property damage	0.218341	0.156179
28 Robbery / Street robbery	0.188597	0.141702
29 Robbery / Home Invasion Robbery	0.045849	0.152078
30 Extortion / Extortion	0.002726	0.143424
31 Injuries / Accidental Injuries	-0.004048	0.12629
32 Robbery / Robbery of delivery people	-0.054524	0.087013
33 Other crimes against property / Other crimes a...	-0.076549	0.139845
34 Sexual harassment / Sexual harassment	-0.126157	0.162789
35 Fraud / Fraud	-0.187616	0.162339
36 Threats / Threats	-0.188564	0.15276
37 Robbery / Other Robbery	-0.220298	0.139085
38 Other crimes of the State Jurisdiction / Other...	-0.227526	0.066913
39 Electoral / Electoral	-0.232462	0.13867
40 Injuries / malicious injuries	-0.309035	0.138732
41 Against the environment / Against the environment	-0.621757	0.100565
42 Gender violence in all its forms different fro...	-0.810025	0.040345
43 Robbery / Comercial robbery	-1.006022	0.112066
44 Robbery / Robbery in individual public transport	-1.146763	0.124403
45 Other crimes against society / Other crimes ag...	-1.157448	0.100623
46 Robbery / Robbery on public roads	-1.776704	0.081811
47 Robbery / Robbery of auto parts	-1.955656	0.071711
48 Robbery / Robbery in individual transport	-2.008686	0.053062
49 Counterfeiting / Forgery	-2.242112	0.063859
50 Robbery / Robbery in collective public transport	-3.678936	0.026626

Notes: The presentation of crime type and subtype is limited to those within the state jurisdiction. In the case of federal crimes, only the concept is presented. It is worth noting that the "Other Crimes" category includes federal crimes such as "huachicoleo" and others.

Source: Own elaboration.

The presented order depicts the evolution of crime, where theft is generally observed to be at a low level of complexity, except for theft of machinery and banks. Following are crimes that threaten the well-being and dignity of individuals, including gender violence, threats, injuries, harassment, and sexual abuse. The next group of crimes escalates to those that threaten freedom and life, such as human trafficking, kidnapping, crimes against health, femicide, and rape, leading to the highest levels of crime complexity, which culminate in the trafficking of minors and organized crime.

The arrangement of crimes in the proposed order reflects a progression in the capabilities required to commit them. While crimes such as theft are characterized by a relatively low level of specialization and resources, the robbery of a banking institution demands a higher degree of both. It is possible that the grouping of intentional and accidental homicide together is due to the similarity of the capabilities required to execute each crime, despite their different motives. Similarly, femicide is ranked closely in terms of complexity, indicating a similar level of capabilities required, even though the motivation behind the crime is distinct. These observations suggest that the proposed model offers a good fit for the data.

The findings depict a hierarchy of crime evolution that describes the escalation of crime in complexity. It is plausible that criminals acquire knowledge and resources to perpetrate more complex crimes, such as transitioning from robbing a collective to joining a criminal organization and subsequently engaging in sex-related offenses and murder. Alternatively, this hierarchy could also reflect the evolution of the institutional environment in a specific region, which may facilitate the production of crimes of a certain complexity. Such an environment could result from corruption at different levels of government. Consequently, states with favorable institutional conditions for certain types of crimes may develop different levels of criminal complexity.

Finally, this ordering of crime evolution represents a significant step toward developing more effective and efficient preventive policies. It provides a deeper understanding of the genesis of crime based on quantitative measures, facilitating the generation of targeted policies to prevent its escalation. This suggests that the most effective way to address crime in the future may be to intervene at the early stages of its complexity. Such preventive policies can take various forms, such as improved social services and education programs, community policing initiatives, targeted interventions to address specific risk factors associated with particular crimes, and broader measures to improve

at-risk populations' social and economic conditions.

4.1 *Testing the crime evolution theory*

A Vector Autoregression (VAR) model was employed to investigate the order of crime evolution. The purpose of this model was to provide evidence supporting the notion that the evolution of crime proceeds from lower to higher complexity. The model was constructed using the vector $\Delta Y_{i,t} = [\Delta Y_{1,i,t} \Delta Y_{2,i,t}]'$, where $\Delta Y_{1,i,t}$ represents the rate of high-complexity crimes per 100 000 inhabitants in state i, and $\Delta Y_{2,i,t}$ represents the rate of low-complexity crimes per 100 000 inhabitants in state i, both in difference to their corresponding quarterly value in the previous year $(\Delta Y_{j,i,t} = Y_{j,i,t} - Y_{j,i,t-4} \ for \ j \in 1, 2)$.

The clusters generated by the methodology were used to distinguish between high and low-complexity crimes, where positive values represent high-complexity crimes and negative values represent low-complexity crimes. Subsequently, the data was processed to create a panel with 384 observations covering quarterly data from 2017 to 2019, aggregated into two categories (high and low-complexity) of crime rates. The estimation was performed using $K = 2$; the model included only two lags because higher orders of lagged variables were found to be insignificant. The Dickey-Fuller test confirmed the stationarity of the series. The estimated model is presented below:

$$\Delta Y_{,i,t} = \propto + \sum_{k=1}^{K} \Pi_{t-k} \Delta Y i_{i,t-k} + u_{i,t} \tag{5}$$

The matrix Π capture the interaction between lagged and present values. Before differencing, the series were detrended using a linear and quadratic trend. A Vector Error Correction (VEC) model was also estimated to account for potential co-integration among the series.

The results of the estimation provide significant evidence regarding the evolution of crime. As expected, the analysis shows that, in general, increases in low-complexity crime rates have a positive impact on future high-complexity crime rates. Table 2 presents the results of the estimation.

Table 2
Estimation VAR model

	VAR		VEC	
	$\Delta Y_{i,1,t}$	$\Delta Y_{i,2,t}$	$\Delta Y_{i,1,t}$	$\Delta Y_{i,2,t}$
$\Delta Y_{i,1,t-1}$	*0.43	-0.09	*0.40	-0.00
$\Delta Y_{i,2,t-1}$	*0.33	*0.76	*0.43	*0.42
$\Delta Y_{i,1,t-2}$	-0.08	0.01	-0.07	0.02
$\Delta Y_{i,2,t-2}$	-0.10	*-0.18	-0.15	-0.00
Error vector			Yes	Yes
Obs.	384			

Notes: *Significance at 90% with a bootstrap procedure.
Source: Own elaboration.

Upon analyzing the VAR model, it was observed that the estimated effect of $\Delta Y_{i,t-1,2}$ on $\Delta Y_{i,t,1}$ was positive and statistically significant at the 90% confidence interval. This indicates that lower rates of low-complexity crimes are associated with an increase in rates of high-complexity crimes in the future. In contrast, the estimate for the effect of $\Delta Y_{i,1,t-1}$ on $\Delta Y_{i,2,t}$ was negative, suggesting the possibility of substitution of crimes from lower to higher complexity. We used the Granger Causality test on both equations to determine any causal relationship between the variables. The results indicate that there is a Granger-causal relation between low-complexity crimes and future high-complexity crimes. However, for the second equation, no significant Granger-causal relationship was detected between high-complexity crimes and future low-complexity crimes. The Granger Causality test was performed on both equations using restricted and unrestricted models. For the first equation, the null hypothesis of no Granger Causality was rejected with an F-statistic of 4.76 (p-value of 0.00), indicating that $\Delta Y_{i,2,t-1}$ Granger-causes $\Delta Y_{i,1,t}$. On the other hand, for the second equation, the null hypothesis was not rejected with an F-statistic of 1.22 (p-value of 0.29), suggesting that $\Delta Y_{i,1,t-1}$ does not Granger-cause $\Delta Y_{i,2,t}$.[14]

---

[14] Typically, time series data in levels exhibit co-integration, which was addressed using stationary differenced series according to the Dickey-Fuller test. However, co-integration can still be present in stationary series. To test this

These results provide evidence to support the proposed order of crime complexity. The estimations suggest that crime complexity progresses from low to high complexity. For instance, it is possible that by preventing low-complexity crimes today, such as robberies, we can prevent high-complexity crimes in the future, such as homicides. This is because the capabilities necessary to commit low-complexity crimes could be a precursor to committing high-complexity crimes. Thus, disrupting initial capabilities to commit low-complexity crimes may also prevent more severe crimes from occurring in the future.

4.2 *Crime network*

In some cases, it may not be feasible to solely address the early stages of crime evolution. Hence, understanding the relationships between different types of crimes can aid in determining the optimal crime to prevent. Such a crime would be related to many other types of crimes and may generate negative externalities. The proximity matrix can be used to identify the relationships between crimes. While every crime is related to all other crimes, not all relationships are equally strong. Figure 1 provides a visual representation of the relationships between 50 different crimes, where each node corresponds to a crime and each line represents a relationship. Only the strongest relationships are depicted, and the nodes are colored according to the harm done defined in the SESNSP catalog. Furthermore, black nodes indicate the grouping of crimes under federal jurisdiction. The proximity matrix can aid in identifying the best crime to prevent, considering its relationship with other types of crimes.

---

possibility, a VEC model was estimated, and the following model was used: $\Delta Y_{i,t} = \alpha + \sum_{k=1}^{K} \Pi_{t-k} \Delta Y_{i,t-k} + \lambda \left(Y_{2,t-h} - \alpha\lambda - \beta\lambda Y_{1,t-h} \right) + u_{i,t}$. Therefore, we introduced a correction term error in the VEC model to account for the potential long-term relationship between $Y_{1,t-h}$ and $Y_{2,t-h}$ that may still exist after differencing, . The original and the VEC models are consistent and yield similar results concerning the interactions between crime complexity and the Granger causality test. The correction term used in the VEC model was $Y_{2,t-h} - \alpha\lambda - \beta\lambda Y_{1,t-h} + u_{i,t}$. It was chosen because it maximizes the R^2 compared to the alternative choice of $Y_{1,t-h} - \alpha\lambda - \beta\lambda Y_{2,t-h}$. Furthermore, to ensure consistency with the differenced data, the error term was estimated with 4 lags ($h = 4$).

Figure 1
Crime space
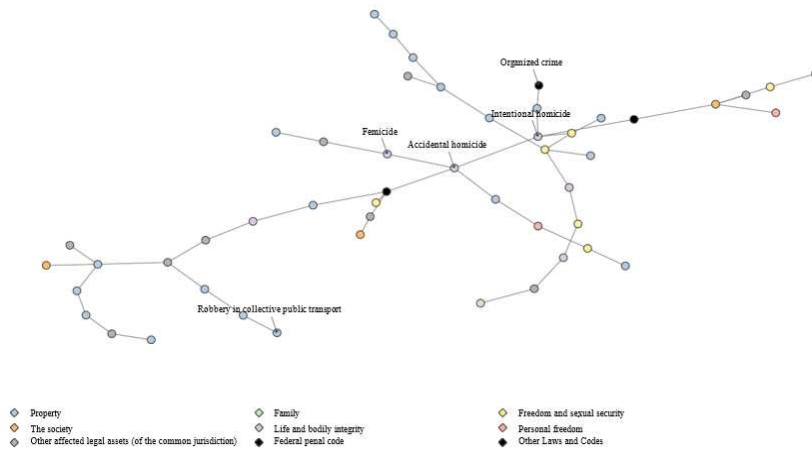
Notes: The color assigned to each node represents the category of goods affected by the crime, as defined by the SESNSP. Only the strongest connections between the nodes, determined using the Kruskal algorithm, are displayed. The Fruchterman-Reingold algorithm was used to position the nodes in the graph. It should be noted that the graph is a simplified representation of the proximity matrix.

Source: Own elaboration.

Figure 1 shows that different types of robberies and crimes, such as threats and fraud are grouped together (as represented by blue and gray nodes), suggesting that these crimes require similar capabilities to be committed. Notably, the blue node in the central area of the graph represents the crime of the least complexity (robbery in collective public transport), while the black node in the upper part of the graph represents the crime of the highest complexity (organized crime). Thus, figure 1 can also be interpreted as a path of complexity and diversification of crimes, highlighting the relationships between them. Furthermore, the graph serves as a two-dimensional evolution map of crimes that considers similar capabilities, not solely based on complexity.

In figure 1, accidental homicide, intentional homicide, and femicide are depicted as immediately connected, indicating that these three crimes share similarities in terms of the capabilities required to commit them (purple nodes at the center of the graph). It is important to emphasize that this clustering occurs despite the model's lack of information regarding the fact that all three crimes ultimately

result in loss of life. As a result, the joint probability of a homicide being committed is more closely related to femicide than to other crimes.

The preceding analysis reveals a significant finding regarding the escalation of gender-based violence. By examining the closest relationships or highest joint probabilities among crimes, it is observed that gender violence, represented by the green node at the bottom of figure 1, is immediately connected to malicious injuries, sexual harassment, and sexual abuse and finally culminates in rape. This pattern of escalation accurately reflects the behavior of sexual violence and is appropriately captured by the model. The result sheds light on the underlying dynamics of gender violence and provides valuable insights for policymakers and practitioners to develop effective prevention and intervention strategies.

Furthermore, the network analysis conducted in this study demonstrates that femicide and rape are the most central crimes in the crime network, as indicated by their high eigenvector centrality (as previously discussed in section 2). Specifically, these crimes exhibit the highest conditional probabilities with other crimes, implying that addressing them can generate positive externalities in preventing other related crimes. Table 1 provides a summary of the centrality metrics for each crime. Notably, in figure 1, these crimes are represented by the central purple nodes and their closest associated yellow node. It is worth noting that the centrality measure was computed using the proximity matrix that considers all interactions among crimes, as opposed to only the strongest ones depicted in figure 1. The use of the closeness centrality measure as an alternative method also resulted in femicide and rape being identified as the top two most central nodes, but not in the same order.

Thus, it is plausible to argue that preventing these two crimes could benefit public safety. Specifically, decreasing the incidence of either of these crimes will likely yield the greatest reduction in the marginal probabilities of other crime incidences.

To address the complexity inherent in the creation of a crime, a comprehensive preventive policy must consider various aspects of society, including economic and social environments, institutional development, and other factors. By doing so, such a policy could effectively reduce the capabilities required for the particular offense under consideration as well as those common to other offenses. Specifically, given the centrality of femicide and rape in the crime network, a desirable policy in this context should prioritize gender violence, with a particular focus on these two offenses. By incorporating these key

elements and addressing the underlying complexity of crime, such a policy could have a significant impact on overall security.

On the other hand, proximity alone does not fully capture the complexity inherent in the evolution of crime. However, even when only considering the complexity, preventing gender violence remains an effective strategy since it can prevent the creation of more complex crimes, such as trafficking in minors, crimes against health, and kidnappings, among others. In other words, preventing femicide and rape has implications for crime prevention not only in the present but also in the future, as it may prevent the occurrence and diversification of more severe crimes. Therefore, implementing policies to prevent gender violence should be a priority in the fight against crime.

In this context, prevention efforts should begin with addressing gender violence, followed by sexual harassment, then rape, and finally, femicide. In this case, femicide and rape incorporate all of the capabilities of the preceding crimes. Consequently, if a criminal does not develop the capabilities incorporated in gender violence, the probability of committing more complex crimes, such as drug trafficking, is diminished. Alternatively, prevention of gender violence may also prevent childhood maltreatment, a risk factor for future violent crimes such as homicides, rape, and drug offenses.

The relationship between femicide and several crimes has been highlighted by the World Health Organization (who). For instance, who (2012) identifies femicide as being associated with verbal harassment, emotional abuse, physical or sexual abuse, and murder. Furthermore, who (2012) also identifies several risk factors that are associated with femicide, including unemployment, gun ownership, problematic alcohol and drug consumption, gender inequality, and low government social spending. These factors can all be considered as capabilities that enable the incidence of femicides.

In the context of violence prevention, the impact of violence against women extends beyond immediate harm to the individuals affected, as it can have lasting effects on future generations and society as a whole. According to Guedes and Mikton (2013), exposure to violence against women during childhood increases the risk of experiencing or perpetrating different forms of violence later in life. Additionally, such disclosure has been linked to child maltreatment and abuse. Moreover, a longitudinal study by Widom and Maxfield (2001) -on abused and neglected children during the 1970s- found that this violence increases the likelihood of future delinquency and criminality by 29%. These findings underscore the need for comprehensive prevention policies that prioritize reducing violence against

women and its associated negative outcomes.

The evidence demonstrates that violence against women can be a central factor in current and future violence and crime. As a result, preventing femicides and violence against women can create a societal environment that discourages criminal behavior now. This approach creates conditions that decrease the likelihood of violent behavior in the future.

4.3 *State crime complexity*

From a geographic standpoint, the states of Mexico, Mexico City, and Jalisco emerge as having the highest crime rates in the country. Over the period from 2015 to 2019, these three states accounted for 48% of all crimes committed in Mexico. However, it is relevant to note that the complexity of the crime does not solely capture the overall aggregation of crime in each state. Instead, the complexity analysis aims to standardize the comparison of each state in terms of the types of crimes committed. Thus, a state's complexity score reflects the prevalence of the crimes committed in that state relative to the entire nation. Specifically, states that demonstrate a higher proportion of complex crimes relative to the whole country will be classified as more complex compared to states where this is not the case. This has significant implications for policy formulation and intervention, allowing policymakers to identify areas of higher complexity and target resources towards more effective crime prevention strategies.

According to the analysis, Tlaxcala, Nayarit, Sonora, Campeche, Sinaloa, Chihuahua, Tamaulipas, Guerrero, Hidalgo, and Michoacán are the most complex states in terms of crime. The complexity of each state is determined by the proportion of homicides, rapes, and organized crime in their total crime rate relative to the national average. Conversely, the states of Mexico, Quer'etaro, and Mexico City are characterized by a basket of less complex crimes due to the high proportion of robberies in their total crimes. This aligns with the observation that regions with higher economic activity tend to experience more robberies. The study found a positive correlation between economic growth, measured by the Quarterly Indicator of State Economic Activity (*Indicador Trimestral de la Actividad Económica Estatal*, ITAEE), and robberies. Moreover, the correlation between robberies and ITAEE is much stronger than that between high-complexity crimes and ITAEE, suggesting that economic growth is more strongly associated with low-complexity crimes. These findings are presented in tables 3 and 4.

Table 3
Correlation between robbery, high-complexity crimes (HCC) and ITAEE

	ITAEE	*Robbery*	*HCC*
ITAEE	1.00		
Robbery	0.41	1.00	
HCC	0.08	0.24	1.00

Notes: The normalization of the Quarterly Indicator of Economic Activity (ITAEE) is based on a standard of 100 for the year 2013, and the average ITAEE was calculated for each state from 2015 to 2019.

Source: Own elaboration.

Table 4
State crime complexity (SCC)
(Order from high-complexity to low-complexity)

1. Tlaxcala	11. Guanajuato	21. Chiapas	31. Querétaro
2. Nayarit	12. Zacatecas	22. Puebla	32. México
3. Sonora	13. Colima	23. Baja California	
4. Campeche	14. Quintana Roo	24. Tabasco	
5. Sinaloa	15. Veracruz	25. Morelos	
6. Tamaulipas	16. Baja California Sur	26. Yucatán	
7. Chihuahua	17. Nuevo León	27. Coahuila	
8. Michoacán	18. Oaxaca	28. Jalisco	
9. Guerrero	19. San Luis Potosí	29. Aguascalientes	
10. Hidalgo	20. Durango	30. Ciudad de México	

Source: Own elaboration.

The relationship between low-complexity crimes and economic growth underscores the interplay between economy and crime evolution. The SCC is strongly connected to underlying attributes not examined in this study, such as the institutional environment and justice system, among other factors.

5. Conclusion

This study aims investigate the development of criminal capabilities using the economic complexity methodology and provide an order for the evolution of criminal behavior. The research also offers a framework to describe the distribution and diversification of crime across different regions and insights into the evolution of criminal behavior and effective preventive measures to combat it. This quantitative approach to characterizing criminal behavior contributes to the literature on economic complexity, machine learning applications, and public policy analysis related to crime.

The findings of this study suggest the need for reevaluating conventional approaches to preventive policy and emphasize the importance of understanding the diversification of crime. A proposed strategy to inhibit crime diversification in Mexico entails preventing central crimes in the crime network. The analysis reveals that femicide and rape are the most influential crimes in the criminal network. Thus, preventing violence against women and femicides could lead to a societal environment that discourages criminal activity overall. The results of this research highlight the pivotal role of violence against women in the realm of crime in Mexico and stress the critical importance of preventing femicide. The significance of preventive measures is considerable, as they present the opportunity for a centralized policy approach to protect women against all forms of violence in Mexico.

Moreover, this study presents evidence that raises awareness of the impact and influence of gender violence and advocates for a comprehensive approach to address it. The results encourage us to reflect on our societal behavior in response to these crimes and the factors contributing to the proliferation of gender violence. This work aims to provide tools for preventive action against criminal conduct and, more significantly, to provide evidence of the significance of gender violence in society.

References

Aihara, K. and B. Kim. 2009. Comparison of immunization strategies in geographical networks, *Physics Letters A*, 373(42): 3877-3882.

Albeaik, S., M. Kaltenberg, M. Alsaleh, and C.A. Hidalgo. 2017. Improving the economic complexity index, ArXiv preprint, arXiv:1707.05826.

Becker, G.S. 1968. Crime and punishment: An economic approach, *Journal of Political Economy*, 76(2): 169-217.

Benyishay, A. and S. Pearlman. 2013. Homicide and work: The impact of Mexico's drug war on labor market participation, SSRN, https://ssrn.com/abstract=2302437.

Benyishay, A. and S. Pearlman. 2014. Crime and microenterprise growth: Evidence from Mexico, *World Development*, 56: 139-152.

Braunstein, A., L. Dall'Asta, G. Semerjian, and L. Zdeborová. 2016. Network dismantling, *Proceedings of the National Academy of Sciences*, 113(44): 12368-12373.

Brown, L. and K. Hibbert. 2017. The effect of crime on foreign direct investment: A multi-country panel data analysis, *The Journal of Developing Areas*, 51(1): 295-307.

Buonanno, P. 2003. The socioeconomic determinants of crime. A review of the literature, Working Paper No. 63, Department of Economics, University of Milano-Bicocca.

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

Chalfin, A. 2015. Economic costs of crime, in W.G. Jennings (ed.), *The Encyclopedia of Crime and Punishment*, New York, John Wiley and Sons.

Che, N.X. 2020. Intelligent export diversification: An export recommendation system with machine learning, IMF Working Paper No. 2020/175.

González, S. 2016. Criminalidad y crecimiento económico regional en México, *Frontera Norte*, 26(51): 75-111.

Guedes, A. and C. Mikton. 2013. Examining the intersections between child maltreatment and intimate partner violence, *Western Journal of Emergency Medicine*, 14(4): 377-379.

Han, J., S. Tang, Y. Shi, L. Zhao, and J. Li. 2021. An efficient layer node attack strategy to dismantle large multiplex networks, *The European Physical Journal B: Condensed Matter and Complex Systems*, 94(3): 1-8.

Hausmann, R., C.A. Hidalgo, S. Bustos, M. Coscia, A. Simoes, and M.A. Yildirim. 2013. *The Atlas of Economic Complexity: Mapping Paths to Prosperity*, Cambridge, MIT Press.

Hausmann, R. and B. Klinger. 2007. The structure of the product space and the evolution of comparative advantage, CID Working Paper No. 146, Kennedy School of Government, Harvard University.

Hidalgo, C. 2021. Economic complexity theory and applications, *Nature Reviews Physics*, 3: 92-113.

Hidalgo, C. and R. Hausmann. 2009. The building blocks of economic complexity, *Proceedings of the National Academy of Sciences*, 106(26): 10570-10575.

Jaitman, L., D. Caprirolo, R. Granguillhome, P. Keefer, T. Leggett, J. Lewis, J.A. Mejía-Guerra, M. Mello, H. Sutton, and I. Torre. 2017. *The Costs of*

Crime and Violence: New Evidence and Insights in Latin America and the Caribbean, Washington D.C., Inter-American Development Bank.

Kertész, J. and J. Wachs. 2020. Complexity science approach to economic crime, *Nature Reviews Physics*, 3: 70-71.

Li, C., G. Jiang, and Y. Song. 2014. Epidemic spreading in dynamic small-world networks with community structure, *Complex Systems and Complexity Science*, 11: 33-39.

Mealy, P., J.D. Farmer, and A. Teytelboym. 2018. A new interpretation of the Economic Complexity Index, INET Oxford Working Paper No. 2018-04.

México Evalúa. 2020. *Fallas de Origen 2020: Índice de Confiabilidad de la Estadística Criminal*, Mexico City, México Evalúa.

Neto, J. 2017. Social network analysis and organised crime investigation: Adequacy to networks, organised cybercrime, Portuguese framework, in E.C. Viano (ed.), *Cybercrime, Organized Crime, and Societal Responses*, Springer.

Olavarria-Gambi, M. 2007. The economic cost of crime in Chile, *Global Crime*, 8(4): 287-310.

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.

Papachristos, A.V., D.M. Hureau, and A.A. Braga. 2013. The corner and the crew: The influence of geography and social networks on gang violence, *American Sociological Review*, 78(3): 417-447.

Prieto-Carrón, M., M. Thomson, and M. Macdonald. 2007. No more killings! Women respond to femicides in Central America, *Gender and Development*, 15(1): 25-40.

Ren, X.L., N. Gleinig, D. Helbing, and N. Antulov-Fantulin. 2018. Generalized network dismantling, *Proceedings of the National Academy of Sciences*, 116(14): 6554-6559.

Ríos, V. 2017. Crime and violence effects on economic diversity: The case of Mexico's drug war, unpublished document.

Salinas, M.G. 2021. Proximity and horizontal policies: The backbone of export diversification, IMF Working Paper No. 2021/064.

Scott, J. and P.J. Carrington. 2011. *The SAGE Handbook of Social Network Analysis*, California, SAGE Publications Ltd.

Sierra-Arévalo, M. and A. Papachristos. 2015. Social network analysis and gangs, in S.H. Decker and D.C. Pyrooz (eds.), *The Handbook of Gangs*, New York, John Wiley and Sons.

SESNSP. 2019. Incidencia delictiva, https://www.gob.mx/sesnsp/acciones-y-programas/incidencia-delictiva-299891?state=published

Torres, R.C., A.V. Mollick, and E. Saucedo. 2018. The impact of crime and other economic forces on Mexico's foreign direct investment inflows, Banco de México, Working Paper No. 2018-24.

Verdasca, J., M. Telo da Gama, A. Nunes, N. Bernardino, J. Pacheco, and M. Gomes. 2005. Recurrent epidemics in small world networks, *Journal of Theoretical Biology*, 233(4): 553-561.

WHO. 2012. Understanding and addressing violence against women: Intimate partner violence, https://www.who.int/publications/i/item/WHO-RHR-

12.36.

World Bank. 2021. Intentional homicides (per 100 000 people), https://data.world
bank.org/indicator/VC.IHR.PSRC.P5.

Widom, C. and M. Maxfield. 2001. An update on the cycle of violence, NIJ
Research in Brief No. 184894, National Institute of Justice.

Appendix A: Data

Table A1

Number of crimes from 2015 to 2019

Year	Crime Type / Subtype	2015	2016	2017	2018	2019	Total	(%)	Used
1	Robbery / Other Robbery	177,081	228,395	228,387	224,188	214,496	1,072,547	10.825	
2	Other crimes of the State Jurisdiction / Other...	186,171	210,198	198,850	180,016	183,420	958,655	9.676	
3	Robbery / Motor vehicle theft	160,944	174,812	210,845	213,991	187,657	948,249	9.571	
4	Family violence / Family violence	127,424	153,893	169,579	180,187	210,188	841,271	8.491	No
5	Injuries / malicious injuries	142,447	138,712	154,778	157,416	166,518	759,871	7.669	
6	Property damage / Property damage	120,246	122,785	129,390	132,692	135,780	640,893	6.468	
7	Robbery / Comercial robbery	81,553	78,256	109,725	114,336	118,167	502,037	5.067	
8	Robbery / Home Invasion Robbery	91,346	87,915	92,947	83,694	82,515	438,417	4.425	
9	Threats / Threats	68,007	76,043	82,682	89,788	110,092	426,612	4.306	
10	Robbery / Robbery on public roads	80,511	68,111	86,813	95,285	83,196	413,916	4.178	
11	Fraud / Fraud	59,416	60,143	61,609	66,248	76,834	324,250	3.273	
12	Injuries / Accidental Injuries	64,669	58,594	50,104	41,121	44,028	258,516	2.609	
13	Other Crimes / Other Crimes	47,298	43,045	54,096	65,600	38,517	248,556	2.509	
14	Drug dealing / Drug dealing	32,398	30,025	45,181	58,588	70,274	236,466	2.387	
15	Other Laws and Codes / Other Laws and Codes	35,459	33,761	34,249	39,502	50,410	193,381	1.952	No
16	Breach of family assistance obligations / Brea...	27,894	25,478	23,558	21,592	24,425	122,947	1.241	No
17	Dispossession / Dispossession	21,618	21,766	24,702	25,618	28,435	122,139	1.233	
18	Homicide / Intentional Homicide	16,121	20,149	25,035	29,098	29,484	119,887	1.21	
19	Breach of trust / Breach of trust	20,013	20,739	24,424	24,741	27,456	117,373	1.185	
20	Counterfeiting / Forgery	14,013	15,084	17,875	21,124	19,888	87,984	0.888	
21	Sexual abuse / Sexual abuse	11,980	15,090	15,899	18,875	23,660	85,504	0.863	
22	Kidnapping / Kidnapping	12,298	14,135	17,218	19,841	21,941	85,433	0.862	
23	Robbery / Robbery in collective public transport	17,097	11,744	16,100	20,044	18,297	83,282	0.841	

Year	Crime Type / Subtype	2015	2016	2017	2018	2019	Total	(%)	Used
24	Crimes committed by public servants / Crimes c...	11,821	13,471	15,164	18,332	21,523	80,311	0.811	
25	Homicide / Accidental Homicide	15,212	15,262	16,291	15,390	15,377	77,532	0.783	
26	Robbery / Robbery of auto parts	10,096	9,037	15,628	18,090	20,263	73,114	0.738	
27	Robbery / Robbery in individual transport	11,709	9,039	14,504	14,806	15,852	65,910	0.665	
28	Rape / Rape	10,538	10,992	10,786	12,360	13,665	58,341	0.589	
29	Other crimes against the family / Other crimes...	8,561	10,753	11,369	10,239	12,056	52,978	0.535	No
30	Robbery / Robbery of delivery people	7,258	8,769	12,031	13,068	11,662	52,788	0.533	
31	Other crimes against property / Other crimes a...	3,813	6,530	6,696	10,902	14,495	42,436	0.428	
32	Against Health / Against Health	8,687	6,211	6,323	6,429	7,507	35,157	0.355	
33	Extortion / Extortion	6,008	5,732	6,143	6,721	8,734	33,338	0.336	
34	Robbery / Theft of livestock	8,440	7,132	6,030	5,739	5,399	32,740	0.33	No
35	Other crimes that threaten life and physical i...	3,122	3,915	5,141	7,232	8,490	27,900	0.282	
36	Other crimes that violate sexual freedom and s...	4,899	4,286	4,819	4,821	6,331	25,156	0.254	
37	Other crimes against society / Other crimes ag...	2,546	3,216	3,403	5,549	7,983	22,697	0.229	
38	Robbery / Street robbery	2,625	2,534	3,105	3,496	4,044	15,804	0.16	
39	Falsehood / Falsehood	2,568	2,768	2,953	3,167	3,512	14,968	0.151	
40	Statutory rape / Statutory rape	2,081	2,547	2,734	2,962	3,677	14,001	0.141	
41	Robbery / Robbery in individual public transport	2,244	1,948	2,254	2,414	2,898	11,758	0.119	
42	Gender violence in all its forms different fro...	1,645	1,893	2,142	2,255	3,180	11,115	0.112	
43	Sexual harassment / Sexual harassment	1,109	1,325	1,700	2,712	4,220	11,066	0.112	
44	Corruption of minors / Corruption of minors	1,940	1,868	1,851	1,848	2,186	9,693	0.098	
45	Against the environment / Against the environment	1,184	1,449	1,807	2,120	2,192	8,752	0.088	
46	General Health Law (L.G.S.) / General Health L...	3,114	1,431	1,065	1,101	1,578	8,289	0.084	No

Table A1
(Continued)

Year	Crime Type / Subtype	2015	2016	2017	2018	2019	Total	(%)	Used
47	Robbery / Robbery of machinery	1,129	1,094	2,195	1,057	1,170	6,645	0.067	
48	Sexual coercion / Sexual coercion	784	929	1,052	1,276	1,862	5,903	0.06	
49	Electoral / Electoral	738	472	1,351	2,315	677	5,553	0.056	
50	Femicide / Femicide	411	605	742	893	943	3,594	0.036	
51	Abortion / Abortion	544	562	547	604	717	2,974	0.03	
52	Robbery / Bank robbery	614	569	535	394	475	2,587	0.026	
53	Federal Law Against Organized Crime (L.F.C.D.O...	1,051	631	221	199	265	2,367	0.024	
54	Human trafficking / Human trafficking	415	383	304	387	544	2,033	0.021	
55	Abduction / Abduction for a sexual purpose	271	254	204	119	116	964	0.01	No
56	Hindering prosecution or apprehension / Hinder...	119	153	104	129	127	632	0.006	
57	Trafficking of minors / Trafficking of minors	96	203	176	40	29	544	0.005	
58	Incest / Incest	17	73	35	10	14	149	0.002	No

Notes: Type and subtype of crimes are presented in case they are crimes of state jurisdiction. In the case of being a federal crime, only the concept is presented.

Source: Own elaboration with data from SESNSP (2019).

Appendix B: Unreported crime robustness test

Estimating economic complexity can be challenging due to unreported crimes (black figures). Therefore, to determine the robustness of the results obtained, an estimate of economic complexity was made that considers the unreported crimes. The data on unreported crimes were obtained from various sources, including the National Survey of Victimization and Perception of Public Safety (*Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública*, ENVIPE) for 2019 and México Evalúa.

A correction matrix, denoted as Z, was constructed based on the black figures to incorporate the unreported crimes into the analysis. Each matrix element is defined as the ratio of 1 to 1 minus the percentage of the black figure for the corresponding crime. The correction can thus be expressed as follows:[15]

$$M^1 = M^o \ o \ Z$$

For each column, a constant value was defined (through rows or states) equal to the black figure corresponding to ENVIPE. In general, the defined black figure was 93.2%.[16] For some crimes, it is defined as a black figure the following: extortion, 97%; theft in transport and from bystanders, 94%; abuse, harassment, and sexual harassment, 96%; and finally, rape in general, 60%. In the case of homicides, the black figure was 10% taken (calculated from the differential percentage between homicides in the SESNSP and INEGI). In the case of femicide, the black figure was 95%. Finally, in the case of vehicle robbery, it was not taken as constant across states. Since it is intended to capture the variance between states, the value of the ratio of 1 over 1 minus the percentage of uninsured vehicles was taken as an approximate variable of the black figure. This results in a 70% black figure for this crime. The auto-insured data was taken from México Evalúa.

The estimate of economic complexity and the associated metrics were computed using the Z matrix. The complexity and centrality of the analysis are presented below with the correction for black figures.

In terms of complexity, the results exhibit a slight modification. However, as in the previous analysis, the centrality results remain consistent for the three most complex crimes. Nevertheless, femicide

---

[15] The symbol o represents the product of arrays element by element.

[16] In the case that the black figure is constant across all crimes, the complexity algorithm is not affected.

and rape were found to be the second most central crimes, in contrast to the previous finding in which they were the most central crimes.

Table B1

Complexity and centrality by crime.
Correction for black figures

Crime Type / Subtype	Complexity	Centrality
1 Trafficking of minors / Trafficking of minors	1.136555	0.0839
2 Homicide / Accidental Homicide	0.967376	0.173492
3 Against Health / Against Health	0.935916	0.169058
4 Homicide / Intentional Homicide	0.901656	0.173851
5 Kidnapping / Kidnapping	0.881143	0.166975
6 Drug dealing / Drug dealing	0.820072	0.13981
7 Federal Law Against Organized Crime (L.F.C.D.O...	0.796079	0.134454
8 Corruption of minors / Corruption of minors	0.763013	0.165931
9 Rape / Rape	0.720199	0.177866
10 Other Crimes / Other Crimes	0.718888	0.168989
11 Robbery / Bank robbery	0.700652	0.168407
12 Femicide / Femicide	0.698257	0.176359
13 Robbery / Motor vehicle theft	0.688323	0.157657
14 Other crimes that violate sexual freedom and s...	0.589788	0.174945
15 Hindering prosecution or apprehension / Hinder...	0.588714	0.168115
16 Property damage / Property damage	0.557447	0.160076
17 Human trafficking / Human trafficking	0.511752	0.138579
18 Robbery / Robbery of machinery	0.505751	0.147968
19 Abortion / Abortion	0.460273	0.14717
20 Robbery / Home Invasion Robbery	0.459713	0.156807
21 Dispossession / Dispossession	0.439741	0.176019
22 Breach of trust / Breach of trust	0.415453	0.166159
23 Falsehood / Falsehood	0.415282	0.178592
24 Crimes committed by public servants / Crimes c...	0.411745	0.144214

Table B1
(Continued)

Crime Type / Subtype	Complexity	Centrality
25 Sexual abuse / Sexual abuse	0.372021	0.173253
26 Sexual coercion / Sexual coercion	0.365902	0.14606
27 Other crimes that threaten life and physical i...	0.354491	0.160006
28 Robbery / Street robbery	0.215575	0.140339
29 Statutory rape / Statutory rape	0.211871	0.160992
30 Extortion / Extortion	0.171198	0.158187
31 Sexual harassment / Sexual harassment	0.089244	0.166037
32 Fraud / Fraud	0.05739	0.156104
33 Threats / Threats	-0.022396	0.146287
34 Other crimes against property / Other crimes a...	-0.074515	0.132631
35 Injuries / malicious injuries	-0.116812	0.144444
36 Injuries / Accidental Injuries	-0.117075	0.125796
37 Robbery / Other Robbery	-0.301969	0.123253
38 Robbery / Robbery of delivery people	-0.385175	0.078776
39 Other crimes of the State Jurisdiction / Other...	-0.44087	0.062192
40 Electoral / Electoral	-0.441485	0.130804
41 Against the environment / Against the environment	-0.507775	0.108255
42 Robbery / Comercial robbery	-0.867636	0.09443
43 Counterfeiting / Forgery	-1.087531	0.079192
44 Robbery / Robbery in individual public transport	-1.199404	0.117535
45 Gender violence in all its forms different fro...	-1.225338	0.037715
46 Other crimes against society / Other crimes ag...	-1.350752	0.0991
47 Robbery / Robbery of auto parts	-1.379445	0.068122
48 Robbery / Robbery on public roads	-1.603959	0.07865
49 Robbery / Robbery in individual transport	-2.523384	0.04891
50 Robbery / Robbery in collective public transport	-4.275963	0.024878

Notes: The type and subtype of the crimes are presented when they are under state jurisdiction. In the case of a federal crime, only the concept is presented. Other Crimes includes "huachicoleo", among other federal offenses.

Source: Own elaboration.