EARLY WARNING PROJECT:
NOT A RELIABLE WARNING FOR INDIA?

A PRELIMINARY REPORT CRITICALLY EVALUATING
THE EARLY WARNING PROJECT MODEL

Acknowledgement

The authors would like to express their sincere gratitude to **Mr. Rajiv Malhotra**, Founder-Director of Infinity Foundation, for his substantial contribution in critically analyzing the EWP's methodology. His insights and perspectives added great value to the report, enriching its depth and rigor

Contributing Authors

Venkataraman Krishnamurthy

Visiting Faculty, Marketing, Operations Management, Quantitative Methods and Information Systems, IIM Udaipur

Dina Banerjee

Faculty, Organizational Behavior and Human Resource Management, IIM Udaipur

Ujjwal Das

Faculty, Operations Management, Quantitative Methods and Information Systems, IIM Udaipur

Shaurya Gupta

Centre for Development Policy & Management, IIM Udaipur

Lavinya Celly

Student, Sciences Po, Paris, France

Centre for Development Policy & Management

An IIM Udaipur Initiative

Importance of This Report

The United States Holocaust Memorial Museum (USHMM) in Washington, DC, issued a report in February 2024 titled *"Risk of Mass Atrocities in India"*, which claimed that India faced a high probability of genocide. This report was subsequently cited by well-known Hinduphobic groups in the United States, raising concerns about its potential impact on India's global perception and international relations. Furthermore, the report called for strong actions against India by the United States and multilateral institutions, including possible sanctions, thereby amplifying its geopolitical significance. Given the high credibility of the USHMM and the potential policy consequences of such a classification, it became imperative to investigate the methodological soundness and academic rigor of the findings presented in the report.

Infinity Foundation, a non-profit organization based in Princeton (USA) that is focused on research and education and specializes in the field of civilization studies, brought the USHMM's report to the attention of its Jewish affiliate, the Hindu Jewish Coalition (HJC) in the United States, where Mr. Rajiv Malhotra, Founder-Director of Infinity Foundation serves as a founding member. Recognizing the gravity of the issue, HJC conducted an investigation in collaboration with the American Jewish Committee (AJC), a prominent Jewish organization in the United States. The AJC, upon review, responded that this was a matter of academic research and clarified that it had not sponsored the report. It was left to Infinity Foundation to analyze the basis of the USHMM's report. This led to a collaborative effort between Mr Rajiv Malhotra and Centre for Development Policy and Management (CDPM), IIM Udaipur. to critically assess the report's claims.

The USHMM's report cited the Early Warning Project (EWP), which had ranked India among the top 15 countries at risk of mass killings since its 2017–18 assessment, with its highest rank being second in the world in the previous year. The EWP, a joint initiative of the Simon-Skjodt Center for the Prevention of Genocide at the USHMM and the Dickey Center for International Understanding at Dartmouth College, has been publishing annual reports titled *"Countries at Risk for Mass Killings" since 2014*. These reports have consistently placed India in a high-risk category, and their ranking was based on a statistical model developed by the EWP. Given the potential policy implications of such assessments, it was essential to examine the theoretical foundations, methodological framework, and empirical robustness of the EWP's predictive model.

The CDPM, IIM Udaipur, undertook a thorough examination of the EWP's methodology, encompassing a literature review on genocide studies, an analysis of the theoretical framework guiding mass killing predictions, an assessment of the selection of risk factors, and a technical evaluation of the statistical model and data sources. The findings of this critical review revealed serious statistical inconsistencies, methodological flaws, in EWP's methodology. These weaknesses significantly undermine the reliability and predictive utility of the EWP model, raising questions about its scientific credibility and policy relevance. Furthermore, this report highlights EWP model's opaque criteria, selection bias, and lack of transparency in data collection resulted in systematic distortions that disproportionately misrepresented India's risk level in comparison to other nations.

Given the stature of USHMM as a globally respected institution and the potential consequences of its reports influencing international policymaking circles, it is critical to develop serious academic rejoinders that challenge biased narratives and ensure intellectual integrity in global discourse. Reports like these, when unchallenged, can shape perceptions in intergovernmental relations, policy formulation in the United States and Europe, and academic discourse on human rights. They may also impact multilateral agencies such as the United Nations (UN), the World Economic Forum (WEF), and corporate entities engaged in Diversity, Equity, and Inclusion (DEI) and Environmental, Social, and Governance (ESG) compliance. Given that many global institutions rely on these assessments when formulating policies and strategic responses, it becomes imperative to re-educate policymakers, academics, and thought leaders by presenting empirical, evidence-based rebuttals to ensure fair and unbiased assessments of India. This report, and similar studies, serve an important academic and strategic function by countering misinformation, ensuring balanced global narratives, and upholding rigorous research standards in assessing the risk of mass atrocities worldwide.

Abstract

This report critically evaluates the Early Warning Project (EWP) Model, a prominent quantitative framework designed to predict and prevent genocide. As a predictive tool widely referenced in genocide prevention efforts, the EWP Model's methodology, transparency, and adherence to academic standards are thoroughly analyzed. The report examines key components, including variable selection, statistical underpinnings, and the model's evolution, to assess its reliability and practical utility. Complementing the model analysis, an extensive literature review contextualizes the EWP Model within the broader field of genocide research. It highlights significant methodological flaws, including reliance on correlational rather than causal data, use of static and redundant variables, and selective adaptation of foundational studies without adequate justification. The report critiques the lack of transparency in variable selection, methodological updates, and scholarly attribution, which compromises the model's credibility. Key issues identified include logical inconsistencies, such as equating extremes (e.g., full civil liberties with complete repression), reliance on outdated or inconsistent data, and counterintuitive conclusions, such as associating multi-party democratic system, civil liberties with increased risk of mass killings. Comparisons with foundational studies reveal substantial methodological divergence. The report calls for greater transparency, theoretical rigor, and adherence to best practices to improve the model's utility as a predictive tool for mass atrocity prevention.

Executive Summary

The Early Warning Project (EWP) Model, developed by the United States Holocaust Memorial Museum and Dartmouth College, is designed to predict and prevent mass killings through a statistical framework assessing socio-political and economic risk factors. The methodology employed by EWP significantly departs from established academic and research standards. While the model claims to offer actionable insights, it exhibits significant limitations that undermine its credibility and practical utility. Chief among these is the reliance on correlational data rather than causal relationships, limiting its ability to provide meaningful guidance for intervention.

The lack of transparency regarding variable selection further exacerbates this issue. Over-reliance on static factors such as "Population Size" and "Geographic Region," traps certain countries like India in high-risk categories regardless of reforms, introducing statistical challenges like multicollinearity that reduce the model's predictive integrity. EWP states that they will not make any effort to explain the relationship between the risk assessment factors used in their model and the occurrence of atrocities: "We make no effort to explain these kinds of relationships in the data, some of which can seem perplexing." This approach is both unusual and problematic, as explaining the rationale for selecting independent variables and its impact on the outcome is a fundamental requirement that any statistical model must fulfil. EWP's position not to address these relationships raises concerns about their methodology and greatly undermines the model's credibility.

EWP has constructed its model using not only the same data across multiple variables but also by including numerous highly correlated variables, both of which raise significant methodological concerns. Using redundant or highly correlated variables in a statistical model introduces several negative effects that undermine its reliability due to multicollinearity, overestimation of effects, overfitting, interpretability, and biased coefficients.

A fundamental flaw in the EWP model lies in its selective adaptation of variables from foundational studies, divorcing them from their original causal context. Foundational works, such as those by Barbara Harff, identified factors like regime type and ideological orientation as key drivers of mass killings but EWP has dropped these variables from its model. Similarly, EWP has excluded several key drivers identified in other foundational studies. EWP Model also cherry-picks and modify variable definitions out of context with respect to the foundational studies while reclassifying causal factors as mere correlates. This deliberate reframing distorts the relationships between variables and outcomes, diminishing the explanatory power of the original research.

The Early Warning Project (EWP) faces critical challenges in maintaining statistical rigor due to its frequent and undocumented changes to model parameters. Parameters such as "Regime Type", "Regime Duration" and "Freedom of Movement of Men" have been removed in recent years, while new variables like women's participation in civil society and media censorship were introduced without clear theoretical justification as they correlate strongly with already existing variables to freedom and civil liberties. Additionally, the calculation of the "Number of Years in Dataset" parameter is flawed, as it inherently relies on temporal comparisons that the EWP itself cautions against, creating a logical inconsistency.

EWP model has significant logical inconsistencies due to statistical and methodological flaws including spurious correlations, modified variable definitions, multicollinearity and improper weighting. A particularly concerning feature of the EWP Model is its deeply anti-democratic orientation, concealed through ambiguous parameter definitions. Criteria such as "Equal Distribution of Civil Liberties across Geographies and Social Groups" fail to differentiate between regimes with no civil liberties and those with full civil liberties. This type of obfuscation places authoritarian states that uniformly suppress freedoms with countries respecting complete civil liberties. Moreover, the model says repressing civil society, banning of all political parties except state-sponsored party and removal of civil liberties across all the territories within a country will lead to drastic reduction in the risk of mass killings.

Table 1: Illustration of EWP Model's Logical Inconsistencies

EWP Model impact due to implementation of anti-democratic and anti-freedom measures	**EWP Model impact due to implementation of democratic reforms and enhancement of freedoms**
India's risk of mass killings reduce from 8% to 3% and it's rank improves exponentially from 5 to 32 if it takes following anti-democratic measures • Become fully autocratic and entirely suppress civil liberties across all its territories • Curtail women's rights in such a way that they are often or always prevented from participating in civil society organizations • Ban all parties except state-sponsored party. • Curtail judiciary's ability to control arbitrary power • Repress civil society organizations • Do not respect freedom of Religion • Do not respect Freedom of Discussion	India's risk of mass killings and it's rank will remain the same even it it implement slew of following democratic reforms: • Enhance the judiciary's ability to control arbitrary power via institutional reform. • Fully respect freedom of discussion • Fully respect freedom of religion • No government censorship • No discrimination against any section of the population
China's risk of mass killings reduce from 3% to 1% and it's rank improves exponentially from 26 to 62 if it takes following anti-democratic measures: • Become fully autocratic and entirely suppress civil liberties across all its territories • Curtail judiciary's ability to control arbitrary power • Repress civil society organizations • Prevent women from participation in civil society organizations	North Korea's risk of mass killings increase from 1% to 3% and it's rank deteriorates from 76 to 21 if it implement slew of following democratic reforms: • Enhance the judiciary's ability to control arbitrary power via institutional reform. • No ban on political parties • Allow civil liberties in some geographical regions • Stop repression of civil society • Freely allow women's participation in civil society organizations

The model's data inputs also raise serious concerns. It relies heavily on subjective ratings, such as those from the V-Dem project, which lack transparency and supporting evidence. EWP demonstrates selectivity in data sources. For instance, for assessing ethnic fractionalization, EWP relies on census data from countries such as the United States, France, Israel, and New Zealand while it relies on outdated data from the 2000 Encyclopedia Britannica for India, despite the availability of the comprehensive and updated 2011 Census. The use of inconsistent data sources across regions reflects a discriminatory practice that affects the objectivity and credibility of the EWP. Discrepancies between the model's assumptions and public reporting further undermine its reliability.

The Early Warning Project (EWP) model faces critical reliability issues due to its high False Positive Rate (FPR) of 95.2% and False Negative Rate (FNR) of 29-36%. These errors significantly undermine its effectiveness as a predictive tool for mass killings. The high FPR means the model generates 20 false alarms for every actual event, driven by reliance on correlational rather than causal relationships, redundant variables causing multicollinearity, and lack of transparency in variable selection and weighting. This results in misallocation of resources and "alert fatigue."

Conversely, the high FNR indicates the model misses up to 36% of actual mass killing onsets, stemming from its inability to establish causal links between risk factors and events. These issues compromise the model's ability to provide accurate warnings and timely interventions.

Overall, the combination of excessive false alarms and missed predictions diminishes the EWP model's credibility and practical utility for policymakers and researchers, limiting its role in atrocity prevention. The lack of transparency and reliance on static parameters further hinders necessary improvements to address these shortcomings.

The model displays drastic fluctuations in intercept values—from a stable range of -8.8 to -11.2 between 2019 and 2022 to an unexpected surge to 67.96 in 2023—and inconsistent coefficients for key variables, both in magnitude and direction. These variations suggest a lack of theoretical grounding, arbitrary adjustments to variable inclusion, and failure to address spurious correlations, leading to unstable relationships between predictors and outcomes. Such instability undermines the model's predictive power, interpretability, and credibility, raising concerns about its reliability and usability for decision-making.

The EWP's final reports contradict its statistical findings, raising questions about the coherence and transparency of its conclusions. The Early Warning Project's (EWP) reports exhibit a substantial reliance on media sources and opinion pieces rather than on its Statistical Risk Assessment Model, raising concerns about the integrity and rigor of its conclusions.

In summary, the EWP model's reliance on correlational data, opaque revisions, logical flaws, ethically questionable assumptions, high FPR and FNR severely limit its effectiveness. To enhance its reliability and policy relevance, the model must:

- Transition from correlation-based predictions to causal frameworks.
- Increase transparency by documenting variable selection criteria and methodological changes.
- Refine variable definitions to capture nuanced socio-political realities.
- Align more closely with foundational research to preserve conceptual integrity and/or provide robust theoretical framework for any changes.
- Reduce reliance on static parameters that skew rankings without offering actionable insights.

Addressing these issues is critical for restoring the model's credibility and ensuring it serves as a valuable tool for preventing mass atrocities. Failure to address these flaws risks misguiding policymakers and undermining their stated mission.

Contents

1. Introduction 10
2. Comprehensive List of Genocide Predictors Based on the Literature Review 11
3. Early Warning Project (EWP): An introduction 13
4. Early Warning Project (EWP): Misalignment between Claims and Model Capabilities of Early Warning Project 13
5. Challenges in Using the EWP Model for Effective Intervention 14
6. Limitations of Correlation-based Prediction Model 14
7. Limitations of Correlation Models for Predictions: A Literature Review 15
8. Criteria for Selecting EWP Model Variables Not Specified 16
9. Lack of Transparency in Methodological Revisions and Scholarly Attribution 16
10. Redundant and Correlated Variables Compromise the Predictive Integrity of the EWP Model 21
11. Selective Variable Adaptation and Methodological Distortions 23
12. Brief Summary of the Foundational Studies Referenced by EWP 24
13. Methodological Divergence in the Early Warning Project 33
14. Low Reliability due to High False Positive Rate (FPR) and High False Negative Rate (FNR) 34
15. Significant Logical Inconsistencies in the EWP model 36
16. EWP's Changing Parameters: A Threat to Statistical Rigor and Credibility 37
17. EWP Model Equates Complete Lack of Civil Liberties with Full-civil Liberties 38
18. EWP Model Adopts the Erroneous Principles of "All Equal None" or "Nothing Equals Something" 40
19. EWP model greatly incentivizes authoritarianism and various forms of repression of liberties 46
20. EWP model permanently traps countries like India as high risk for mass killings, irrespective of major reforms due to high weightage for static factors like "Population Size" and "Region" 47
21. Correlated Variables Contradict: Structural Issues in the EWP Model 48
22. Errors/Logical Inconsistencies / Lack of Evidences in the EWP Data Points 49
23. No Data, Just Anonymous Opinions for Many Parameters 51
24. Factual Errors and Logical Inconsistencies in the EWP Model 55
25. Changes in the Intercept and Coefficients 56
26. Discrepancies in EWP's Model Assumptions vs. Public Reporting 57
27. Final Report Different from the Model's Findings 59
28. Conclusion 61
29. References 63
30. Index of Tables 67

Introduction

Predicting and preventing genocide remains a critical concern for governments, international organizations, and researchers, as such atrocities inflict profound and lasting harm on societies. *While the extensive literature on genocide predictors reveals a complex interplay of social, political, and economic factors, forecasting systems like the Early Warning Project (EWP) Model represent a quantitative approach to assessing the risk of such events.* This report critically examines the EWP Model, delving deeply into its methodology, transparency, and alignment with academic standards.

The EWP Model, widely referenced in the field of genocide prevention, is scrutinized in this report to evaluate its reliability and utility as a predictive tool. Key aspects such as its selection of variables, its underlying statistical framework, and its modifications over time are analyzed in detail.

In addition to its primary focus on the EWP Model, this report incorporates insights from an extensive literature review. By examining a wide range of academic articles and studies, the report contextualizes the EWP Model within the broader discourse on genocide predictors. The report highlights the methodological differences between the EWP Model and other studies, emphasizing the distinct approaches employed by the EWP compared to the qualitative and socio-political frameworks that dominate existing genocide research.

Through this dual lens of in-depth model analysis and a comprehensive literature review, the report aims to contribute to the advancement of early warning systems. It identifies gaps in current predictive methodologies and offers recommendations to reconcile quantitative and qualitative approaches, ultimately enhancing the effectiveness of genocide prevention efforts.

Comprehensive List of Genocide Predictors Based on the Literature Review

Genocide is a complex and multifaceted phenomenon often driven by a wide range of *political, social, and economic factors*. The literature on genocide predictors emphasizes the importance of understanding these variables to effectively anticipate and prevent such atrocities. The genocide predictors identified through academic literature provide a broader and more nuanced perspective on the socio-political dynamics that lead to mass violence. These indicators often focus on deeper systemic issues, such as *political instability, ethnic and religious divisions, and exclusionary ideologies*, which contribute to the conditions under which genocides are likely to occur.

For example, *regime changes, civil wars, and political upheavals* create unstable environments that can trigger violence, while historical grievances and long-standing discrimination can escalate into genocidal actions when tensions remain unresolved. Additionally, economic inequality, crises such as famine or disease outbreaks, and the involvement of non-state armed groups further exacerbate the risk of genocide.

Many of these predictors highlight the role of state-led actions, where authoritarian regimes and manipulative legal frameworks are often used to justify repressive policies aimed at specific groups. The presence of dehumanizing ideologies and rhetoric also plays a critical role in fueling genocidal intent, allowing perpetrators to morally disengage and justify their actions. Similarly, struggles among political elites for control of state power can lead to heightened tensions and, ultimately, mass violence. The involvement of foreign powers and ineffective international responses can further inflame internal conflicts, making external intervention either ineffective or a contributing factor to genocide.

Overall, the predictors from the literature provide a comprehensive understanding of the complex interplay of factors that drive genocides. These indicators are summarized in the following table, providing a clear outline of the main predictors identified through the literature and their explanations.

Table 2: List of Genocide Predictors from Literature Review and Their Explanations		
SL	**Literature-based Genocide Predictors**	**Explanation**
1.	Political Instability and Upheaval	Regime changes, civil wars, and political crises can trigger genocide. Includes state perception of threats and political regime type.
2.	Ethnic and Religious Divisions	Deep-rooted divisions in society, along with ethnic polarization and proximity, increase risks.
3.	Exclusionary and Extremist Ideologies	Extremist and exclusionary ideologies, whether political or religious, drive mass violence.
4.	Historical Grievances and Discrimination	Countries with past genocides and long-standing grievances are more likely to experience genocide. Includes state-led discrimination.
5.	Economic Conditions and Inequality	Poor economic conditions, economic shocks, and significant inequality between groups are precursors to violence.
6.	Ongoing Conflicts and Armed Groups	Active conflicts and the involvement of armed non-state actors collaborating with the state lead to genocide.
7.	Dehumanization and Ideological Beliefs	Dehumanizing rhetoric, genocidal ideologies, and moral disengagement are significant contributors to genocidal behavior.
8.	Crisis Situations and Social Unrest	Extreme crises such as famine, disease outbreaks, or civil unrest can push states to genocidal actions.
9.	Legal Frameworks and Governance	Laws and legal structures can be manipulated to justify genocidal actions. Authoritarian regimes are more prone to abuse of power.
10.	Cultural and Biological Genocide Predictors	Destruction of cultural heritage, targeted attacks on specific populations, and denial of resources or health care are indicators of genocidal intent.
11.	Elite Competition and Power Struggles	Struggles for state control, especially among ethnic or political elites, lead to heightened tensions and genocide.
12.	Foreign Influence and International Response	External political and economic pressures, combined with ineffective peacekeeping, can exacerbate internal tensions.
13.	Military Strategies and State Actions	Use of military force, including scorched-earth tactics and forced displacement, indicate state policy aimed at eliminating specific groups.
14.	Psychological Mindset and Dehumanization	The psychological mindset of perpetrators, including moral disengagement, allows for justification of genocidal actions.

Early Warning Project (EWP): An introduction

The Early Warning Project (EWP) Model, developed by the United States Holocaust Memorial Museum and Dartmouth College, is designed to predict and prevent mass killings through a statistical framework assessing over 30 socio-political and economic risk factors. The methodology employed by EWP significantly departs from established academic and research standards. While the model claims to offer actionable insights, it exhibits significant limitations that undermine its credibility and practical utility.

Early Warning Project (EWP): Misalignment between Claims and Model Capabilities of Early Warning Project

There is a notable disconnect between the EWP's claims, objectives, and what its model actually delivers. While EWP aims to provide actionable insights to prevent mass atrocities, its model remains purely correlational, lacking causal explanations for the risk factors used. EWP positions itself as a provider of reliable intelligence on potential mass atrocities, claiming to identify warning signs that can help prevent such events. It says on its website that genocide and mass atrocities are never spontaneous and can be predicted by studying patterns from past situations. and claim that they have identified patterns and circumstances that often precede mass killings. About its statistical model, EWP says that, ***"in partnership with Dartmouth College, we study past situations where governments and nonstate groups systematically targeted and killed thousands of civilians. In doing so, we identified a range of patterns and circumstances that often precede such violence."***

However, the model is purely correlational and does not identify any such precedents or causal factors. EWP ranks ***c***ountries based on certain variables, which are called risk assessment factors. However, they state that these factors are not causal but merely correlational, meaning that the identified predictors of atrocities do not necessarily cause them. EWP says that ***"Readers should keep in mind that our model is not causal: the variables identified as predicting higher or lower risk of mass killings in a country are not necessarily the factors that drive or trigger atrocities".*** "***We emphasize that these risk factors should not be interpreted as causes or "drivers" of risk but simply as correlates of risk that have proven useful in forecasting."*** This approach limits the model's utility for fulfilling EWP's objective of preventing atrocities, as it neither identifies causal factors nor offers actionable insights for intervention, creating a clear gap between their goals and what the model actually delivers.

Challenges in Using the EWP Model for Effective Intervention

EWP concedes that addressing the risk factors identified in their model may not actually reduce the likelihood of mass atrocities. ***"An important consequence of the non-causal nature of these forecasts is that actions aimed at addressing risk factors identified in the model would not necessarily be effective ways of mitigating the risk of mass atrocities; this assessment does not seek to evaluate atrocity prevention policy prescriptions."*** This further complicates the practical value of their model for guiding prevention.

While EWP claims to provide insights to prevent mass atrocities, their model's non-causal nature and failure to explain the relationship between variables and events weaken its reliability. While EWP claims that "***As the first public early warning system for mass atrocities, this project aims to provide governments, civil society groups, and other influential actors early and reliable warnings of mass atrocities and, as a result, greater opportunities to take preventive action",*** there is a clear gap in terms of the deliverables of their model for prediction or even prevention.

Limitations of Correlation-based Prediction Model

Purely correlation-based models, such as those employed by EWP, have significant limitations when used to predict future outcomes. These models do not offer the depth, stability, or reliability needed for understanding the true drivers of complex phenomena, leading to potentially flawed conclusions and strategies. Some of the limitations are as follows:

- **Spurious Relationships:** Correlators can result in spurious correlations, where two variables appear related but are not causally connected. This can lead to false inferences, as the model might overfit to patterns that do not hold in different datasets or contexts.
- **Vulnerability to Confounding Variables:** Correlational models are highly sensitive to confounding variables that may be influencing both the predictors and the outcome. Without considering causal mechanisms, the model can misinterpret relationships due to the presence of these hidden or unaccounted-for variables.

- **Unstable Predictions and Non-generalizability:** Models based on correlators may not perform well outside the dataset on which they were trained. Since correlation does not establish a stable or fundamental connection, small changes in data or context can drastically alter predictions, making such models unreliable in new or evolving environments.
- **Ethical and Policy Implications:** Relying on correlational models in areas such as policymaking, health care, or security could lead to misinformed decisions. Policies designed based on correlational findings might fail or have unintended consequences because they do not address the true drivers of the issue.

Limitations of Correlation Models for Predictions: A Literature Review

Several research papers and books have pointed out the limitations of using purely correlation-based model for predictions, such as:

- **Pearl, J., & Mackenzie, D. (2018).** ***The Book of Why: The New Science of Cause and Effect*****. Basic Books.**
  In this book, Judea Pearl explores the pitfalls of correlation-based approaches in machine learning and statistical modeling. He argues that understanding causality is essential for making reliable predictions, especially when trying to forecast the effects of actions or interventions.
- **Holland, P. W. (1986). Statistics and Causal Inference.** ***Journal of the American Statistical Association*****, 81(396), 945-960.**
  Holland's paper critiques the use of correlational data for making causal inferences and predictions. He highlights that without establishing causal links, predictions based on correlations can lead to incorrect or unstable conclusions.
- **Silver, N. (2012).** ***The Signal and the Noise: Why So Many Predictions Fail—but Some Don't*****. Penguin Books.**
  Nate Silver discusses the limitations of using purely correlational models for prediction, particularly in areas like economics and politics. He emphasizes the importance of understanding the underlying causes of events rather than just identifying statistical correlations.

Criteria for Selecting EWP Model Variables Not Specified

The EWP employs over 30 variables, referred to as correlators, to assess the risk of mass killings but provides no explanation for how these variables were selected, stating that only their selection is *"based on prior empirical work and theory."*

In its annual reports, EWP references foundational studies by Barbara Harff, Political Instability Task Force, and Jay Ulfelder, yet closer examination reveals significant differences between the risk assessment criteria from these sources and the variables currently employed by EWP. A request for clarification on these discrepancies yielded no additional insights, as EWP directed the inquiry back to its website without further explanation.

EWP also states that they will not make any effort to explain the relationship the risk assessment factors used in their model and the occurrence of atrocities: *"We make no effort to explain these kinds of relationships in the data, some of which can seem perplexing."* This approach is both unusual and problematic, as transparency regarding the selection of variables and the theoretical or empirical basis underpinning them is fundamental for ensuring the rigor and credibility of any model. It is already a significant limitation that the EWP model does not identify the causal factors behind mass killings. However, the absence of any explanation for the selection criteria of even the correlational factors further compromises the model's validity.

Lack of Transparency in Methodological Revisions and Scholarly Attribution

The EWP) was initially designed in 2014 with Dr. Jay Ulfelder's assistance. Over the years, the model has undergone substantial revisions, with a major transformation in 2017 that expanded its scope to include mass killings by non-state actors and extended the forecasting window.

Subsequent iterations have been introduced even after 2017 through the introduction of new variables, and removal of others. By 2024, the variables and structure of the EWP model had diverged considerably from the 2014 design.

Table 3: Comparison of Ulfelder Variables (used by EWP in 2014) and EWP Variables (2024)		
SL	**Ulfelder Variables[1] (used by EWP in 2014)**	**EWP Variables (2024)[2]**
1.	Region	Region
2.	Population size	Population Size
3.	Any ongoing episodes of state-led mass killing	Number of Ongoing Mass Killings
4.	Infant mortality rate	Infant Mortality Rate
5.	Annual % change in GDP per capita	Annual Percentage Change in GDP Per Capita
6.	Any coup attempts in past 5 years	Any Coup Attempts in Past five years
7.	Ethnic Fractionalization	Ethnic Fractionalization
8.	Share of population subjected to state-led discrimination	Percentage of Population Discriminated Against
9.	ICCPR 1st Optional Protocol signatory	State Signatory of First Optional Protocol to the International Covenant on Civil and Political Rights
10.	Trade openness	Number of Years since independence
11.	Post-Cold War period (year ≥ 1991)	Any Mass Killing (Since 1945)
12.	Polity (Autocracy/Anocracy/Democracy/Other)	Average Annual Rate of Mass Killing Onsets
13.	Regime Duration	Battle-Related Deaths
14.	Salient elite ethnicity (Majority Rule/Minority Rule)	Mass Killing Onset in 2022
15.	Ruling elites espouse an exclusionary ideology	Civil Society Repression
16.	Sum of max annual magnitudes of PITF instability other than genocide from past 10 yrs	Freedom of Discussion
17.	Armed conflict in geographic region	Freedom of Religion
18.	Violent civil conflict	Government Censorship
19.	Any state-led mass killing since WWII (cumulative)	Inequality in Civil Liberties — Geographic Region
20.	Country age	Inequality in Civil Liberties — Social Group
21.		Political Killing
22.		Prevention of Women's Participation in Civil Society Organizations
23.		Judicial Reform
24.		Minority Control
25.		Party Ban
26.		Power Distributed by Social Group
27.		Power Distributed by Socioeconomic Position
28.		New Country
29.		Ongoing Mass Killing Targets Multiple or Broad Groups
30.		Ongoing Mass Killing(s) Targets Exactly One Protected Group

[1] Ulfelder variables have been taken from the Source: https://github.com/ulfelder/earlywarningproject-statrisk-replication/blob/master/data.out/EWP%20Data%20Dictionary%2020140909.pdf

[2]EWP Variables (2024) have been taken from the source: https://earlywarningproject.ushmm.org/methodology-risk-factors

In 2024, the following eight out of the twenty variables used by Ulfelder as well as EWP (2014) were removed from the list of EWP's variables.

Table 4: Variables used by Ulfelder as well as EWP (2014) were removed from the list of EWP's variables.

SL	Variables
1.	Trade Openness
2.	Post-Cold War period (year ≥ 1991)
3.	Polity (Autocracy/Anocracy/Democracy/Other)
4.	Regime Duration
5.	Ruling elites espouse an exclusionary ideology
6.	Sum of max annual magnitudes of PITF instability other than genocide from past 10 yrs
7.	Armed conflict in geographic region
8.	Violent civil conflict

The following nineteen out of the thirty variables used by EWP in 2024 were either not included in or modified from the variables list of Ulfelder (2014):

Table 5: List of Variables used by EWP in 2024, either not included in or modified from the variables list of Ulfelder (2014)			
SL	**Variables**	**SL**	**Variables**
1.	Freedom of Discussion	**14.**	Ongoing Mass Killing Targets Multiple or Broad Groups*(Ulfelder's dataset considered only state-led mass killings)*
2.	Freedom of Religion	**15.**	Ongoing Mass Killing(s) Targets Exactly One Protected Group *(Ulfelder's dataset considered only state-led mass killings)*
3.	Government Censorship	**16.**	Battle-related Deaths
4.	Civil Society Repression	**17.**	Mass Killing Onset in 2021 *(Ulfelder's dataset considered only state-led mass killings)*
5.	Inequality in Civil Liberties —Geographic Region	**18.**	Prevention of Women's Participation in Civil Society Organizations
6.	Inequality in Civil Liberties — Social Group	**19.**	Judicial Reform
7.	Political Killing		
8.	Party Ban		
9.	Minority Control		
10	Power Distributed by Social Group		
11.	Power Distributed by Socioeconomic Position		
12.	Average Annual Rate of Mass Killing Onsets *(Ulfelder's dataset considered only state-led mass killings)*		
13.	Number of Ongoing Mass Killings (*Ulfelder's dataset considered only state-led mass killings)*		

However, theoretical reasoning and methodological justification for the modification of variables are not documented on EWP's website or available in their reports. When contacted in September 2024, Dr. Ulfelder clarified that he has not been involved with the EWP project for over a decade, directing queries to the United States Holocaust Memorial Museum (USHMM). Furthermore, details regarding the contributions of scholars such as Professors Valentino and Hazlett, who are reportedly credited by EWP for assisting with the 2017 revisions, are notably missing. The academic reference available on their website remains the 2013 paper by Dr. Jay Ulfelder, "A Multimodel Ensemble to Forecast Onsets of State-Sponsored Mass Killing."

Despite claims of methodological updates, no subsequent research paper or article detailing the changes made in 2017 and subsequently has been made available, raising concerns about the transparency of the modifications underlying the current model. Since significant changes were indeed implemented in 2016, it would be logical to reference both the 2013 and 2016 publications, rather than relying solely on the 2013 paper by Dr Ulfelder;s which EWP do not follow currently. It is inexplicable that EWP has published the research article whose model they do not follow rather than the ones that they follow. This lack of clear attribution and comprehensive documentation of scholarly contributions falls short of established academic and research standards.

Redundant and Correlated Variables Compromise the Predictive Integrity of the EWP Model

The EWP has constructed its model using not only the same data across multiple variables but also by including numerous highly correlated variables, both of which raise significant methodological concerns. For example, in the case of India, data on the Maoist insurgency, Sikh insurgency, and Kashmir insurgency are represented under two separate independent variables—"Battle-Related Deaths" and "Past Mass Killing Onsets Since 1945."

The Early Warning Project (EWP) has utilized two independent variables, social_power_dist and minority rule, both derived from the same ordinal variable, v2pepwrsoc_ord, from the V-Dem dataset. Additionally, the variables within the EWP dataset demonstrate significant levels of correlation, which undermines the reliability of its predictive model. For example, variables related to civil liberties and political freedoms—such as Civil Society Repression, Government Censorship, Inequality in Civil Liberties (Geographic Region), Inequality in Civil Liberties (Social Group), Freedom of Discussion, Freedom of Religion, Women's Participation in Civil Society Organizations, and Party Ban—are all inherently interconnected. Repressive actions like government censorship or party bans often coincide with restrictions on freedom of discussion and religion, as well as broader inequalities in civil liberties across regions and social groups. Similarly, women's limited participation in civil society is typically a consequence of these broader repressive dynamics, making these variables reflections of the same underlying trends.

In the domain of social power and discrimination, variables like Minority Control, Power Distributed According to Social Groups, and Percentage of Population Discriminated Against are deeply interrelated. Minority control over political or economic resources often directly correlates with unequal power distribution among social groups and an increase in the percentage of the population facing systemic discrimination. These variables essentially capture overlapping facets of social inequality, making their inclusion as separate parameters redundant.

The variables such as New Country and Number of Years since independence, are by design, highly correlated. A "new country" will inherently have fewer years represented in the dataset, creating a direct relationship between these two variables.

Using redundant or highly correlated variables in a statistical model introduces several negative effects that undermine its reliability and interpretability. One primary issue is **multicollinearity**, which occurs when independent variables are highly correlated and thus supply overlapping information. This redundancy makes it challenging for the model to differentiate the unique impact of each variable, leading to unstable coefficients and inflated standard errors. Such instability impedes the model's ability to accurately identify significant predictors, often resulting in misleading conclusions about variable importance.

Another consequence is **overestimation of effects**. When similar data appears across multiple variables, the model may erroneously inflate the influence of certain factors, misinterpreting their importance. This can lead to a distorted view of which variables are most significant in determining the outcome, potentially skewing interpretations and reducing the model's accuracy.

Additionally, **overfitting** becomes a concern with redundant variables, as the model may fit too closely to the training data, capturing random noise rather than meaningful patterns. This over-complexity reduces the model's performance on new data, limiting its utility in real-world applications where adaptability is essential.

Including highly correlated variables also impacts **interpretability**. When predictors overlap, it becomes difficult to separate their individual contributions, which obscures understanding of how each variable relates to the outcome. This ambiguity weakens the model's insights and clarity. Finally, redundant variables can lead to **biased coefficients**, as the model struggles to assign appropriate weights to overlapping predictors. Such bias distorts estimates of each variable's true impact, further compromising the model's reliability and usefulness in decision-making.

In summary, using the same data across different independent variables or highly correlational variables can compromise the model's accuracy, reliability, and interpretability. To avoid these pitfalls, it is essential to reduce redundancy by selecting distinct, non-overlapping predictors and addressing multicollinearity through proper statistical techniques.

Selective Variable Adaptation and Methodological Distortions

The EWP has not articulated the criteria underpinning its selection of independent variables, raising concerns about the model's methodological transparency. An examination of the studies EWP references as foundational sources suggests that EWP has selectively incorporated assessment variables from these works, while omitting several critical variables identified in the original research. Moreover, EWP appears to have significantly modified some variables cited from these sources, while a number of other assessment factors cannot be directly traced back to the studies EWP references.

According to EWP, its initial model draws on the research of Barbara Harff and the Political Instability Task Force (PITF), which Dr. Jay Ulfelder assisted in its implementation. Ulfelder's paper relied on PITF/Harff Model, Colaresi and Carey Model and Elite Threat Model.

Brief Summary of the Foundational Studies Referenced by EWP

----

1. **Genocide and Political Mass Murder Since 1955 by Barbara Harff**

This paper examines the factors that elevate the risk of genocide and political mass murder, aiming to identify patterns that may help prevent future atrocities. Using historical data and a statistical model, Harff contrasts countries that have experienced mass killings with those that have not. Her findings emphasize the challenges of preventing such violence but stress the importance of early identification of high-risk situations. Six key predictors of genocide or political mass murder are outlined here.

1.1. *Political Upheaval:* Internal political crises, such as revolutionary wars, state collapse, or significant instability, raise the risk of mass killings. The greater the intensity of internal wars and regime crises over the previous 15 years, the more likely a new state failure could lead to genocide or politicide. In cases of severe upheaval, the risk of genocide or politicide is nearly twice as high, especially when these crises involve armed conflicts within the state.

1.2. *Prior Genocides:* A history of past genocide or mass killings in a country strongly correlates with future occurrences. Countries with previous genocides are more than three times as likely to experience new episodes of genocide or politicide following state failure. Harff's analysis suggests that past occurrences contribute more significantly to the risk of future genocide than even the intensity of recent political upheaval.

1.3. *Ethnic and Religious Cleavages:* The dominance of a political elite based on an ethnic minority often leads to exclusionary policies and heightens the risk of mass killings. Harff's model shows that countries with a political elite rooted in an ethnic minority face a risk of genocide or politicide that is two and a half times higher than in states without such minority-based dominance.

1.4. *& 1.5 Elite Ideology and Regime Type:* Authoritarian or exclusionary ideologies among ruling elites significantly increase the likelihood of mass killings. Countries governed by elites with exclusionary ideologies are two and a half times more likely to experience state failures that lead to genocide or politicide. Additionally, autocratic regimes are three and a half times more likely to encounter mass killings than democratic ones, particularly when other risk factors are present.

1.6. *International Interdependence (Trade Openness):* Lower trade openness correlates with a higher risk of genocide, as countries with closed economies lack external economic engagement and pressure that might otherwise discourage mass killings. Countries with low trade openness are two and a half times more likely to experience genocide or politicide. In contrast, high trade openness minimizes the risk of state failure and substantially reduces the likelihood that state failures would lead to severe human rights violations.

2. A Global Model for Forecasting Political Instability by Jack A. Goldstone et al.

This study introduces a global model to forecast political instability with a two-year lead time, analyzing instability events from 1955 to 2003. It aims to provide a simple yet accurate framework that focuses on political structures and elite relationships rather than traditional economic or geographic factors. The model identifies four key predictors of instability and demonstrates over 80% accuracy in forecasting various forms of political instability, including civil wars and adverse regime changes.

Key Predictors of Political Instability

2.1 ***Regime Type:*** A newly developed regime typology identifies partial democracies with factionalism as the most unstable. These regimes exhibit over 30 times the likelihood of instability compared to full autocracies. Partial democracies without factionalism and partial autocracies also show elevated risks, while full democracies and autocracies are relatively more stable.

2.2 **Infant Mortality:** High infant mortality rates serve as a significant risk factor. Countries at the 75th percentile of infant mortality rates are 6.5 times more likely to experience instability compared to those at the 25th percentile.

2.3 ***Conflict-ridden Neighborhoods:*** The presence of armed conflicts in four or more neighboring states substantially increases the likelihood of instability.

2.4 ***State-led Discrimination:*** High levels of political or economic discrimination against minority groups increase the risk of instability, particularly civil wars. Countries exhibiting state-led discrimination face three times the odds of instability onset compared to those without such practices.

3. To Kill or to Protect Security Forces, Domestic Institutions, and Genocide by Colaresi and Carey

In "To Kill or to Protect: Security Forces, Domestic Institutions, and Genocide," Michael Colaresi and Sabine C. Carey investigate how the structure and nature of domestic security forces, along with political institutions, influence the likelihood of genocide. They analyze data to understand when security forces might enable mass killings or, alternatively, protect populations from such violence. Their study underscores the role of state-controlled security forces and governance structures in either fueling or preventing genocide. Colaresi and Carey's findings suggest that reducing the likelihood of genocide involves strengthening democratic institutions, improving civilian oversight, and limiting the autonomy of security forces. This model highlights both the preventive and risk-enhancing roles of domestic institutions and security forces in shaping the incidence of mass violence. The authors identify several parameters that influence the probability of mass killings:

3.1 Security Force Autonomy: When security forces operate with high autonomy, they are more likely to act against the population without state oversight, which can lead to increased risks of genocide. Autonomous security forces often escape accountability and may act on internal biases.

3.2 Civilian Oversight of Security Forces: Strong civilian control and oversight of military and police forces correlate with a reduced risk of genocide. Such control helps to restrain security forces and ensures they operate under legal and institutional checks that protect citizens.

3.3 Political Institutions: The type of political regime plays a crucial role in the probability of mass killings. Democracies, with transparent institutions and accountability mechanisms, are less likely to experience genocide. In contrast, authoritarian regimes, where leaders have fewer constraints, are at a higher risk.

3.4 State Capacity and Cohesion: A state with high administrative and institutional capacity, which allows for cohesive control over security forces and adherence to laws, is better able to prevent mass killings. Weak state structures often lead to fragmented control, increasing the likelihood of violence.

3.5 Presence of Conflict or Rebellion: Active conflict or rebellion within a state elevates the risk of genocide, as governments may perceive certain groups as threats and use security forces to target them under the pretext of maintaining order.

4. A Multi-model Ensemble to Forecast Onsets of State-Sponsored Mass Killing by Jay Ulfelder

In "A Multimodel Ensemble to Forecast Onsets of State-Sponsored Mass Killing," Jay Ulfelder combines four different forecasting models to identify the onset of mass killings, each based on distinct methodologies and sets of variables. The following is an analysis of each model and the parameters used to predict mass killing.

4.1 ***Forecast 1- Colaresi and Carey Model:*** This model, developed by Michael Colaresi and Sabine Carey, focuses on political instability and the likelihood of genocide or politicide occurring within episodes of political instability. The model's core premise is that the risk of state-led mass killing increases based on the interaction between security forces' size and executive constraints.

4.2 ***Forecast 2- PITF/Harff Model:*** This model combines two separate forecasts: the Political Instability Task Force (PITF) model for political instability and Barbara Harff's model for mass killings conditional on political instability. By integrating these, Ulfelder aims to assess mass killing risks more holistically.

4.3 ***Forecast 3- Elite Threat Model*:** This model assumes that mass killings are more likely during elite crises, such as civil wars, coups, and regime collapses. It uses logistic regression to assess mass-killing onset based on indicators for civil war and coup propensity.

Parameters:

- ***Civil War:*** The risk of mass killings rises significantly during ongoing civil conflicts.
- ***Infant Mortality:*** High mortality rates indicate instability and poor governance.
- ***Political Regime Type:*** Autocratic or unstable regimes correlate with a higher risk.
- ***Natural Resource Income:*** Resource-dependent economies that can fuel internal strife, are at greater risk.
- ***GDP Growth:*** Annual GDP growth below 2% indicates economic instability.
- ***Regional Conflict:*** Conflict in neighboring countries elevates the risk.

This model assesses mass killings as more probable in contexts where civil wars and coups are likely.

4.4 ***Forecast 4- Random Forests Model:*** The Random Forests model employs machine learning to predict mass killings by creating multiple decision trees based on subsets of the covariates used in the other models. It aggregates these trees for a more comprehensive prediction.

Parameters: This model uses all the parameters from the other three models, encompassing variables related to;

- **Executive constraints**
- **Ethnic diversity**
- **Infant mortality**
- **Security force size**
- **Economic factors**
- **Regional conflict**
- **Trade openness**
- **Regime type**

The Random Forests model offers a broader, data-driven approach by identifying patterns in the cumulative data from the other models.

Divergence of the EWP Model from Jay Ulfelder's Model

Ulfelder's multimodel ensemble leverages these four models to capture a diverse range of factors influencing the likelihood of state-sponsored mass killings.

1. **Executive authority and security force dynamics** (Colaresi and Carey)
2. **Political instability and conditional probability of violence** (PITF/Harff)
3. **Crisis-driven mass killing risks** (Elite Threat)
4. **Comprehensive data-driven analysis** (Random Forests)

The foundational studies referenced by the EWP employed causal-based models, explicitly establishing clear relationships between risk factors and the occurrence of atrocities or political instability. However, EWP, while incorporating parameters derived from these studies, characterizes these factors as correlational rather than causal. A plausible explanation for this discrepancy lies in EWP selectively incorporating parameters from various foundational studies introducing numerous additional variables, without providing a clear theoretical justification or methodological rationale for their inclusion. This selective integration likely compromises the causal validity of the original parameters by altering their contextual relevance and theoretical grounding. Consequently, the causal relationships identified in the foundational studies may be diluted or obscured, prompting EWP to redefine these parameters as correlational. This methodological shift represents a significant departure from the causal frameworks of the foundational studies and highlights challenges in preserving the conceptual integrity of parameters when they are combined within a broader, less focused analytical framework.

"Regime Type" has been consistently recognized as a significant predictor of mass killings in foundational studies, notably those by Barbara Harff and the Political Instability Task Force (PITF). Until 2019-20, EWP included "Regime Type" as a critical parameter in its risk assessment model. EWP reports from 2018-19 and 2019-20 explicitly cite the role of anocratic regimes in elevating the mass killing risk rankings of countries such as Egypt, Angola, Côte d'Ivoire, Yemen, Turkey, Sudan, Ethiopia, the Republic of Congo, and Cameroon. However, this parameter was excluded from the EWP model beginning in 2020-21 due to the discontinuation of updated data from the Center for Systemic Peace, the source of this information. Although the EWP argued that removing this variable did not significantly affect the model's overall accuracy, it acknowledged potential shifts in the risk rankings of specific countries. Given its prior importance in explaining high-risk rankings and its foundational relevance, the decision to discard this parameter is surprising. The exclusion of "Regime Type", a variable validated as a key predictor in foundational studies, raises serious concerns about the methodological rigor of the EWP model. By discarding a parameter of such importance based on the lack of updated data, the model appears to adopt a casual approach to parameter selection. It has significant impact on the rankings of several countries. For instance, in 2018-19, several countries like Democratic Republic of the Congo, Yemen, Turkey, Burundi were ranked high due their "Regime Type". In 2018-19, EWP classified North Korea as one of the high-risk countries (28th) with risk probability of 4.2%. It appears that since "Regime Type" was dropped after 2018-19, North Korea's ranking has improved significantly. It never featured in the top 30 countries after 2018-19 and currently ranks 76th with risk probability of just 0.3%.

Additionally, the EWP model omits the "Ideological Orientation of the Ruling Elite," a factor that according to Barbara Harff's research, demonstrates a strong correlation with the occurrence of mass killings. This omission is particularly striking given that the founding charter of the United States Holocaust Memorial Museum (USHMM)—the institution overseeing the EWP—was authored by Holocaust survivor Elie Wiesel. The exclusion of "Ideological Orientation of the Ruling Elite" as a risk factor seems especially ironic, considering that this factor was a primary driver behind the Holocaust and numerous other historical mass killings.

In the foundational work by the Political Instability Task Force (PITF), the variable "Armed Conflict in 4+ Bordering States" was identified as a critical predictor of instability and mass killings. The Early Warning Project (EWP), however, appears to have modified this variable to "Region" without providing a rationale for this change. This modification represents a significant departure from the original parameter definition. While the PITF emphasized the presence of armed conflict in at least four neighboring countries as a high-risk indicator, EWP's model altered this parameter by designating certain regions as inherently high-risk, even in the absence of armed conflicts in 4+ neighboring countries. This expansion of the parameter's definition could lead to the classification of numerous countries as high risk, diverging from the more targeted approach of the original study. In a statistical model, independent variables are typically selected based on rigorous research and are grounded in empirical evidence. EWP's alteration of the variable's scope, without providing justification, introduces potential biases that may significantly impact the model's outcomes. This modification raises concerns about the credibility and validity of the EWP model, as the lack of a clear criterion risks shifting focus away from genuinely high-risk countries.

The modification of the variable "Region" exerts a markedly different influence on the EWP outcomes compared to the original study's model. In the original framework, the presence of armed conflicts in four or more neighboring countries significantly increases a nation's risk of mass killing, affecting the rankings of countries accordingly. Consequently, countries such as India, which lack armed conflicts in adjacent nations, are positioned with relatively low risk in the original model based on this parameter. However, in the EWP model, because it is pigeonholed in the Central and South Asian region, it increases the risk of mass killing in India significantly.

Apart from the fact that this variable, "Geographical Region", has been significantly modified from its original study without any explanation, it is a very ambiguous variable that can be looked at in different ways. International organizations classify regions in various ways.

- **UNCTAD**, for instance, classifies the nations into six main categories: Africa, America, Asia , Europe and Oceania .
- **The International Monetary Fund (IMF)** classifies countries into Africa, Asia and Pacific, Europe, Middle East and Central Asia and Western Hemisphere.
- **The World Trade Organization (WTO)** classifies regions such as North America, South/Central America, Europe, Commonwealth of Independent States, Africa, Middle East and Asia.
- **The World Trade Organization (WTO)** classifies regions such as North America, South/Central America, Europe, Commonwealth of Independent States, Africa, Middle East and Asia.
- **The World Health Organization (WHO)** organizes countries into six health-specific regions: the African Region, the Region of the Americas, South-East Asia, Europe, the Eastern Mediterranean, and the Western Pacific.
- **The International Labour Organization (ILO)** groups countries into Africa, Arab States, Asia and the Pacific, and Europe and Central Asia.

In the EWP model, the classification of countries by geographic region follows the U.S. Department of State's framework, whose mission is "To protect and promote U.S. security, prosperity, and democratic values and shape an international environment in which all Americans can thrive." The decision of EWP to choose the U.S. Department of State's definition over alternatives from the United Nations, IMF, WTO, or ILO is surprising, as there appears to be no clear rationale for particularly favoring this classification. While regional classifications typically should not matter, EWP assigns considerable weight to location, significantly influencing a country's risk score for mass killings. For instance, India's geographic region alone ranks as one of the top three factors driving its elevated risk level. Classifying India within the Americas would result in a 50% reduction in its risk score—from 8% to 4%—and an improvement in ranking from 5th to 16th place. (EWP has an interactive tool, "Change a Country's Risk Factors", **https://earlywarningproject.shinyapps.io/risk/** where one can check how changes to a country's risk factors would affect its risk of mass killing, holding all other aspects constant).

The EWP dataset appears to be riddled with errors even in classifying the countries according to regions. For instance, the Maldives is included under Central and South Asia by the U.S. Department of State, while EWP classifies it under East Asia Pacific. Also, given that geographic classification is static and unchangeable by nature, using it as a key parameter risk generating misleading alerts, as countries have no control over their regional designation once assigned.

Jay Ulfelder's article, "A Multimodel Ensemble to Forecast Onsets of State-Sponsored Mass Killing," serves as an academic reference for the EWP model, and Ulfelder contributed to its initial implementation. While he incorporated some variables from Michael Colaresi and Sabine Carey's 2008 study in the Journal of Conflict Resolution, key factors from this research, such as Executive Constraints, Security Forces, and the interaction between these variables, were omitted in EWP's adaptation. Colaresi and Carey's study examined conditions under which state security forces could protect or harm citizens, emphasizing that factors like executive constraints are crucial in understanding the dual role of security forces in genocide risk. The EWP model, however, selectively incorporated "Population Size", a variable minimally emphasized in the original study, as it was controlled primarily to prevent misinterpretation of security force measures.

This selective adaptation of "Population Size", again a static variable, results in significant weight within the EWP model, particularly affecting the classification of countries at high risk. For example, "Population Size" is the most important factor contributing to India's elevated risk ranking. Without "Population Size" as a parameter, India's rank would move exponentially from 5 to 52, or if it had been a least populated country, its risk would reduce by 75% from 8% to 2%, dropping its risk ranking from 5th to 48th.

The inclusion of variables like "*Geographical Region*" and "*Population Size*" in the EWP model raises significant methodological concerns. "*Geographical Region*" appears to be a modified version of the Political Instability Task Force's parameter, "Armed Conflict in 4+ Bordering States," yet it is ambiguous and open to multiple classifications. Meanwhile, "Population Size" is incorporated out of context from Colaresi and Carey's study, omitting critical variables central to that research. Both variables are static and beyond a country's control, yet they disproportionately influence the risk rankings in the EWP model.

For instance, if India had been a country with low population and in American region , its probability of risk of mass killings would reduce from 8% to 1%, and its rank would improve exponentially from 5 to 77. This illustrates how using static, unmodifiable parameters like "*Population Size*" and "*Geographical Region*", about which a country can do nothing in terms of corrective action, can permanently trap countries like India as high risk for mass killings.

The other way of looking at it is, if a low-population country in North America is identical to India on the remaining 30+ socio-economic-political parameters in the EWP model, still that country would rank as 77th, while India would rank as 5th. Even Afghanistan's rank with all it's current turmoil would improve exponentially from 8 to 60 (risk of mass killings reduce by 83% from 6% to 1%) if it were a low population country in American region.

Ethnic Fractionalization is another factor used by EWP, which greatly deviates from the foundational studies. The EWP model defines Ethnic Fractionalization as *"corresponds to the probability that two randomly drawn individuals within a country are not from the same ethnic group"*. The definition is highly problematic, as the EWP model suggests that mere existence of any diversity and pluralism by itself increases the risk of mass killings. It appears that EWP has taken this variable, ignoring the definition and the relevance of ethnicity in genocide by Harff and Colaresi and Sabine Carey on the same parameter. Barbara Harff says that ***"Ethnic heterogeneity is likely to lead to geno-/politicide only if an ethnic minority dominates the elite,"*** but EWP says the mere existence of any diversity is a correlating factor for mass killings by defining Ethnic Fractionalization as *"corresponds to the probability that two randomly drawn individuals within a country are not from the same ethnic group"*.

Colaresi and Sabine Carey did not find ethnic fractionalization to be a statistically significant predictor of the log odds of genocide. They say that *" While there is a fruitful debate concerning the operationalization of ethnic variables in the civil war literature, we are unsure how to translate these into the context of genocide. To be conservative, we have respecified our model with seven different measures of ethnic and religious differences within a country. Those reported in the article are coded from Fearon and Laitin's (2003) measure of ethnic fractionalization (1). We also have run the analysis with Fearon and Laitin's measure of the largest ethnic group in a country (2), the second largest group in the country (3), and our own measure of the ratio of the largest group in the country to all groups smaller than the second largest group (4). Drawing on data collected by Reynal-Querol and Montalvo (2005), we have also analyzed ethnic polarization (5), religious polarization (6), and religious fractionalization (7). In no case did we find any of these variables to be statistically significant predictors of the log odds of genocide. Furthermore, the results and inferences reported in this article are consistent regardless of this specification. A useful discussion of the role of ethnicity in civil wars can be found in Sambanis (2001).that they have respecified their model with seven different measures of ethnic and religious differences within a country. Those reported in the article are coded from Fearon and Laitin's (2003) measure of ethnic fractionalization*___. In no case did we find any of these variables to be statistically significant predictors of the log odds of genocide.___

The EWP model's use of Ethnic Fractionalization appears to distort the variable's interpretation from the foundational studies it has referenced , oversimplifying diversity as a risk factor for mass killings without considering the nuanced contexts highlighted by Harff, Colaresi, and others.

Many assessment factors in the EWP model cannot be traced back to any of the foundational studies cited on its website. Using variables from other studies without considering their original context can diminish a model's relevance and accuracy, as these variables are often tailored to specific frameworks. When adapted without the conditions they were designed for, they may not fully capture the dynamics of the new model's focus, potentially leading to overfitting—where the model appears accurate on past data but fails to generalize to new scenarios. This can reduce predictive validity, making the model's assessments unreliable when applied to diverse real-world situations.

Changing the definitions of variables from the original studies further compounds these issues, as it introduces a departure from the empirical foundations and interpretive frameworks established by prior research. Such modifications alter the intended meaning and role of these variables, often broadening or shifting their scope in ways that distort their predictive accuracy. For instance, redefining a variable like "*Geographical Region*" from its original focus on neighboring armed conflicts to a more general regional classification overlooks specific conflict dynamics, potentially inflating risk scores without evidence-based grounding.
Cherry-picking variables while omitting others, can introduce biases and distort the relationships among variables, disrupting the balance intended in the original research and creating skewed or overly simplified risk assessments. This practice can increase the likelihood of model instability, where predictions fluctuate with minor data changes, further compromising the model's reliability and consistency.

Assigning high coefficients to static parameters like "*Population Size*" and "*Region*" can lead to inflexible, biased predictions that do not adapt to evolving conditions. Such parameters can produce false positives by marking countries as high-risk based on unchangeable factors, limiting the model's practical utility. This approach constraints policy relevance, as countries are unfairly impacted by fixed characteristics they cannot alter, providing little actionable insight for risk mitigation or prevention. Together, these issues undermine the model's ability to deliver accurate, dynamic, and contextually relevant assessments.

Methodological Divergence in the Early Warning Project

The EWP has adopted a statistical tool fundamentally different from those recommended in foundational studies, raising significant methodological concerns. While Ulfelder's foundational work emphasized the use of logistic regression and perhaps Random Forest (RF), EWP has chosen penalized logistic regression (elastic net).

Ulfelder's work recommended logistic regression as a baseline and RF as an advanced ensemble method, with model performance evaluated using the Area Under the Curve (AUC) metric. These approaches prioritized simplicity, robustness, and transparency. For example, Harff's model, which used fewer covariates, consistently demonstrated higher AUC values, showing that simpler models could outperform more complex alternatives.

In contrast, EWP adopted elastic net regression, which introduces regularization penalties to mitigate overfitting. This departure from the foundational studies raises questions about the compatibility of EWP's results with the theoretical and methodological framework established by Ulfelder. Notably, EWP has not provided a clear explanation of the specific issues it found in the original methodologies to justify this shift.

Furthermore, in elastic net, alpha=0.5 is chosen, and the other penalizing parameter, lambda, is selected by cross-validation. In statistics literature, it is recommended to use two-way cross-validation for selecting alpha and lambda.

A critical issue lies in EWP's reliance on cherry-picking of parameters identified in foundational studies within a different statistical framework. Parameters optimized for simpler models, like logistic regression or RF, may not perform equivalently in elastic net, which employs unique penalty structures. This inconsistency risks distorting the relationships between variables and outcomes established in the original research, reducing the validity and reliability of the findings.

The treatment of population size as a predictor highlights the methodological challenges introduced by EWP's approach. The following three scenarios were analyzed related to India's ranking:

1. Allowing population size weights to vary by region.
2. Applying a uniform weight across regions.
3. Excluding population size altogether from the model

Table 6: Scenarios Analyzed for Population Size as a Predictor in EWP				
Model	**Method**	**Rank 19**	**Rank 20**	**Rank 21**
Pop. By reg	Logistic Regression	14	13	39
	Penalized Logistic Regression	13	12	48
Pop. Only	Logistic Regression	7	6	22
	Penalized Logistic Regression	7	7	32
Drop Pop.	Logistic Regression	22	25	42
	Penalized Logistic Regression	23	22	52

The analysis revealed that excluding population size caused dramatic shifts in country rankings. For instance, India's ranking fluctuated significantly from 5 to 52. The choice of elastic net regression introduces additional complexities. It complicates variable weighting and introduces ranking inconsistencies. These issues undermine the transparency and interpretability of EWP's findings.

Low Reliability due to High False Positive Rate (FPR) and High False Negative Rate (FNR)

The Early Warning Project (EWP) model demonstrates significant limitations in predictive performance, as reflected in its high False Positive Rate (FPR) and False Negative Rate (FNR). These shortcomings raise serious concerns about the model's reliability as a forecasting tool for identifying the risk of mass killing onsets.

The model exhibits an exceptionally high FPR of 95.2%, indicating that for every actual mass killing onset, 20 false alarms are generated among countries classified as high risk, with risk scores of 4% or higher. This high rate of false positives is primarily attributed to several factors. First, the EWP model heavily relies on correlational rather than causal relationships among its variables. This approach increases the likelihood of spurious correlations, where countries are erroneously flagged as high-risk due to patterns that do not hold causally. Second, the inclusion of static variables like "Geographical Region" and "Population Size" disproportionately influences risk scores, trapping certain countries in high-risk categories regardless of significant on-ground reforms. Third, the presence of redundant and highly correlated variables, such as those related to civil liberties and political freedoms, exacerbates issues of multicollinearity, which overstates the impact of some factors. Finally, the lack of contextual calibration and transparency in the selection and weighting of variables introduces further inconsistencies, amplifying the rate of false positives.

Similarly, the model's FNR, ranging between 29% and 36%, signifies that nearly 30% of actual mass killing onsets in countries classified as high risk (with risk scores of 4% or higher) are missed by the model. This high rate of false negatives stems from the EWP's non-causal framework, which fails to establish definitive relationships between risk factors and the actual occurrence of atrocities. The implications of these high error rates on the reliability of the EWP model are substantial. A high FPR diminishes the model's practical utility by creating a flood of false alarms, leading to resource misallocation and "alert fatigue" among policymakers and practitioners. Efforts may be disproportionately directed toward countries with minimal actual risk, detracting from timely interventions in genuinely high-risk scenarios. Conversely, a high FNR undermines the model's core objective of early warning and prevention, as more than a third of actual mass killing onsets remain unpredicted, reducing the opportunity for proactive measures. Together, these inaccuracies erode the model's credibility among researchers and policymakers, potentially limiting its influence in shaping strategies for atrocity prevention. Moreover, the reliance on static and correlational parameters, coupled with a lack of transparency in model updates, hampers iterative improvements that could address these shortcomings.

Significant Logical Inconsistencies in the EWP model

EWP reports say that some of the relationships between the data and the output will be perplexing: *"We make no effort to explain these kinds of relationships in the data, some of which can seem perplexing; we only use them for their predictive value."* This is because the EWP model has significant logical inconsistencies due to statistical and methodological flaws, including spurious correlations, modified variable definitions, multicollinearity and improper weighting.

Cherry-picking parameters from various studies without accounting for each factor's original context may also have compounded this problem.

Further, modifying variable definitions from foundational studies has led to inconsistent effects. Using highly correlated variables (multicollinearity) as independent predictors—like freedom of speech and civil society repression—also skews results, distorting risk estimates and causing erratic changes in rankings. Another problem is the improper weighting of variables, where some factors may have an exaggerated influence, leading to illogical rank changes. Together, these issues weaken the model's coherence and predictive validity, undermining its reliability as a tool for assessing risk.

EWP's Changing Parameters: A Threat to Statistical Rigor and Credibility

The EWP has made significant changes to its assessment parameters over the years, yet it lacks a consolidated and transparent record of these modifications on its official website. While the annual reports provide some information regarding parameter inclusions and exclusions for specific years, they do not clarify whether these were the only changes implemented, thereby creating ambiguity about the evolution of the model. For instance, the parameters "*Regime Type*" and *"Regime Duration"* were reportedly utilized until 2019-20 but were excluded from 2020-21 onwards due to the discontinuation of updated data from the Center for Systemic Peace, the source of these variables. This is particularly concerning, as "*Regime Type*" has been a critical factor in explaining the likelihood of mass killings as per the studies referenced by EWP. Similarly, the parameter *"Freedom of Movement of Men,"* used until 2022-23, was removed in 2023-24 because it produced large year-to-year shifts in risk estimates that EWP stated it could not adequately explain. It is also important to note that *"RegimeType"* and *"Freedom of Movement of Men,"* were very important factors for high rankings of many countries in the EWP model till thev were discontinued.

The Early Warning Project (EWP) reports asserts that the exclusion of these variables did not significantly impact the overall accuracy of the model based on most measures while this adjustment may have contributed to changes in the risk rankings of specific countries. Given the EWP model's already limited predictive accuracy, characterized by a false positive rate of 95.2% and a false negative rate between 29% and 36%, the assertion that the removal of these two variables has not affected the model's accuracy appears logically inconsistent as the model's accuracy is already very low.

In 2023-24, EWP has also introduced certain new variables, including women's participation in civil society organizations, government censorship of the media, and discrimination against ethnic groups. In 2024-25, EWP dropped the tradeshare variable (sum of imports and exports divided by GDP) because of a significant degree of missing data. However, no theoretical justification has been provided for the inclusion or exclusion of these parameters. This lack of theoretical grounding undermines the robustness of the statistical model, as decisions about variable selection appear ad hoc rather than being guided by a clear framework or empirical rationale. Such an approach risks compromising the credibility of the model, particularly given the historical importance of parameters like "*Regime Type*" in understanding mass killing dynamics. Treating variables as interchangeable or dispensable without proper justification is inconsistent with the principles of robust statistical modeling, where the inclusion or exclusion of variables should be firmly anchored in theoretical and empirical evidence. The Early Warning Project appears to prioritize a more ad hoc approach to the selection and treatment of independent variables, rather than grounding its model in a rigorous theoretical and empirical framework.

EWP Model Equates Complete Lack of Civil Liberties with Full-civil Liberties

One of the risk assessment factors employed by the EWP model is the measure of government respect for civil liberties across different geographic regions within a country. EWP captures this as below:

"Is government respect (or lack of respect) for civil liberties equal across different geographic areas of the country?" The model classifies countries into one of the following options:

- ➔ Not equally distributed
- ➔ Equally distributed

There is no option to identify non-existence of civil liberties in any geography. This classification couches reality as it equates countries with no civil liberties with those with full civil liberties, though the reality of civil liberties in the two categories of countries will be very different. However, because of EWP's model, countries where civil liberties are either uniformly upheld or completely denied are clubbed together. In the EWP model, this has resulted in the paradoxical classification of autocratic and totalitarian countries, such as North Korea, which do not have any civil liberties whatsoever, and fully democratic nations like the United Kingdom, Australia, France, Germany, and Japan, which come out on top in this parameter. In contrast, democracies like India and the United States, where civil liberties may vary by region, score poorly on this parameter as below[4]:

4. **Source:** https://earlywarningproject.ushmm.org/downloads

Table 7: Civil Liberty Scores of Selected Countries in the EWP Model	
risk_in_2024_25country	**even_civil_rights**
India	0
North Korea	1
United States of America	0
Japan	1
United Kingdom	1
Germany	1
France	1
Australia	1

This inconsistency in methodology is problematic, as it skews the comparative ranking of countries significantly. This classification disturbingly supports repression and the complete removal of civil liberties to reduce mass killings. For instance, if India were to become fully autocratic and entirely suppress civil liberties across all its territories, its risk of mass killing would be reduced dramatically, with its rank improving exponentially by 12 places from 5th to 17th and its estimated risk decreasing from 8% to 4%. Conversely, if North Korea were to democratize partially and allow civil liberties in some geographic regions, its risk ranking would deteriorate significantly by 30 places, moving from 71 to 40 as below:

Table 8: Impact of Policy Actions on Country Rankings in the EWP Model		
Country	**Action**	**Country Ranking in EWP Model**
India	Suppress civil liberties throughout its geographical territory	5 to 17
North Korea	Democratize and allow civil liberties in some geographic regions	71 to 40

Such results raise important questions about the logic and internal consistency of the EWP's risk model, particularly regarding its treatment of civil liberties. This methodological choice to equate two opposite extremes—no civil liberties and full civil liberties —rewards autocratic governance.

EWP Model Adopts the Erroneous Principles of "All Equal None" or "Nothing Equals Something"

The EWP model has clubbed the extremes not just with civil liberties across the geography but across several other parameters. Moreover, this binary approach stands in contrast to the EWP model's graded approach for other parameters. For example, in assessing "freedom of religion," the model employs a scale from "0" to "4," where "0" represents no freedom of religion, "4" represents full freedom, and intermediate values (1, 2, 3) reflect varying degrees of religious freedom. However, for the parameter, "Do some social groups—as distinguished by language, ethnicity, religion, race, region, or caste—have fewer civil liberties than others?", the EWP model classifies the countries in only one of the following two options:

- Members of all salient social groups enjoy the same level of civil liberties
- Members of some social groups enjoy fewer civil liberties than the general population

Here too, the EWP model also equates two extremes: no social group enjoys any civil liberties with all social groups enjoying all civil liberties. There is no option to identify the non-existence of civil liberties for any social group. Hence, authoritarian regimes like North Korea, in which civil liberties are non-existent, are clubbed along with full democracies like the United Kingdom. For the parameter, *"Does the government directly or indirectly attempt to censor the print or broadcast media?",* the EWP model classifies the countries in one of the following two options:

- Attempts at censorship are rare and/or limited.
- Attempts at censorship are direct and/or routine.

The EWP model does not differentiate between the countries where there is no Free Media at all and the countries where there is Free Media, and hence will categorize countries like Poland and India along with North Korea, which does not have any Free media at all.

For the parameter, "Were the judiciary's formal powers altered this year in ways that affect its ability to control the arbitrary use of state authority?" the EWP model classifies the countries in one of the following options:

- Judiciary's ability to control arbitrary power was reduced via institutional reform.

- There was no change to the judiciary's ability to control arbitrary power via institutional review.
- Judiciary's ability to control arbitrary power was enhanced via institutional reform.

The EWP model does not separately identify the countries where the judiciary has no power at all to control arbitrary power.

By adopting a **binary or overly simplified classification system**, the model collapses a wide range of realities into a limited number of categories. For instance, treating "no civil liberties" and "uniform civil liberties" as equivalent misses crucial variations within and between countries. This reduces the **granularity** of data input into the model, which impacts the **predictive power** of the statistical algorithms. When countries with vastly different realities are lumped together (e.g., North Korea and the United Kingdom in the civil liberties parameter), the **weight assigned to that parameter** becomes **misleading**. In statistical models, proper weighting is critical for accurate forecasting. Misclassifying nations can lead to **incorrect weight distributions** for the parameters, causing **biased risk assessments** and distorting the model's overall performance.

By treating different extremes as identical (e.g., "no civil liberties" and "uniform civil liberties"), the model introduces **statistical noise**. This makes it harder to identify genuine risk signals because the underlying data is too noisy and generalized to be useful. This generalization leads to a loss of **statistical significance** in the model, as important distinctions between country contexts are erased. It results in **overfitting** to certain patterns that do not generalize well across different types of regimes. Although EWP utilizes data from the Varieties of Democracy (V-Dem) project for certain parameters, it has modified the original V-Dem scales, resulting in a loss of data granularity and nuance. By condensing V-Dem's multi-point scales into simplified binary or limited classifications, EWP risks distorting the source data, potentially leading to inaccurate country assessments.

statistical noise. This makes it harder to identify genuine risk signals because the underlying data is too noisy and generalized to be useful. This generalization leads to a loss of **statistical significance** in the model, as important distinctions between country contexts are erased.

It results in **overfitting** to certain patterns that do not generalize well across different types of regimes. Although EWP utilizes data from the Varieties of Democracy (V-Dem) project for certain parameters, it has modified the original V-Dem scales, resulting in a loss of data granularity and nuance. By condensing V-Dem's multi-point scales into simplified binary or limited classifications, EWP risks distorting the source data, potentially leading to inaccurate country assessments.

Table 9: EWP's Modifications of Scales to V-Dem Data	
V-DEM	**EWP**
Do all social groups, as distinguished by language, ethnicity, religion, race, region, or caste, enjoy the same level of civil liberties, or are some groups generally in a more favorable position? 0: Members of some social groups enjoy much fewer civil liberties than the general population. 1: Members of some social groups enjoy substantially fewer civil liberties than the general population. 2: Members of some social groups enjoy moderately fewer civil liberties than the general population. 3: Members of some social groups enjoy slightly fewer civil liberties than the general population. 4: Members of all salient social groups enjoy the same level of civil liberties.	"Do some social groups—as distinguished by language, ethnicity, religion, race, region, or caste—have fewer civil liberties than others?", 0: Members of all salient social groups enjoy the same level of civil liberties. 1: Members of some social groups enjoy fewer civil liberties than the general population.
Does the government directly or indirectly attempt to censor the print or broadcast media? 0: Attempts to censor are direct and routine. 1: Attempts to censor are indirect but nevertheless routine. 2: Attempts to censor are direct but limited to especially sensitive issues. 3: Attempts to censor are indirect and limited to especially sensitive issues.	"Does the government directly or indirectly attempt to censor the print or broadcast media?" 0: Attempts at censorship are rare and/or limited. 1: Attempts at censorship are direct and/or routine.

4: The government rarely attempts to censor major media in any way, and when such exceptional attempts are discovered, the responsible officials are usually punished.	
Is political power distributed according to social groups? 0: Political power is monopolized by one social group comprising a minority of the population. This monopoly is institutionalized, i.e., not subject to frequent change. 1: Political power is monopolized by several social groups comprising a minority of the population. This monopoly is institutionalized, i.e., not subject to frequent change. 2: Political power is monopolized by several social groups comprising a majority of the population. This monopoly is institutionalized, i.e., not subject to frequent change. 3: Either all social groups possess some political power, with some groups having more power than others, or different social groups alternate in power, with one group controlling much of the political power for a period of time, followed by another — but all significant groups have a turn at the seat of power. 4: All social groups have roughly equal political power, or there are no strong ethnic, caste, linguistic, racial, religious, or regional differences to speak of. Social group characteristics are not relevant to politics.	Is political power distributed according to social groups? 0: All social groups have roughly equal political power or there are no strong ethnic, caste, linguistic, racial, religious, or regional differences to speak of. Social group characteristics are not relevant to politics. 1: Political power is monopolized by one or a few social groups.

As illustrated in the table 7, countries with markedly different levels of civil liberties, media censorship, or political power distribution receive similar scores under EWP's oversimplified framework, masking important contextual distinctions captured in V-Dem's approach. This divergence introduces discrepancies in country scores between the two models, ultimately compromising the EWP model's reliability. By reducing complex socio-political realities to broad categories, EWP's classifications fail to accurately represent the variability and depth of conditions on the ground, affecting the model's capacity to provide credible risk assessments.

EWP model says banning political parties reduces risk of mass killings

"Party Ban" is one of the parameters in the EWP model. For this parameter, the EWP model classifies the countries into one of the following options:

- All parties except the state-sponsored party (and closely allied parties) are banned, or elections are non-partisan, or there are no officially recognized parties.
- Some parties are banned.
- No parties are officially banned.

The EWP model suggests multi-party democracy worsens the risk of mass killings, whereas banning of all parties except the state-sponsored party reduces the risk of mass killings. For example, according to EWP data:

- Some political parties are banned in Chad. If Chad were to ban all political parties except the state- sponsored party or if there are no officially recognized parties, it's risk of mass killings would be reduce by 42% from the current 12% to 7% and the country rank would improve from 1 to 5.
- No parties are officially banned in the UK, France and India. If these countries were to ban all the political parties except the state-sponsored party or if there are no recognized parties, then the risk of mass killings in these countries would reduce significantly:
    - The UK's rank would improve by 16 points from 113 to 129.
    - France's rank would improve by 18 points from 122 to 140.
    - India's rank would improve by 3 points from 5 to 8.
- All parties except the state-sponsored party are banned in China, North Korea and Saudi Arabia. If these countries were to allow multi-party democracy and incorporate this change in EWP's interactive tool, "Change a Country's Risk Factors", then the risk of mass killings in these countries would increase significantly.
    - China's rank would deteriorate from 26 to 18.
    - North Korea's rank would deteriorate from 71 to 58.
    - Saudi Arabia's rank would deteriorate from 103 to 90.

EWP says, repressing civil society organizations reduces mass killings

"Civil Society Repression" is one of the parameters in the EWP model. As per EWP,

- Currently, according to EWP, civil society is not repressed in China and if China were to repress the civil society, it's rank would improve by two points from 26 to 28.
- Currently, Civil Society is repressed in North Korea and if North Korea give freedom to it's civil society, its rank would deteriorate by 6 points from 71 to 65.

EWP says increasing Judiciary's ability to control arbitrary power increases mass killings

"Judicial Reform" is one of the parameters in the EWP model. For this parameter, the EWP model classifies countries into one of the following options:

- The judiciary's ability to control arbitrary power was reduced via institutional reform.
- There was no change in the judiciary's ability to control arbitrary power via institutional review.
- The judiciary's ability to control arbitrary power was enhanced via institutional reform.
- According to EWP's model, if the judiciary's ability to control arbitrary power was reduced, it reduces the risk of mass killings and vice versa. For example, If China's judiciary's ability to control the arbitrary power is reduced, the country's rank would improve from 26 to 28, and if the judiciary's ability to control the arbitrary power increases, the country's rank would deteriorate from 26 to 24.

EWP model greatly incentivizes authoritarianism and various forms of repression of liberties

---

According to the EWP model, the risk of mass killing significantly reduces if India and China become more authoritarian, curtail women's rights and repress various forms of liberties.

Table 10: Impact of Political and Civil Actions on Risk and Ranking in the EWP Model - India and China

Country	Action	Impact in EWP Model	
		Risk of Mass Killing	Country Rank
India	○ Become fully autocratic and entirely suppress civil liberties across all its territories ○ Curtail women's rights in such a way that they are often or always prevented from participating in civil society organizations ○ Ban all parties except state-sponsored party ○ Curtail judiciary's ability to control arbitrary power ○ Repress civil society organizations ○ Do not respect freedom of Religion ○ Do not respect Freedom of Discussion	8% to 3%	5 to 32
China	○ Become fully autocratic and entirely suppress civil liberties across all its territories ○ Curtail judiciary's ability to control arbitrary power ○ Repress civil society organizations ○ Prevent women from participation in civil society organizations	3% to 1%	26 to 62

According to the EWP model, the risk of mass killing significantly increases if the countries become more free, democratic, and allow various forms of Civil liberties, shown as follows:

Table 11: Impact of Political and Civil Actions on Risk and Ranking in the EWP Model – China and North Korea

Country	Action	Impact in EWP Model	
		Risk of Mass Killing	Country Rank
China	○ Enhance the judiciary's ability to control arbitrary power via institutional reform. ○ No ban on political parties	3% to 4%	26 to 17
North Korea	○ Enhance the judiciary's ability to control arbitrary power via institutional reform. ○ No ban on political parties ○ Allow civil liberties in some geographical regions ○ Stop repression of civil society ○ Freely allow women's participation in civil society organizations	1% to 3%	76 to 21

EWP model permanently traps countries like India as high risk for mass killings, irrespective of major reforms due to high weightage for static factors like "Population Size" and "Region"

The EWP model has several factors, which are either static or historical with high coefficients like Population Size, Geographical Region etc. In addition, the model is built on the logic that the risk of mass killing increases with the increase in various forms of freedoms. Given this peculiar aspect of the EWP model, countries with large populations or certain geographic locations are trapped in high-risk categories permanently even if they implement a slew of democratic reforms. For instance, even if India undertakes significant initiatives towards freedom and civil rights, there will be no change in the country's risk and ranking in EWP model:

Table 12: No Effect of Institutional and Social Reforms on Risk and Ranking of India in the EWP Model

Country	Action	Impact in EWP Model	
		Risk of Mass Killing	Country Rank
India	○ Enhance the judiciary's ability to control arbitrary power via institutional reform. ○ Fully respect freedom of discussion ○ Fully respect freedom of religion ○ No government censorship ○ No discrimination against any section of the population	8% to 8% (No change)	5 to 5 (No change)

The EWP model's reliance on heavily weighted static factors like population size, geographical region etc creates a structural bias that perpetually classifies certain regions as high-risk, regardless of ongoing socio-political reforms. These unchangeable factors overshadow the influence of dynamic variables, such as policy changes and democratic reforms, limiting the model's responsiveness to real-time developments and resulting in an inflexible risk profile. Furthermore, the model's anti-democratic logic penalizes countries for enacting reforms by assuming that democratization efforts, such as enhancing judicial independence, civil liberties, and minority protections, increase instability and mass-killing risk. This approach often produces misleading "false positives" that inaccurately flag reform-oriented democracies as high-risks.

Additionally, the disproportionate weight given to static factors systematically biases the model against countries with large populations or high ethnic diversity like India. This inherent bias undermines the universality of the model, suggesting it may reinforce geopolitical stereotypes rather than serve as an unbiased tool for policy guidance.

Correlated Variables Contradict: Structural Issues in the EWP Model

In the EWP model, factors that should usually work together in the same direction end up with opposite effects in the model and vice versa. For instance,

- **Repression of Civil Society and Party Bans:** Non-repression of civil society organizations, freedom to start and run political parties are directly aligned and should have the same directional impact on the outcome. According to the EWP model, there is no repression of civil society in China. If China were to repress its civil society, its rank in the risk of mass killings improves 2 points from 26 to 28 while allowing all political parties increase the risk of mass killings and deteriorates it's rankings from 26 to 18.

The above logical inconsistencies of the EWP model could be due to several statistical and methodological flaws including spurious correlations, cherry-picked parameters, modified variable definitions, multicollinearity and improper weighting. Further, modifying variable definitions from foundational studies can lead to inconsistent effects. Using highly correlated variables (multicollinearity) as independent predictors also skews results, distorting risk estimates and causing erratic changes in rankings. Improper weighting of variables is another problem, where some factors may have an exaggerated influence, leading to illogical rank changes. Such illogical rank changes make the model difficult to interpret and trust, as rankings often shift counterintuitively. Together, these issues weaken the model's coherence and predictive validity, undermining its reliability as a tool for assessing risk.

Errors/Logical Inconsistencies / Lack of Evidences in the EWP Data Points

Beyond the logic-related issues discussed above, the EWP model's risk assessment framework is impacted by data capture errors either in the source data or in capturing the source data into their dataset.

Selectivity Related to Battle-related Deaths

It is one of the important parameters contributing to the classification of India as a country with high risk of mass killing. While the PRIO and UCDP data has recorded battle deaths in many countries, the EWP model says battle deaths have taken place in only 33 of the 168 countries in its dataset[3]. EWP model has not acknowledged the mass killings and battle deaths that occurred in several countries, which would have made a significant change in the rankings. To illustrate:

- Battle deaths in Sri Lanka due to conflicts with LTTE have not been considered, and hence EWP classifies Sir Lanka as a country with no mass killings.
- Several battle deaths in China due to various conflicts have not been considered and hence EWP classifies China as a country with no mass killings.
- Battle deaths in Greece due to conflicts with DSE have not been considered, and EWP classifies Greece as a country with no mass killings.
- Battle deaths in France due to conflicts with OAS have not been considered, and EWP classifies France as a country with no mass killings.

The credibility of the EWP model is undermined by its failure to include battle-related death data that is already well-documented and widely known in the public domain. Despite significant battle deaths recorded in countries like China, Sri Lanka, Greece, France and several other countries, the EWP model omits these figures, which are well-established both in these datasets and in public knowledge.

3. **Source:** https://earlywarningproject.ushmm.org/downloads

Selectivity and Factual Errors Related to Mass Killings Data

EWP defines mass killings as, *"A mass killing occurs when the* ***deliberate actions*** *of armed groups in a particular country (including but not limited to state security forces, rebel armies, and other militias) result in the deaths of at least* ***1,000 noncombatant civilians*** *in that country over a period of* **one year or less***."*

EWP has published the Mass Killings Events Master List on its website. For India, it says that the Maoist insurgency by several groups with communist ideologies in several Indian states, targeting civilians, has been an ongoing mass killing since 2004. However, neither the data sources relied upon by EWP (Peace Research Institute Oslo (PRIO) and Uppsala Conflict Data Program (UCDP)) nor the Indian government's official data[1] substantiate this classification. According to all these sources, deaths due to Maoist insurgency have been much less than 1000 per year and recently have come down to less than 100 per year.

While including Maoist killings in India as one of the past and ongoing mass killing events, even though the data says otherwise, the EWP Mass Killings Events Master List does not even list various mass killings that have happened in many countries that are very much part of public knowledge. For example, deaths in Greece and France due to conflicts with DSE and OAS are neither accounted for battle deaths nor mass killings.

The omission of such widely recognized mass killing events in several countries, along with the inconsistent application of the mass killing criteria (as seen with the Maoist insurgency in India), raises significant concerns about the accuracy and completeness of EWP's dataset. These exclusions, coupled with unsubstantiated inclusions, call into question the integrity of EWP's data and the validity of its statistical model, as the exclusion of such events, which are very much part of public knowledge from the mass killings list, raises serious concerns about the integrity of the EWP's data set and the results of the statistical model.

No Data, Just Anonymous Opinions for Many Parameters

EWP has rated countries on the following 15 parameters based on inputs from Varieties of Democracy (V-DEM):

- Civil Society Repression
- Freedom of Discussion
- Freedom of Religion
- Government Censorship
- Inequalities in Civil Liberties – Geographical Region
- Inequality in Civil Liberties — Social Group
- Political Killing
- Prevention of Women’s Participation in Civil Society Organizations
- Judicial Reform
- Minority Control
- Party Ban
- Power Distributed by Social Group
- Power Distributed by Socio-economic Position

V-DEM does not provide any evidence for their ratings, which are based on subjective opinions. It appoints a few people as country or subject matter experts who are anonymous and whose identity is not revealed. These experts give their ratings on different parameters, but no evidence is asked by VDEM or provided by those experts for their ratings. As per the communication from VDEM, *“The country experts do not provide any reasoning for their coding* decisions”. However, they have the possibility to comment on certain aspect if they like. However, these comments are not publicly available for data protection issues.“ and “we do not collect information about what events / developments drive the different democracy scores”.

This reliance on VDEM’s ratings, which are nothing more than the opinion of anonymous individuals is contrary to the claim made by EWP on its website related to the Data Sources. *“Our risk assessment relies on publicly available data on a variety of country characteristics. In 2017, we updated our data sources as new datasets became available—for example, measures of civil liberties and government repression. We also took special care to avoid using data that, in our judgment, could be susceptible to bias when coded or recoded retrospectively. This should help ensure that our model performs as well in practice as it does on historical data.”*

The EWP model makes a guesswork to selectively validate VDEM’s ratings, which cannot be considered as an evidence in a statistical model.

For instance, EWP says the following regarding Chad in its 2022-23 report: *"Chad continues to move up our risk list, landing at fourth with a risk estimate above nine percent for 2022 - 23 after being ranked tenth last year and 23rd the year before. Chad has consistently ranked in the high-risk (top-30) category, with fourth marking its highest ranking to date. This shift can be most attributed to a decrease in freedom of movement for men, according to V-Dem. Although V-Dem does not indicate the reason for specific changes in their coding, it may have been related to the government's crackdown on opposition in advance of the April 11, 2021 presidential elections."*

Source inconsistency in Capturing "Ethnic Fractionalization" Data

The EWP categorizes India's population into "Indo-Aryans" (72%) and "Dravidians" (25%)—a colonial construct based on the discredited Aryan invasion theory, which has not been used by India since its independence. EWP's inconsistency in data sourcing compounds this issue. While census data is used for countries like the US, France, Israel and New Zealand, it has used outdated Encyclopedia Britannica Data from 2000 for India, despite the availability of the 2011 Census. This reliance on fundamentally flawed data and inconsistent use of sources introduces biases and affects the statistical integrity of the model, potentially distorting risk assessments.

Contradictions among Data Sources in "Power Sharing among Social Groups"

The EWP utilizes the Ethnic Power Relations (EPR) dataset to evaluate levels of discrimination and power dynamics within countries by examining the distribution of political authority among various social groups. The EPR employs an ordinal scale to categorize populations based on their access to executive power, distinguishing groups into classifications such as *"Monopoly," "Dominance," "Senior Partner," "Junior Partner," "Powerless," "Discrimination," "Self-exclusion,"* and *"Irrelevant".* These classifications capture the degree of control, power-sharing, or exclusion experienced by each group in relation to executive decision-making.

For instance, groups classified as *"Dominant"* wield substantial executive power while allowing only minimal, token representation from other groups, who lack real decision-making influence. In contrast, groups labelled *"Powerless"* hold no political power at the national level, although they may not face explicit discrimination.

United States of America: EPR data says that whites, who form 60% of the US population, are DOMINANT, while Latinos, African Americans, Asian Americans, American Indians and Arab Americans, who form 39.6% of the population, are POWERLESS. Despite this data, the EWP model says that ***all social groups in the US have roughly equal political power or there are no strong ethnic, caste, linguistic, racial, religious, or regional differences to speak of. Social group characteristics are not relevant to politics***. In essence, while EPR data says that 40% of US citizens are powerless, hold no political power, and do not have influence on decision-making, EWP in its model recorded a completely contrary idea that all groups have equal political power. It appears that for this parameter, EWP has chosen to rely on opinion-based V-DEM data rather than evidence-based EPR data. This error can be observed for many countries, which is illustrated in the following sample shown in table 13.

<table>
<tr><th colspan="4">Table 13: Analysis of Social Group Power Distribution and Discrepancies in EWP Data</th></tr>
<tr><th>Country</th><th>Percentage of Population classified as "Powerless /Discriminatory" as per EPR</th><th>EWP data on "Power Distributed by Social Group"</th><th>Remarks</th></tr>
<tr><td>United States of America</td><td>39.6%</td><td rowspan="4">All social groups have roughly equal political power, or there are no strong ethnic, caste, linguistic, racial, religious, or regional differences to speak of. Social group characteristics are not relevant to politics.
All social groups have roughly equal political power, or there are no strong ethnic, caste, linguistic, racial, religious, or regional differences to speak of. Social group characteristics are not relevant to politics.</td><td rowspan="4">While EPR has identified a substantial part of the population in these countries as powerless or discriminated against, the EWP model erroneously states that all social groups have equal political power.</td></tr>
<tr><td>Kenya</td><td>31%</td></tr>
<tr><td>United Kingdom</td><td>6%</td></tr>
<tr><td>Nepal</td><td>50%</td></tr>
</table>

Selectivity in Capturing the "Discrimination" Data

The EWP faces a critical issue in its measurement of discrimination due to its selective reliance on the *"Discriminated"* category from the Ethnic Power Relations (EPR) Dataset, excluding groups classified as *"Powerless"* and *"Junior Partner".* According to EPR, *"Powerless"* groups lack political power but are not explicitly discriminated against, while *"Junior Partner"* groups hold limited influence. Despite these groups' political marginalization, EWP's focus solely on active, intentional exclusion overlooks their exclusion from power. This selective classification raises concerns about EWP's comprehensiveness in assessing discrimination. For example, EPR categorizes 39.6% of the U.S. population as *"Powerless,"* yet EWP report no discrimination in the country, ignoring substantial political marginalization.

Furthermore, EWP claims an absence of any discrimination in 128 of 168 countries, disregarding large segments of populations classified by EPR as *"Powerless"* or *"Junior Partner."* Such omissions may underreport systemic exclusion in societies where power is concentrated among dominant groups, while others remain marginalized. By excluding *"Powerless"* and *"Junior Partner"* groups, EWP risks overlooking political dynamics that contribute to societal tensions and conflict. Political exclusion, even when not explicitly labeled as discrimination, often generates grievances that can escalate. This lack of transparency and exclusion of politically marginalized groups undermine EWP's model reliability in forecasting potential risks of violence, as it fails to capture the full spectrum of discrimination necessary for accurate early warning assessments.

Factual Errors and Logical Inconsistencies in the EWP Model

It appears that EWP has not done the basic verification of the logical inconsistency and factual correctness in its dataset. For instance:

- EWP data says that civil society is not repressed in China. This is factually incorrect, as there is no civil society and civil liberties in China in the first place.
- EWP data says that women are rarely or never prevented from participating in civil society organizations in China. This is factually incorrect because when there is no existence of civil society in China, it is logically inconsistent to say that women are never prevented from participating in civil society organizations.
- EWP data says that civil society in North Korea is repressed, women are often or always prevented from participating in civil society organizations and political power is monopolized by one or few social groups, while simultaneously claiming that members of all salient social groups and all geographies enjoy the same level of civil liberties. These two assertions are contradictory, as repression and exclusion imply absence or unequal distribution of civil liberties.

Changes in the Intercept and Coefficients

The statistical model under discussion demonstrates significant weaknesses, particularly in its intercept values and the instability of coefficients across key variables. These issues raise concerns about the model's stability, reliability, and overall validity. For instance, the intercept values, which represent the baseline level of the dependent variable when all predictors are zero, have fluctuated drastically over time. Between 2019 and 2022, the intercept remained relatively stable, ranging between -8.8 and -11.2. However, in 2023, this value abruptly escalated to 67.96, a monumental shift that is difficult to justify within a theoretically sound framework. Such a drastic change suggests that underlying issues in the data or the model's structure have not been adequately addressed, undermining confidence in the model's ability to provide consistent results.

This instability extends to the coefficients of key variables, which show significant directional and/or magnitude shifts over time. Variables such as anymk.ever (indicating any mass killing ever), countryage.ln (log-transformed age of the country), even_civilrights (measuring civil rights equality), reg.sca (regional social cohesion), and repress_civilsoc (repression of civil society) have all demonstrated substantial changes. Moreover, some variables have even changed direction, further complicating the interpretation of their effects. These shifts suggest that the relationships between predictors and the dependent variable are not only inconsistent but also potentially influenced by arbitrary changes in the model's structure, including the inclusion or exclusion of variables.

The variability in the model can be attributed to several underlying issues. First, the selection of variables appears to lack a solid theoretical foundation. Many variables with negligible or zero impact are included in the model, adding noise and diluting its predictive capacity. Second, the frequent adjustments to variable inclusion and their associated weightages seem arbitrary and lack clear justification, which undermines the model's coherence. Third, there appears to be no systematic effort to eliminate spurious correlations, which further degrades the model's reliability. These weaknesses have profound implications for the model's reliability and usability. The drastic changes in coefficients reduce the model's predictive power, as the relationships between variables and outcomes are not robust. This lack of stability also makes it difficult for stakeholders to interpret the importance or direction of variables, particularly when coefficients change unpredictably over time. Additionally, the absence of a clear theoretical basis for variable selection and the lack of transparency regarding adjustments to the model's structure erode its credibility. Without documented justifications for these changes, the findings of the model become questionable and unreliable. The statistical model suffers from foundational weaknesses that significantly compromise its reliability and interpretability. The drastic changes in intercept values, the variability in coefficients, and the apparent lack of a systematic, theoretically grounded approach to variable selection all contribute to the model's instability.

Discrepancies in EWP's Model Assumptions vs. Public Reporting

The Early Warning Project (EWP) publishes its *Statistical Risk Assessment Results* as annual reports, titled "Countries at Risk for Mass Killing Statistical Risk Assessment Results." These reports include a section highlighting high-risk countries, where explanations for elevated risk rankings are provided. While the EWP's statistical model implies that anti-democratic actions—such as the repression of civil liberties and banning political parties—lower the risk of mass killings, the reports aim to position the model as supportive of democratic norms and civil liberties. For example, according to the EWP's statistical model, if Chad bans all political parties except State-Sponsored party, repress Civil Society and remove civil liberties across all it's territories, it's risk of mass killing will reduce by 75% from the current 12% to 3% and it's rank would improve exponentially from 1 to 22.

Currently, women in Afghanistan lack civil rights and face restrictions on participating in civil society organizations. However, if Afghanistan were to grant full civil rights to women, there will be no change in the risk probability and the rankings of Afghanistan in the EWP model. But annual EWP reports laments about curtailing of women's rights by Taliban regime in Afghanistan while their statistical model do not correlate curtailment of women's rights to increase risk of mass killings. EWP's 2024–25 notes, *"Following the Taliban's takeover in August 2021, the de facto authorities have significantly eroded rights and freedoms, including by severely restricting women's rights and suppressing political opposition…The Taliban continues to restrict women's rights, severely curtailing their freedom to work, travel, pursue education beyond the sixth grade, and, under a new law, even speak in public."* The 2022–23 report similarly notes, *"The Taliban has eroded rights for women and girls, systematically excluding them from Afghan society, including by limiting their freedom of movement and access to employment, education, and other essential services."* These statements create an impression of alignment with pro-democratic principles, despite the model's underlying statistical inference that expanding women's rights would not lead to any reduction in mass killings.

While the Simon-Skjodt Center, under which EWP operates, asserts that its mandate is "to alert the United States' national conscience, influence policymakers, and stimulate worldwide action to prevent and work to halt acts of genocide or related crimes against humanity, and advance justice and accountability," the EWP model statistically supports the restriction of civil liberties as a means of reducing mass killing risk. The reports, however, tend to obscure this stance, presenting the absence of civil liberties across a country as "equal levels of respect for civil liberties across areas within a country." This framing may mislead casual readers about the model's actual implications.

By framing repressive policies as risk-lowering measures without clarifying the model's underlying assumptions, the EWP reports attempt to portray an impression of alignment with human rights norms while statistically advocating for the restriction of civil liberties.

Final Report Different from the Model's Findings

The EWP reports exhibit a substantial reliance on media sources and opinion pieces rather than on its Statistical Risk Assessment Model, raising concerns about the integrity and rigor of its conclusions. For instance, in its 2023-24 report on India, the EWP states, *"India has ranked in the top-15 highest-risk countries of our assessment for several years. Reports of targeted violence against ethnic and religious minorities have continued in 2023. The Hindu nationalist-led government has espoused hate speech against the country's Muslim minority. Hindu-nationalist groups have engaged in anti-Muslim mob violence with impunity. Government authorities have repeatedly responded to these attacks by arresting Muslims and destroying Muslim property. Ahead of the general elections scheduled for 2024, some journalists have raised concerns about an increased risk of violence against Muslim communities. Attacks, including sexual violence targeting women and girls, against primarily Kuki ethnic communities in Manipur resulted in an estimated 160 people killed, over 300 injured, and thousands displaced by mid-August 2023. Additionally, Christian, Dalit, and Adivasi communities have faced targeted abuses."* However, these claims are not generated from the EWP's statistical model but are sourced from various external media outlets, including *Le Monde*, *Bloomberg*, *Human Rights Watch* , *Al Jazeera*, *The Atlantic*, *The Wire*, *VOA News* , among others. This reliance on media reporting rather than model-derived findings suggests that the EWP's conclusions on India are influenced by external narratives rather than empirical risk assessments.

Similarly, the EWP's 2022-23 report on India asserts that *"In 2022, the Hindu nationalist-led government's systematic discrimination against the country's Muslim minority has continued to intensify amid mounting reports of violence—met with impunity—and efforts to restrict Muslim rights. Hindu nationalist leaders have continued to propagate hate speech, including religious leaders' calls for mass killings of Muslims in December 2021. Several states saw large-scale and violent incidents targeting Muslims in recent months, which involved Hindu nationalist processions engaging in derogatory anti-Muslim chants and the desecration of mosques. In response to these violent provocations, local authorities bulldozed Muslim-owned property across several states, which rights groups cited as an apparent attempt at collective punishment. Reports indicate continued abuses in the disputed Muslim-majority territory of Jammu and*

Kashmir, including increased targeting of Hindu civilians by militants and the Indian government's crackdown on journalists and human rights defenders. Other minorities and persecuted groups, including Christians and Dalits, continue to face violence and discrimination." Once again, none of these claims are outputs of the EWP's Statistical Risk Assessment Model; instead, they are extracted from various media sources, including *CFR*, *The Diplomat*, *The Indian Express*, *Al Jazeera*, *Washington Post*, *Vice*, *The Guardian*, *Article-14*, *Human Rights Watch*, *Amnesty*, and *The Atlantic*. This approach suggests that the EWP's reports do not adhere strictly to the model's empirical findings but rather incorporate and amplify perspectives from journalistic and advocacy sources. EWP report 2024-25 inexplicably did not discuss India despite ranking it 5th among of the countries. By embedding media narratives into its country reports, EWP's reports risk conflating opinion-based observations with data-driven conclusions. This practice compromises the objectivity expected of a statistical risk assessment and introduces a potential bias aligned with media discourse. Such practice, whereby EWP's reports selectively present information external to its statistical model, may mislead stakeholders, fostering perceptions not grounded in the model's empirical outputs.

Conclusion

The statistical inconsistencies and methodological flaws in the Early Warning Project (EWP) model significantly undermine its reliability and utility as a predictive tool. The drastic fluctuations in intercept values, coupled with unstable coefficients for key variables, expose foundational weaknesses in the model's design and implementation. The arbitrary inclusion and exclusion of variables, without a clear theoretical basis, exacerbate these issues, leading to logical inconsistencies and questionable outcomes. Furthermore, the reliance on static and spurious correlations, coupled with inadequate efforts to address multicollinearity and eliminate redundant variables, compromises the model's predictive accuracy and interpretability.

The Early Warning Project (EWP) model exhibits profound methodological weaknesses that compromise its reliability and practical utility as a predictive tool for assessing the risk of mass killings. Chief among these issues are drastic fluctuations in intercept values—ranging from a stable -8.8 to -11.2 between 2019 and 2022 to an inexplicable rise to 67.96 in 2023—and unstable coefficients for key variables. These changes undermine the model's internal consistency, suggesting a lack of robust theoretical or empirical grounding. Moreover, the frequent inclusion and exclusion of variables, often without adequate justification, introduces further inconsistencies, raising questions about the coherence of the model's design.

The model's performance metrics, notably a False Positive Rate (FPR) of 95.2% and a False Negative Rate (FNR) of 29-36%, highlight its limited predictive accuracy. The high FPR generates an overwhelming number of false alarms, diverting resources toward low-risk scenarios and contributing to "alert fatigue." Conversely, the high FNR signifies that nearly one-third of actual mass killing onsets remain undetected, undermining the model's core objective of early warning and prevention. These deficiencies are compounded by the use of static variables such as "Population Size" and "Geographical Region," which disproportionately influence risk rankings and trap certain countries in high-risk categories irrespective of real-world reforms.

The lack of transparency in variable selection and weighting, coupled with the reliance on correlational rather than causal relationships, further weakens the model's theoretical foundation. Parameters with overlapping meanings and high intercorrelations, as well as the introduction of spurious correlations, result in multicollinearity and inflated coefficients, distorting the interpretability of results. These flaws render the model vulnerable to overfitting and reduce its generalizability across diverse contexts.

Given these fundamental issues, attempting piecemeal solutions would only perpetuate the model's deficiencies. Instead, the EWP should consider starting afresh, establishing a robust framework grounded in strong theoretical principles and empirical evidence. Such a comprehensive reevaluation is essential for transforming the EWP model into a reliable and effective tool in the field of forecasting of mass killings.

References

1. Braun, R. (2016). Religious minorities and resistance to genocide: The collective rescue of Jews in the Netherlands during the Holocaust. The American Political Science Review, 110(1), 127-147. https://doi.org/10.1017/S0003055415000544
2. Cingranelli, D., & Filippov, M. (2018). Are human rights practices improving? The American Political Science Review, 112(4), 1083-1089. https://doi.org/10.1017/S0003055418000254
3. Fariss, C. J. (2014). Respect for human rights has improved over time: Modeling the changing standard of accountability. The American Political Science Review, 108(2), 297-318. https://doi.org/10.1017/S0003055414000070
4. Pepinsky, T. B., Goodman, S. W., & Ziller, C. (2024). Modeling spatial heterogeneity and historical persistence: Nazi concentration camps and contemporary intolerance. The American Political Science Review, 118(1), 519-528. https://doi.org/10.1017/S0003055423000072
5. Grasse, D. (2024). State terror and long-run development: The persistence of the Khmer Rouge. The American Political Science Review, 118(1), 195-212. https://doi.org/10.1017/S0003055423000382
6. Harff, B. (2017). No lessons learned from the Holocaust? Assessing risks of genocide and political mass murder since 1955. Genocide and Human Rights, 329–345. https://doi.org/10.4324/9781351157568-11
7. Moses, A. D. (2010). Toward a theory of critical genocide studies. Sciences Po Mass Violence and Resistance - Research Network. Retrieved from https://www.sciencespo.fr/mass-violence-war-massacre-resistance/en/document/toward-theory-critical-genocide-studies.html
8. Goldsmith, B. E., Butcher, C. R., Semenovich, D., & Sowmya, A. (2013). Forecasting the onset of genocide and politicide: Annual out-of-sample forecasts on a global dataset, 1988–2003. Journal of Peace Research, 50(4), 437-452. https://doi.org/10.1177/0022343313479789
9. Corb, N. (2024). "An island of Jewish autonomous life": Paul Rosner's diary and the story of the Young Maccabi Movement in Germany. Holocaust and Genocide Studies, 38(1), 81–101. https://doi.org/10.1093/hgs/dcae002
10. Bachman, J. S. (2019). A 'synchronised attack' on life: The Saudi-led coalition's 'hidden and holistic' genocide in Yemen and the shared responsibility of the US and UK. Third World Quarterly, 40(2), 298–316. https://doi.org/10.1080/01436597.2018.1539910
11. Baines, E. K. (2003). Body politics and the Rwandan crisis. Third World Quarterly, 24(3), 479–493. https://doi.org/10.1080/0143659032000084429
12. Udombana, N. J. (2007). Still playing dice with lives: Darfur and security council resolution 1706. Third World Quarterly, 28(1), 97–116. https://doi.org/10.1080/01436590601081906
13. Rosenberg, S. P. (2012). Genocide is a process, not an event. Genocide Studies and Prevention: An International Journal, 7(1), 16-23. https://doi.org/10.3138/gsp.7.1.16

14. Hendrickson, P. (2020). “Old” and “new” mass killing? Genocide and politicide occurrence and severity during and after the Cold War. Terrorism and Political Violence, 34(4), 856–876. https://doi.org/10.1080/09546553.2020.1745776
15. Kizilhan, J. I., Berger, T., Sennhauser, L., & Wenzel, T. (2023). The psychological impact of genocide on the Yazidis. Frontiers in Psychology, 14. https://doi.org/10.3389/fpsyg.2023.1074283
16. Kebranian, N. (2020). Genocide, history, and the law: Legal performativity and recognition of the Armenian genocide in France and Germany. Holocaust and Genocide Studies, 34(2), 253–273. https://doi.org/10.1093/hgs/dcaa027
17. Kamola, I. A. (2007). The global coffee economy and the production of genocide in Rwanda. Third World Quarterly, 28(3), 571–592. https://doi.org/10.1080/01436590701192975
18. Jones, A. (2024). Diffusing genocide studies, defusing genocides. Genocide Studies and Prevention, 18(2), 34-56.
19. Uzonyi, G. (2018). Interstate rivalry, genocide, and politicide. Journal of Peace Research, 55(4), 476–490. https://doi.org/10.1177/0022343317741186
20. Bachman, J. (2020). Cases studied in genocide studies and prevention and Journal of Genocide Research and implications for the field of genocide studies. Genocide Studies and Prevention: An International Journal, 14(1), 2-20. https://doi.org/10.5038/1911-9933.14.1.1706
21. Ahram, A. I. (2014). The role of state-sponsored militias in genocide. Terrorism and Political Violence, 26(3), 488–503. https://doi.org/10.1080/09546553.2012.734875
22. Gavriilidis, A. (2016). Lemkin's Greek friends: Abusing history, constructing genocide—and vice versa. Holocaust and Genocide Studies, 30(3), 488–505. https://doi.org/10.1093/hgs/dcw064
23. Travis, H. (2012). On the original understanding of the crime of genocide. Genocide Studies and Prevention: An International Journal, 7(1), 30-55. https://doi.org/10.3138/gsp.7.1.30
24. Straus, S. (2012). “Destroy them to save us”: Theories of genocide and the logics of political violence. Terrorism and Political Violence, 24(4), 544–560. https://doi.org/10.1080/09546553.2012.700611
25. Hossain, M. P. (2021). Stages of the Rohingya genocide: A theoretical and empirical study. Holocaust and Genocide Studies, 35(2), 211–234. https://doi.org/10.1093/hgs/dcab0333
26. Wilson, C. (2020). Nostalgia, entitlement and victimhood: The synergy of white genocide and misogyny. Terrorism and Political Violence, 34(8), 1810–1825. https://doi.org/10.1080/09546553.2020.1839428
27. Williams, T. (2016). More lessons learned from the Holocaust - Towards a complexity-embracing approach to why genocide occurs. Genocide Studies and Prevention: An International Journal, 9(3), 137-153. http://dx.doi.org/10.5038/1911-9933.9.3.1306
28. Nyseth Brehm, H. (2019). Moving beyond the state: An imperative for genocide prediction. Genocide Studies and Prevention: An International Journal, 13(3), 64-78. https://doi.org/10.5038/1911-9933.13.3.1673

29. Bowser, B. P., Word, C. O., & Shaw, K. (2021). Ongoing genocides and the need for healing: The cases of Native and African Americans. Genocide Studies and Prevention: An International Journal, 15(3), 83–99. https://doi.org/10.5038/1911-9933.15.3.1785
30. Levene, M. (1999). The Chittagong Hill Tracts: A case study in the political economy of “creeping” genocide. Third World Quarterly, 20(2), 339–369. https://doi.org/10.1080/01436599913794
31. Williams, T., & Neilsen, R. (2016). “They will rot the society, rot the party, and rot the army”*: Toxification as an ideology and genocidal strategy. Terrorism and Political Violence, 29(6), 1190–1214. https://doi.org/10.1080/09546553.2015.110402
32. Verdeja, E. (2016). Predicting genocide and mass atrocities. Genocide Studies and Prevention: An International Journal, 9(3), 13-32. http://dx.doi.org/10.5038/1911-9933.9.3.1314
33. Parent, G. (2016). Genocide denial: Perpetuating victimization and the cycle of violence in Bosnia and Herzegovina (BiH). Genocide Studies and Prevention: An International Journal, 10(2), 38-58. http://dx.doi.org/10.5038/1911-9933.10.2.1369
34. Uvin, P. (2001). Difficult choices in the new post-conflict agenda: the international community in Rwanda after the genocide. Third World Quarterly, 22(2), 177–189. https://doi.org/10.1080/01436590120043291
35. Sanfilippo, A., McGrath, L., & Whitney, P. (2011). Violent frames in action. Dynamics of Asymmetric Conflict, 4(2), 103–112. https://doi.org/10.1080/17467586.2011.627933
36. Uzonyi, G. (2019). Bureaucratic quality and the severity of genocide and politicide. Dynamics of Asymmetric Conflict, 13(2), 125–142. https://doi.org/10.1080/17467586.2019.1650387
37. Saucier, G., & Akers, L. (2018). Democidal thinking: Patterns in the mindset behind organized mass killing. Genocide Studies and Prevention: An International Journal, 12(1), 80-97. https://doi.org/10.5038/1911-9933.12.1.1546
38. Luck, E. C. (2018). Why the United Nations underperforms at preventing mass atrocities. Genocide Studies and Prevention: An International Journal, 11(3), 32-47. http://doi.org/10.5038/1911-9933.11.3.1516
39. Jones, A. (2011). Diffusing genocide studies, defusing genocides. Genocide Studies and Prevention: An International Journal, 6(3), 8. https://digitalcommons.usf.edu/gsp/vol6/iss3/8
40. McDoom, O. S. (2014). Predicting violence within genocide: A model of elite competition and ethnic segregation from Rwanda. Political Geography, 42, 34-45. https://doi.org/10.1016/j.polgeo.2014.05.006
41. Fortuin, E. (2022). Ukraine commits genocide on Russians: The term "genocide" in Russian propaganda. Russian Linguistics, 46(3), 313-347. https://doi.org/10.1007/s11185-022-09258-5
42. Soundararajan, M., Jaishankar, K., Bushell, M., Telford, L., & Treadwell, J. (2024). The victimization of Rohingyas in Myanmar and Bangladesh: Breaking the silence - postcolonial criminology, ethnography and genocide. British Journal of Criminology, 64(4), 881-895. https://doi.org/10.1093/bjc/azad069

43. Goyes, D. R., South, N., Abaibira, M. A., Baicué, P., Cuchimba, A., & Ñeñetofe, D. T. R. (2021). Genocide and ecocide in four Colombian indigenous communities: The erosion of a way of life and memory. British Journal of Criminology, 61(4), 965-984. https://doi.org/10.1093/bjc/azaa109
44. Campbell, B. (2015). Genocide as predation. International Journal of Law, Crime and Justice, 43(3), 310–325. https://doi.org/10.1016/j.ijlcj.2015.05.004
45. Smeulers, A., & Hoex, L. (2010). Studying the microdynamics of the Rwandan genocide. British Journal of Criminology, 50(3), 435-454. https://doi.org/10.1093/bjc/azq004
46. Bosdriesz, H., & Wirken, S. (2014). An Imperfect Success – The Guatemalan genocide trial and the struggle against impunity for international crimes. International Criminal Law Review, 14(6), 1067-1094. https://doi.org/10.1163/15718123-01406007
47. Campbell, B. (2015). Genocide as predation. International Journal of Law, Crime and Justice, 43(3), 310-325.
48. Kreß, C. (2006). The crime of genocide under international law. International Criminal Law Review, 6(4), 461-502.
49. Findlay, M. (2013). Enunciating genocide: crime, rights and the impact of judicial intervention. International Criminal Law Review, 13(1), 297-317.
50. Akhtar, Z. (2010). Canadian genocide and official culpability. International Criminal Law Review, 10(1), 111-135.
51. Lingaas, C. (2021). Dehumanising Ideology, metaphors, and psychological othering as evidence of genocidal intent. International Criminal Law Review, 22(5-6), 1044-1067.

Index of Tables

Table 1 - Illustration of EWP Model's Logical Inconsistencies

Table 2 - List of Genocide Predictors from Literature Review and Their Explanations

Table 3 - Comparison of Ulfelder Variables (used by EWP in 2014) and EWP Variables (2024)

Table 4 - Variables used by Ulfelder as well as EWP (2014) were removed from the list of EWP's variables.

Table 5 - List of Variables used by EWP in 2024, either not included in or modified from the variables list of Ulfelder (2014)

Table 6 - Scenarios Analyzed for Population Size as a Predictor in EWP

Table 7 - Civil Liberty Scores of Selected Countries in the EWP Model

Table 8 - Impact of Policy Actions on Country Rankings in the EWP Model

Table 9 - EWP's Modifications of Scales to V-Dem Data

Table 10 - Impact of Political and Civil Actions on Risk and Ranking in the EWP Model - India and China

Table 11 - Impact of Political and Civil Actions on Risk and Ranking in the EWP Model – China and North Korea

Table 12 - No Effect of Institutional and Social Reforms on Risk and Ranking of India in the EWP Model

Table 13 - Analysis of Social Group Power Distribution and Discrepancies in EWP Data

Centre for Development Policy & Management

An IIM Udaipur Initiative