Analysis of opioid drug transmission combined with BVAR model

Bing Tu[a], Lihui Li

Jiangxi University of Finance and Economics, Nanchang, China

[a]841348401@qq.com

Keywords: NFLIS; opium control; socio-economic; Matlab

Abstract: This article is about the issue of opium control management in the United States. Our aim is to analyze a series of models such as differential equations through NFLIS and the socio-economic data provided by the US Census, and to optimize them. When solving this problem, we use Python comprehensively, MATLAB, Eviews, stata, Excel software to determine the propagation characteristics of opioids in a given area and predict the possible location, and combined with the given socio-economic data improvement model to give corresponding strategies to effectively control opioids The amount of use is reduced.

1. Introduction

The United States is now facing an opioids crisis, a variety of synthetic and non-synthetic opioids are rampant, and opioids are eroding people's mind and body. It has seriously affected the economic development of the United States, how to effectively and rationally prevent drugs, and help different levels of drug addicts. Treatment, control of the spread of drugs, etc. is the main problem of reducing opium flooding.

In this passage, we build SIS model to describe the transmission of opioids according to the idea of epidemiological dynamics compartment model. We prove the stability of SIS model by using Lyapunov method. Then, based on latitude and longitude coordinates, using MATLAB software to draw the distribution map of opioid use in each state. In the next time, the less used red label will be the place where specific opioids may have been used.

2. Problem Analysis

We analyzed the problems above. Then we took some important features into considered. According to those features, we made necessary assumptions. In order to simplify the model and highlight the main features of the problem, we made the following assumptions.

1) Assume that a person who is addicted to drugs is not immune, that is, an addict is cured and becomes a healthy person, and a healthy person can be infected and become a patient;

2) Assume that death due to illness is snot considered, and migration is not considered, that is, the input population is constant, and the birth rate is equal to the mortality rate.;

3) In the SIR model, assume that the immunized person is evenly distributed among the entire population;

4) The filtered and cleaned data is correct and robust for further analysis;

5) Assume that model error terms follow blue rules;

6) The following PA stands for Tennessee.

  DOI: 10.25236/iclsmh.2019.032

3. Symbol Description

Table 1. Symbol Descriptio

Symbol	Description
t	Time
$S(t)$	Proportion of the easily infected in total number of people at time t
$i(t)$	Proportion of the infected in total number of people at time t
λ	Daily contact rate
N	Total number of people
μ	Proportion of the number of people cured in total number of people
δ	Number of contacts
S	Compartment of the easily infected
$I_1 \ldots I_n$	Compartment of the infected
R	Compartment of the infected after cured
S_n	Average metastatic rate of the infected from the nth stage to recovery stage
γ	Infection rate
$S_{k,j,p}$	Coefficient prior standard deviation of jth variable that lagged p stages in kth equation
η	Total tightness
g(p)	Tightness of p-order lag variable relative to first-order variable
d	Decay factor
f(k,j)	Relative tightness
$\frac{Se_k}{Se_j}$	Ratio of the autoregressive residual standard deviation of the kth sequence $\{y_{kt}\}$
$c_{kj}^{(q)}$	Impulse response function
m	Region
ε	Random disturbance term

4. Synthetic opioid and heroin propagation model based on epidemiology

4.1 Building a compartment model

According to the epidemiological dynamics compartment modeling idea, the population can be divided into two types: Susceptible and Infective (take the first letter of the two words, called SI model), time t The proportion of these two types of people in the total number of people is recorded as s (t) and i (t). The average number of effective contacts per patient per day is a constant λ, called the daily exposure rate. When the patient is in effective contact with the healthy person, the healthy person is infected and becomes a patient. According to the hypothesis, each patient can turn λs(t) healthy people into patients every day. The number of patients is Ni(t), a total of λs(t) healthy people are infected every day, so the rate of increase in the number of patients λNs(t)i(t). That is said:

$$N\frac{di}{dt} = \lambda Nsi - \mu Ni \tag{1}$$

The proportion of patients who are cured every day is a constantμ, called the daily cure rate. After the patient is cured, it becomes a healthy person who can still be infected. 1/μ is the average infection period of this infectious disease. It is not difficult to see in this time, the (1)formula of the SI model should be corrected to

$$N\frac{di}{dt} = \lambda Nsi - \mu Ni$$

Then

$$s(t) + i(t) = 1$$

At the beginning ,the proportion of patients in population is i_0, then

$$\frac{di}{dt} = \lambda i\ (1 - i) - \mu i,\ i(0) = i_0$$

The introduction of δ represents the average number of effective contacts per patient throughout the infection period, called the number of contacts, expressed as $\sigma = \lambda / \mu$, so the equation can be further rewritten as

$$\frac{di}{dt} = -\lambda i[i - (1 - \frac{1}{\sigma})]$$

On the basis of the above modeling, the whole population is divided into four compartments: the vulnerable person's compartment S, the infected person's compartment $I_1, \cdots, I_n$, the healthy person's compartment R, and then the infected person's compartment S'. Therefore, the compartment model can be established as shown in Figure 1.

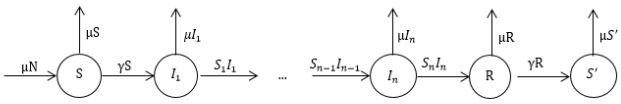


Figure 1 Compartment model

The model of the synthesis of opioids and heroin can be restated as:

$$\begin{cases} \dfrac{di(t)}{dt} = -\lambda s(t)i(t) + \mu i(t) \\ \dfrac{ds(t)}{dt} = \lambda s(t)i(t) - \mu i(t) \end{cases} \tag{2}$$

Among them i(t) is the number of people who have been infected at the moment t, s(t) is the number of vulnerable people at the moment t, and N(t) is the total population of the region at the moment t. The curve based on Python that can draw the SIS model is shown in Figure 2. It seems that the ratio of opioid addicts to the total population and the ratio of addicts to the total population tend to be a stable value over time, and the sum of the two is equal to one.

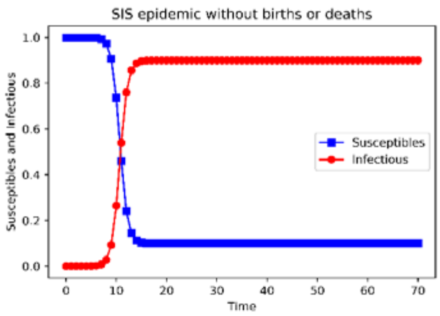


Figure 2 SIS model

4.2 Balance point existence and stability analysis

Let the right end of the system be 0, that is, let:

$$\begin{cases} -\lambda s(t)i(t) + \mu i(t) = 0 \\ \lambda s(t)i(t) - \mu i(t) = 0 \end{cases}$$

We can get the solution: $s(t) = \frac{1}{\lambda}$, for λ is a constant and $\lambda > 0$, Therefore, this solution must exist, so it proves the existence. Two equilibrium solutions are $i(t) = N - \frac{1}{\lambda}, s(t) = \frac{1}{\lambda}$.

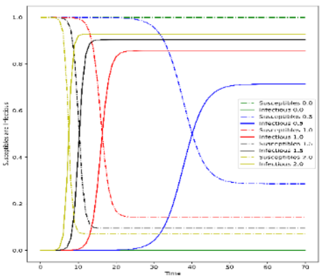
Then we use Lyapunov method to analyze its stability, we can get the equation set:

$$\begin{cases} \frac{d^2 i(t)}{dt^2} = -\lambda s'(t)i(t) - \lambda s(t)i'(t) + \mu i'(t) \\ \frac{d^2 s(t)}{dt^2} = \lambda s'(t)i(t) + \lambda s(t)i'(t) - \mu i'(t) \end{cases},$$

4.3 Numerical Simulation

Numerical simulations were performed on the different cases of the differential equation system to verify the stability of the model. For the sake of convenience, the test is performed 5 times for the parameter λ and the parameter μ respectively, and the values are 0, 0.5, 1, 1.5, 2 respectively. Figure 3 shows the threshold level and final trend of the model as the value of the parameter λ changes. The rate of steady state is different; Figure 4 shows the rate of the model threshold and the rate of the final steady state as the value of the parameter μ changes.

Although the expressions of the model are different with the change of parameters, they all reach a steady state with time under the limitation of a certain threshold.



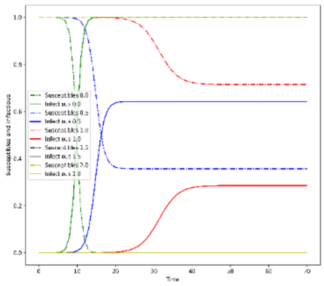
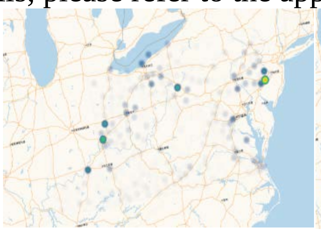


Figure 3 λ changes Figure 4 μ changes

5. A model of propagation characteristics in counties of the five States based on Bayesian vector autoregressive model

Based on the data provided by NFLIS, the distribution of DRUGREPORTS in the US map is shown based on Excel software. As shown in the figure 5 and figure 6, the distribution is divided into 2010 and 2017. For details, please refer to the appendix.



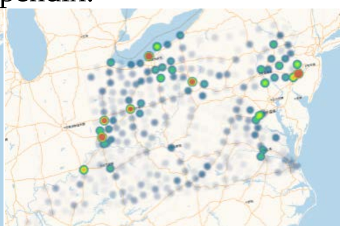


Figure 5 Initial state Figure 6 Final state

From the above evolutionary trends of drugreports between the five states in 2010-2017, it can be seen that there are more central points of opioid use and spread nearby. The amount of opioid use varies from state to state. So the next step was to develop a BVAR model for the use of opioids between the five states.

5.1 Theory basis of BVAR Model

Vector auto regressive (VAR) models are often used to analyze dynamic linkage relationships.But in VAR, the number of parameters to be estimated by the model is large, and the model should have higher accuracy based on large samples, and tend to be less effective in small samples. With the current data support, the available data is not sufficient, and the Bayesian analysis method has obvious advantages in processing small sample data. The vector autoregressive model, Bayesian vector autoregressive model, is built in the Bayesian theory framework. The model not only makes full use of the model information and sample information, but also combines the prior information in the model parameters to effectively overcome the traditional autoregression. Defects in the model and show good accuracy and predictive results.

5.2 Construction of BVAR model

In the constructed BVAR model, the j-th parameter of the kth equation lags the coefficient of the p-period with a priori standard deviation of:

$$s_{k,j,p} = \eta \cdot g(p) \cdot f(k,j) \cdot \frac{s_k}{s_j}$$

Among them, it is the overall tightness, and its value reflects the degree of grasp of the a priori information by the researcher. The smaller value indicates that the researcher has a greater grasp of the prior information. g(p) is the tightness of the p-order lag variable relative to the first-order variable, indicating that the past information is less useful than the current information, usually set to $g(p) = p^{-d}$, is attenuation coefficient，d>0。f(k , j) is relatively tight, usually taking the following form:: $f(k,j) = \begin{cases} 1, k = j \\ w_{kj}, k \neq j \end{cases}$,w_{kj} is an constant between 0 and 1.Its value reflects the relative tightness of other variables in the kth equation (excluding its lagging variables). $\frac{se_k}{se_j}$ is the ratio of the autoregressive residual standard deviation of the kth sequence $\{y_{kt}\}$.

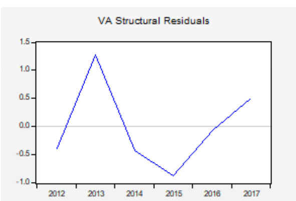
Based on the Eviews 10 software, a series of stationarity tests are carried out (see the appendix for the autocorrelation and partial correlation functions of each sequence). After the selection of the prior distribution and the determination of the lag order, the parameters of the BVAR model can be estimated and will be obtained. The BVAR model is written in the form of a matrix:

$$\begin{pmatrix} VA \\ OH \\ KY \\ WV \\ PA \end{pmatrix}_t = \begin{pmatrix} -0.00973 & 0.19037 & -0.00083 & 0.03047 & -0.01664 \\ -0.00348 & 0.61524 & -0.01118 & -0.03443 & -0.00193 \\ 0.00102 & -0.01193 & 0.00453 & 0.00076 & 0.00064 \\ -0.00087 & -0.08563 & -0.002319 & 0.00488 & 2.36795e-06 \\ 0.01036 & -0.19052 & -0.00494 & 0.00941 & -0.00199 \end{pmatrix} \begin{pmatrix} VA \\ OH \\ KY \\ WV \\ PA \end{pmatrix}_{t-1}$$

$$+ \begin{pmatrix} -0.01317 & -0.27676 & -0.000823 & -0.01645 & 0.011503 \\ 0.00094 & 0.165212 & -0.00366 & -0.00325 & -0.00270 \\ -0.00158 & 0.023721 & -0.00037 & -0.00640 & 0.001595 \\ -0.000815 & -0.04764 & 0.00133 & 0.00061 & 0.00112 \\ 0.00414 & 0.03589 & 0.00121 & 0.012001 & -0.00437 \end{pmatrix} \begin{pmatrix} VA \\ OH \\ KY \\ WV \\ PA \end{pmatrix}_{t-2}$$

$$+ \begin{pmatrix} 41987.42885 \\ 31455.36974 \\ 26020.79728 \\ 19042.60453 \\ 89134.94467 \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \\ e_{4t} \\ e_{5t} \end{pmatrix}$$

As shown in the figure 7-11, the residual model of the model is available. The residual of this model is mostly less than 1 and the model fits well.



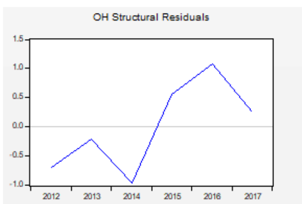
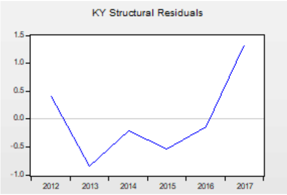


Figure 7 VA residuals　　Figure 8 OH residuals



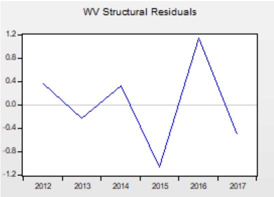


Figure 9 KY residuals   Figure 10 WV residuals

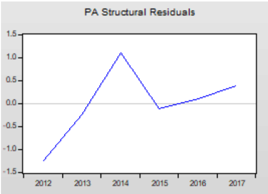


Figure 11 PA residuals

5.3 Impulse response analysis

For the BVAR model, it is often analyzed that the dynamic impact on the system when the model is subjected to a certain impact, that is, the influence of the variation or impact of each endogenous variable on itself and all other endogenous variables, can be achieved through the Impulse Response Function (IRF) is portrayed. The Bayesian analysis method can improve the accuracy of the model parameter estimation, so that the accuracy of the whole system is greatly improved, and the impulse response analysis made is closer to reality. Before doing the impulse response analysis, it is necessary to determine that the model is stable. The unit root model of each variable in Figure 12 is known in the unit circle. This model is stable and can be further analyzed by impulse response.

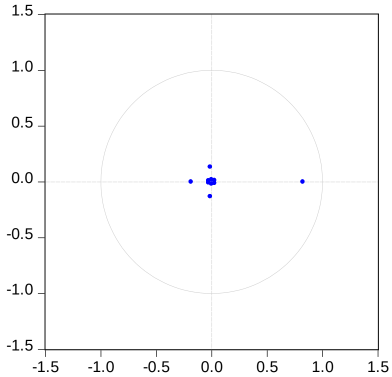


Figure 12 Inverse roots of AR characteristic polynomial

The impulse response function can be represented by $c_{kj}^{(q)}$:

$$c_{\mathrm{kj}}^{(q)} = \frac{\partial\, y_{k,t+q}}{\partial\, \varepsilon_{jt}}, q = 0,1,\ldots; t = 1,2,\ldots, T$$

It describes the change or impact of each endogenous variable on itself or other variables in the period t, when the other perturbation terms are constant, and the perturbations in other periods are constant, $y_{k,t+q}$ is the response to ε_{jt} The reaction produced. The pulse function can be used to dynamically link the opioid reports between the five states. Figure 13 shows the variables under a positive standard deviation exogenous shock, the variables themselves and other variables from 0 to 10 Impulse response results.

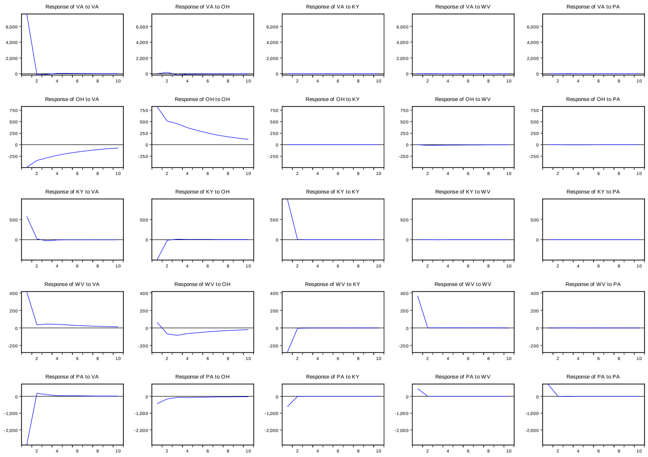


Figure 13 Inverse roots of AR characteristic polynomial

5.4 Characteristics of opioid drug transmission between five states based on impulse response analysis

(1)From the impulse response results of Section 3.3, it can be seen that the five states of drug-reports will have a strong positive response under their own impact. And when the shock begins, the response reaches its maximum and then gradually weakens. Among them, VA State drug-reports responded most to itself. This shows that the spread of drugs between the five states firstly shows the characteristics of the states' mutual spread and mutual influence among the counties in the region.

(2)As shown in Figure 14, the situation of five states' drug-reports. It can be seen that the OH state's drug-reports is on the rise, and the other states' drug-reports are less volatile. The opioids of OH are most affected by the spread of other states. Opioids travel between five states and are radiated from other large-scale cluster centers to other states.

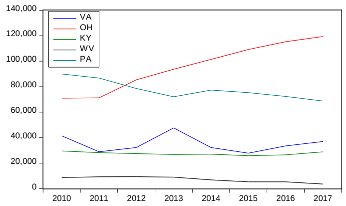


Figure 14 The five states drugereports.

5.5 Association rule mining to determine possible locations for specific opioids

Based on SPSS Modeler software, the counts of different drug names in each year and the number of initial opioids can be drawn. As shown in the figure, different colors represent different drugs, and the circle size represents different counts.

From the red scatter, we can get a preliminary understanding of the possible location of specific opioids. Based on the differential equation model, further use the correlation analysis method in data mining to output the network graph nodes on the modeler platform,as shown in the figure 15 Observe the strong link distribution field.

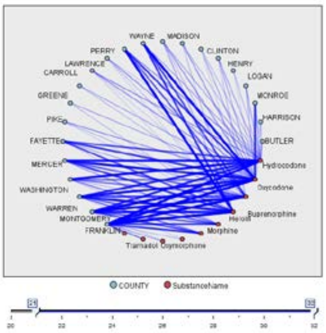


Figure 15 Network graph

6. Response to opioid crisis strategy based on Pareto analysis

(1) Based on the part1 differential equation model 2.2, in order to cope with opioids, the drug crisis can improve the health level and medical level, and improve the daily cure rate μ ;

(2) Improve the government's supervision, that is, reduce the daily contact rate λ. The healthy opioid addicts can be gradually withdrawn from the infection system to form an SIR model, at which time $s(t)+i(t)+r(t)=1$. For outpatients who are immune to illness

$$N\frac{dr}{dt}=\mu Ni$$

At this point, the SIR model is

$$\begin{cases} \dfrac{di}{dt}=\lambda si-\mu i, & i(0)=i_0 \\ \dfrac{ds}{dt}=-\lambda si, & s(0)=s_0 \end{cases}$$

And the ratio of healthy people and patients at the initial moment, s_0>0, i_0>0.

In Figure 16, the red curve represents the situation of the local infected person after the government imposes control. At this time, the opioid infected person approaches a stable state with a small value over time, and the society is in a good state. This strategy is effective but difficult to implement and requires a variety of human and financial costs.

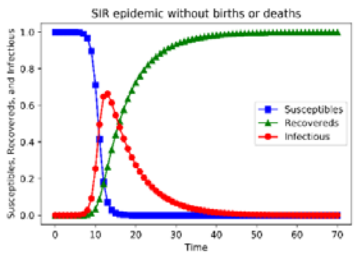


Figure 16 SIR model

(3) On the basis of strategies 1 and 2, the control parameter $\sigma<1$ is realized. At this time, the threshold $1/\sigma$ becomes larger. When $s_0\le 1/\sigma$, the opioid crisis will not spread.

(4) Pareto analysis

Taking the socio-economic factor indicator as the abscissa, the socio-economic data index affects the state-specific drug-reports to map the value to 0-150 according to the weight, and draws a histogram as the ordinate. Then, based on the cumulative composition, the map is plotted. The rise of the curve indicates the increase of the cumulative composition ratio, and finally the Pareto chart is obtained, as shown in Figure 17. Based on this, we identify the major and minor factors that contribute to the socioeconomic factors associated with the use of opioids. On the basis of the above Pareto analysis results, the corresponding strategies are proposed for the main factors and the secondary factors.

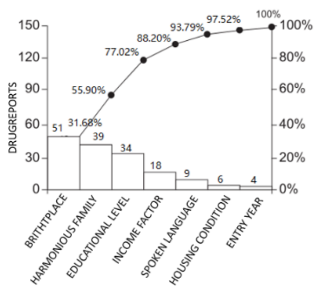


Figure 17 The Pareto chart

Therefore, in the face of the opioid crisis, it is necessary to:

- Build a harmonious family atmosphere, promote the communication between relatives and enhance mutual feelings;
- Fully popularize early childhood education and cultural education, and improve the cultural quality of family members;
- Under the influence of the birthplace factors, promote the coordinated development of counties and states, and narrow the gap between the populations of different regions due to birthplaces;
- Promote multilingual, spiritual and cultural exchanges in various states, and thus relieve and release people's pressure to some extent;
- Improve housing conditions, living environment, infrastructure construction and social welfare protection;
- Promote some local people to change the inherent concept of discrimination against migrants;
- The government should focus on increasing redistribution and promote fair income distribution.

References

[1] Lin W. The quantitative and qualitative analysis of a Tuberculosis model [D].Xian: Xian University of Science and Technology,2012

[2] Kong X, Qi Y, Song X, et al. Modeling Disease Spreading on Complex Networks [J]. Computer Science & Information Systems, 2011, 8(4):1129-1141.

[3] Xiaoning L, Ancha X. The Dynamic Interaction Between the Interest Rate Of Informal Finance and Formal Finance——Based on Empirical Research of Bayesian VAR Model[J]. Journal of Financial Development Research, 2017.

[4] Shuai Z, Van d D P . Global Stability of Infectious Disease Models Using Lyapunov Functions [J]. SIAM Journal on Applied Mathematics, 2013, 73(4):1513-1532.