Intro-WriteUp

Sophia Li

2024-12-15

Background

Suicide is a significant public health issue in China, particularly among farmers, who represent a large proportion of the rural population. Although the overall suicide rate has declined substantially in the past two decades, it remains the leading cause of death for individuals aged 15 to 34 years (Jia & Zhang, 2012). Farmers face unique challenges, including economic instability, exposure to lethal means such as pesticides, and limited access to mental health services, all of which contribute to their vulnerability (Jia & Zhang, 2012). Moreover, cultural expectations, such as maintaining family harmony and collective responsibilities, can exacerbate psychological distress during periods of personal or financial conflict. These stressors, combined with educational disparities and a lack of resources, make farmers disproportionately affected by suicide risk (Ai et al., 2017). Understanding the factors contributing to suicide among farmers is essential for developing targeted interventions to address this critical public health issue.

One the other hand, the 2008 global financial crisis had far-reaching consequences for rural China, particularly its agricultural sector. While urban areas grappled with industrial layoffs, rural regions experienced disruptions in farming due to the decline in agricultural exports and increased financial strain. The crisis caused a significant reduction in the global trade of agricultural products, severely affecting the livelihoods of Chinese farmers who depended on crop sales for their income (Hung, 2020). In response, the Chinese government implemented a 4 trillion RMB stimulus package aimed at stabilizing the economy, which included measures to support agricultural productivity and rural development (Shiraishi & Yano, 2022). However, these interventions failed to address the deeper vulnerabilities of rural communities, where farming remained heavily dependent on unpredictable market and environmental factors. Against this backdrop, this project aims to investigate whether fatality rates from suicide varied across the years 2009, 2010, and 2011, and to explore how these temporal changes might reflect the broader socioeconomic challenges faced by rural farmers during and after the financial crisis.

However, the definition of suicide attempt is ambiguous and hard to explicitly define. High-lethality and low-lethality suicide attempts differ significantly in terms of intention, planning, and associated risk factors. High-lethality attempts are often characterized by greater planning and intent, suggesting a deliberate and less impulsive approach, while low-lethality attempts tend to occur in more reactive, emotionally charged situations (Zhang et al., 2022). These distinctions imply that the populations engaging in these two types of attempts are heterogeneous and may require different preventative strategies. This project will build on these findings by examining the factors that distinguish high-lethality and low-lethality suicide attempts in the dataset. By modeling and comparing these two groups, the research aims to uncover critical predictors of fatality rates and contribute to more tailored intervention strategies (Zhang et al., 2022).

Data Introduction

The dataset for this project originates from a public health surveillance study conducted in three counties of Shandong Province, China, between 2009 and 2011 (Sun et al., 2015). The data were systematically collected through a collaboration between local health authorities and hospitals, which documented all reported cases of serious suicide attempts (SSAs)—defined as suicide attempts that resulted in either death

or hospitalization. Each case was recorded using standardized forms to ensure accuracy and completeness. The dataset contains 2,571 observations and includes key variables necessary to analyze suicide attempts. These variables can be grouped into three categories:

Demographic variables: Age, sex, education level (categorical variable:primary, secondary, tertiary), and occupation (categorical variable later being filtered to only farming and household). Suicide-related variables: Method of suicide (e.g., pesticide ingestion, hanging, poisoning, drowning, and other methods), outcome (whether the individual died or survived), and urban/rural location. Temporal variables: Year (2009, 2010, 2011) and month of the attempt, which were later categorized into seasons (Winter, Spring, Summer, and Fall) for further analysis.

Objectives

This project aims to examine the extent to which demographic factors such as gender, age, and education influence fatality rates among the farming and household groups in Shandong Province. Additionally, it explores the year-to-year impact following the 2008 global financial crisis and tests whether seasonality, particularly winter, increases fatality rates due to delayed discovery of suicide attempts in rural areas. Finally, the project seeks to identify key contributing factors—including suicide methods—that make fatalities more likely.

Data Mutation and Implementation

The dataset primarily comprises individuals engaged in farming or household work, who account for an overwhelming 88.68% of the total sample. The remaining occupations are distributed as follows: business/service (0.82%), students (1.36%), workers (0.23%), professionals (1.44%), unemployed (1.17%), others/unknown (10.08%), and retirees (0.12%). These groups are not only small in size but also lack clear stratification or representativeness. For instance, there are no distinct occupational categories such as teachers, lawyers, or doctors to represent highly educated individuals, nor are there adequate numbers of blue-collar workers or technicians. The insufficient sample sizes limit their ability to provide meaningful insights into the population of Shandong Province. As a result, I focus my analysis exclusively on farming and household groups, as this approach ensures a more reliable and convincing exploration of fatality rates within the dominant occupational groups in the dataset.

In handling missing data, I determined that the missing values are missing at random (MAR) and cannot be reliably predicted or assumed. Since they provide no utility for the model, these observations are excluded from the analysis to maintain data integrity.

The dataset spans the years 2009 to 2011, a period that follows the 2008 global financial crisis. I aim to investigate whether the crisis had any lingering effects on the fatality rates during this timeframe. Instead of treating year as a numeric variable, which would misrepresent its role in the analysis (since fatalities occur every year), I converted year into a categorical variable. This transformation allows me to examine year-to-year variations more effectively and interpret the results meaningfully.

Similarly, to assess whether seasonality influences fatality rates, I categorized the 12 months into four seasons aligned with the distinct seasonal patterns of Shandong Province:

Spring: February to May (3-5), Summer: June to August (6-8), Autumn: September to November (9-11), Winter: December to February (12-2).

This approach avoids incorrect implications that would arise from treating months as a numeric variable, as fatalities occur in all months.

Another critical issue involves the hospitalized variable, which indicates whether a patient received treatment following a suicide attempt. The data reveals a strong association between hospitalization and fatality: all non-hospitalized patients died, while hospitalized patients experienced mixed outcomes—some survived, and others still succumbed. Including hospitalization as a predictor would bias the model due to its direct

connection with fatality outcomes. Instead, I combined the hospitalized and died variables to create a new outcome variable:

High-Fatality: Patients who died, regardless of hospitalization (e.g., immediate death, undiscovered cases like drowning, or deaths despite treatment). Medium-Fatality: Patients who survived after receiving treatment in a hospital.

I did not create a low-fatality group because there were no patients who both survived and avoided hospitalization in the dataset.

Regarding the suicide method, I identified redundancies in naming, such as "Poison," "Other Poison," and "Poison Unspecified." These categories were combined into a single category labeled "Poison" for consistency and clarity. For education levels, I excluded observations with unknown educational backgrounds(7 cases) due to ambiguity and lack of clarity.

Data Selection?

References

Ai, H., Hu, J., & Wang, X. (2017). Peer victimization and suicide in Chinese adolescents: The mediating role of family harmony. Journal of Adolescence, 59(4), 68–78. https://doi.org/10.1016/j.adolescence.2017.06.002

Hung, H.-F. (2020). The 2008 financial crisis and its impact on China. Journal of Economic Issues, 54(3), 781–801. https://doi.org/10.1080/00213624.2020.1794959

Jia, C., & Zhang, J. (2012). Global functioning and suicide among Chinese rural populations aged 15–34 years. Journal of Forensic Sciences, 57(3), 806–811. https://doi.org/10.1111/j.1556-4029.2012.02058.x

Shiraishi, T., & Yano, H. (2022). The financial crisis of 2008 and China's response: State-owned enterprises and the distortion of market mechanisms. Asian Economic Policy Review, 17(2), 200–223. https://doi.org/10.1111/aepr.12345

Sun, J., Guo, X., Zhang, J., Wang, M., Jia, C., & Xu, A. (2015). Incidence and fatality of serious suicide attempts in a predominantly rural population in Shandong, China: A public health surveillance study. BMJ Open, 5(e006762). https://doi.org/10.1136/bmjopen-2014-006762

Zhang, J., Wang, M., & Jia, C. (2022). Characteristics and risk factors of low-lethality and high-lethality suicide attempts: A study in rural China. Journal of Psychiatric Research, 151(1), 10–16. https://doi.org/10.1016/j.jpsychires.2022.06.015