Determinants of Adolescent Suicidal Ideation: A Comparative Analysis of Universal Upbringing versus Contextual Moderators

Author: Youssef El Bassumi

Affiliation: Sabanci University

Course: DSA210

Date: January 9, 2026

Abstract

This study investigates the relative impact of family dynamic variables (parental warmth, monitoring, and adverse childhood experiences) versus contextual demographic factors (urbanicity, race, and neighborhood income) on adolescent suicidal ideation. Utilizing a multi-stage analytical framework comprising Exploratory Data Analysis (EDA), stratified hypothesis testing, and weighted logistic regression, the study evaluates the validity of context-specific risk theories. Results indicate that while exploratory visualizations suggested complex demographic interactions—specifically regarding rural monitoring and racial resilience—these effects did not survive rigorous statistical testing ($p > 0.05$). The final predictive model (AUC = 0.65) identifies a "Triad of Risk" driven by cumulative trauma, counterintuitively **high** parental warmth, and high neighborhood income. Notably, a suppressor effect was observed wherein neighborhood affluence emerged as a risk factor (OR ~1.1) only after controlling for family dysfunction, supporting the "affluent distress" hypothesis.

1. Introduction

Adolescent suicide remains a critical public health crisis, yet the etiology of suicidal ideation is often debated between two frameworks: **Universalist theories**, which posit that core factors like trauma and parental attachment operate similarly across all populations, and **Contextualist theories**, which argue that risk factors vary significantly by race, geography, and socioeconomic status.

This study aims to adjudicate between these frameworks by testing three specific sociological hypotheses:

1. **The "Rural Monitoring" Hypothesis:** That parental monitoring is a protective factor primarily in rural environments where independence is the norm, but less relevant in urban settings.

2. **The "Racial Resilience" Hypothesis:** That adverse childhood experiences (ACEs) are less predictive of suicidality in Black youth due to protective community buffering effects.
3. **The "Status Inconsistency" Hypothesis:** That low maternal education serves as a distinct risk factor specifically within high-income neighborhoods due to relative deprivation.

2. Methodology

2.1 Data Source and Collection

The data for this study were drawn from Wave 1 of the National Longitudinal Study of Adolescent to Adult Health (Add Health), a nationally representative longitudinal study of adolescents in grades 7-12 in the United States.

Data collection for Wave 1 took place between September 1994 and December 1995. The study employed a systematic, clustered sampling design, initially selecting 80 high schools and 52 feeder middle schools from across the U.S. to ensure representation of diverse regions, urbanities, school types, and ethnic backgrounds.

From this sample frame, a core sample of adolescents participated in detailed in-home interviews. These interviews were conducted using Audio Computer-Assisted Self-Interviewing (ACASI) technology for sensitive topics—including suicidal ideation, parental abuse, and substance use—to maximize respondent privacy and reduce social desirability bias. This dataset provides the granular "microdata" necessary to analyze the interplay between individual family dynamics and broader environmental contexts.

2.2 Data and Measures

The analysis utilized a dataset of adolescent respondents. Key measures included:

- **Outcome Variable:** Suicidal Ideation (Binary: 0/1).
- **Predictors (Upbringing):** Cumulative ACE Score, Parental Warmth, and Parental Monitoring.
- **Moderators (Context):** Median Neighborhood Income, Urbanicity (Urban vs. Rural), and Race.

2.3 Analytical Strategy

The study employed a three-phase approach:

1. **Exploratory Data Analysis (EDA):** Visual inspection of interaction effects between upbringing and context variables.

2. **Stratified Hypothesis Testing:** Verification of visual patterns using Independent Samples T-tests and Chi-Square tests to determine statistical significance (alpha = 0.05).
3. **Machine Learning:** Construction of a Logistic Regression model (with balanced class weights to address dataset imbalance) to quantify the adjusted odds ratios of surviving predictors.

3. Results

3.1 Rejection of Contextual Hypotheses

Despite visual trends suggesting demographic interactions, stratified hypothesis testing revealed these patterns to be statistically non-significant.

- **Urbanicity:** Parental monitoring levels did not differ significantly between suicidal and non-suicidal teens in either urban ($p=0.86$) or rural ($p=0.13$) settings.
- **Race:** Trauma (ACEs) proved to be a significant predictor across all racial groups, with no statistical evidence supporting a "buffering" effect for Black youth ($p < .05$ for all groups).
- **Maternal Education:** The distribution of maternal education did not differ significantly by suicide risk within the high-income bracket ($p=0.68$), rejecting the status inconsistency hypothesis.

3.2 Determinants of Risk: The Multivariate Model

The final Logistic Regression model excluded the non-significant demographic moderators. The model achieved a Recall of **57%** and an ROC-AUC of **0.65**. Feature importance analysis yielded the following adjusted Odds Ratios (OR):

- **Parental Warmth (OR ~1.5):** This variable emerged as the strongest predictor. Higher scores on the scale were associated with increased risk, indicating that **high** warmth is the primary driver of ideation.
- **Cumulative Trauma (OR ~1.3):** Each additional unit on the standardized ACE score increased the odds of suicidal ideation by approximately 30%.
- **Median Income (OR ~1.1):** In a significant reversal from univariate testing (where income appeared neutral, $p=.059$), the multivariate model identified neighborhood wealth as a risk factor.

4. Discussion

4.1 The "Suppressor Effect" of Wealth

A primary finding of this study is the emergence of neighborhood income as a risk factor only after controlling for trauma. This suggests a classic statistical **suppressor effect**. Wealthy

adolescents in the sample likely experience lower levels of household trauma, which masks the inherent risks of their environment in simple bivariate analyses. However, when holding trauma constant, the "pure" effect of the high-income environment is deleterious.

4.2 The Universality of Trauma and Attachment

The rejection of the stratified hypotheses supports a **Universalist** interpretation of adolescent risk in this sample. The impact of a dysfunctional home environment—characterized by high trauma and low emotional warmth—appears to transcend the boundaries of race and geography. The failure of the "Rural Monitoring" and "Black Resilience" hypotheses suggests that while cultural contexts differ, the fundamental psychological mechanisms linking abuse and detachment to suicidality remain robust.

5. Conclusion

This study concludes that demographic profiling (e.g., targeting rural youth or specific racial groups differently) may be less effective than screening for family dynamic factors. The risk profile identified is a "Triad of Risk" consisting of **High ACEs**, **Low Parental Warmth (if we were to explain that this is reactive parenting)**, and **High Neighborhood Income**. Prevention efforts should prioritize the identification of cumulative trauma and family emotional disconnect, while maintaining vigilance in high-socioeconomic status communities where environmental risks may be masked by material privilege.

6. Limitations and Future Directions

While this study provides significant insights into the determinants of adolescent suicidality, several limitations must be acknowledged. First, the sheer volume and granularity of the microdata present inherent challenges to validity; despite the rigorous statistical controls applied, the complexity of the dataset means that unobserved confounding variables may still influence the findings.

Second, the scope of this analysis was intentionally constrained to demonstrate fundamental data science workflows. Consequently, while the dataset offered a rich array of potential upbringing and contextual variables, the analysis prioritized a streamlined Logistic Regression approach over more complex, high-dimensional modeling techniques. This decision, while necessary to maintain project feasibility, potentially obscures deeper, non-linear interactions between environmental context and family dynamics.

Future iterations of this work would benefit from expanding the feature set to include a broader spectrum of socioeconomic indicators. Additionally, as methodological capabilities advance, the application of non-linear machine learning algorithms—such as Random Forests or Gradient Boosting—could offer a more nuanced understanding of how specific risk factors interact within the "Triad of Risk" identified in this study.