

1 Introduction

The deaths and disappearances of migrants while on the move are a key concern across the world. There is a serious need to investigate the factors behind it properly. Since 2014, statistics indicate that at least 4,000 people have perished every year on migration routes, and 2022 was the worst year, with more than 8,500 deaths. [Council of Europe (2024)](Migration Data Portal, 2024) In particular, children need special attention, as this demographic is particularly vulnerable due to their age. (Global Migration Data Portal, 2024)

The present study begins by illustrating the temporal and spatial distribution of these events using polar diagrams and an animated globe. Following this, an econometric panel data analysis will be employed to provide more detailed insights, to inform policy interventions to prevent future occurrences. (Thorne, 2021)

1.1 Research question

What are the factors that influence the total number of deaths and missing children in migrant incidents across continents and time?

2 Methodology

The current research reports data was collected from the Missing Migrants Project (International Organization for Migration, 2024) for the period from 2014 to 2024, which resulted in 5,462 entries after some preprocessing. Preprocessing involved filtering for high-quality source information, and the selection of relevant variables, identified through literature, theoretical considerations and variables available in the dataset.

We explored the possibility of merging our dataset with additional sources to enrich our analysis. However, due to the specific nature of the topic, no relevant complementary data was found.

We transformed our variables by using dummy variables for the Cause of Death and logarithmic transformations for non-normal distributions considering skewness and kurtosis analyses, that demonstrated asymmetry and leptokurtic distribution. Linear relationships were incorporated into the model due to theoretical reasoning and literature on the topic.

For our analysis, we employed panel data regressions, analyzing Pooled OLS, Fixed Effects, and Random Effects models. Econometric tests (including Breusch-Pagan and Hausman tests) were conducted to determine which model was most appropriate for the data. This approach allows us to isolate key factors and provide insights to inform interventions aimed at preventing future incidents.

3 Initial Analysis

The polar charts (Figure 3.1) reveal notable temporal patterns. The number of deaths shows a general increase over the years, but a decrease in 2018 and 2019. The number of missing persons presents some oscillations, with peaks around 2016 and 2023. The number of documented incidents shows a steady increase until 2023. Finally, the survivors count show an overall increasing trend with a peak around 2016.

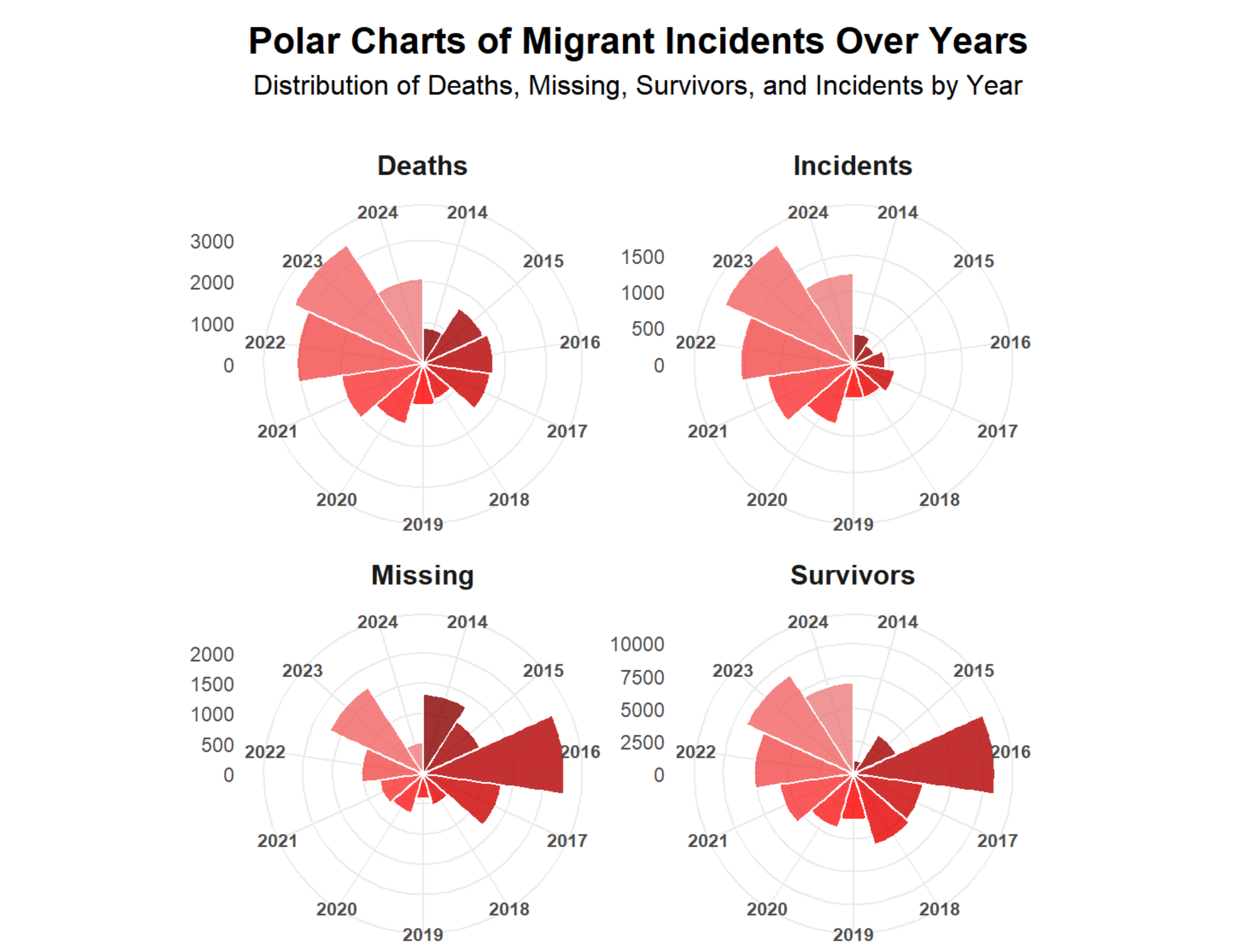


Figure 3.1: Polar Charts of Migrant Incidents Over Years. Distribution of Deaths, Missing, Survivors, and Incidents by Year. The animated map (Figure 3.2) effectively shows the geographical distribution of incidents over time. It highlights the Mediterranean Sea as a critical area for migrant deaths, with notable incidents occurring in Central America and along the borders between Mexico and USA.

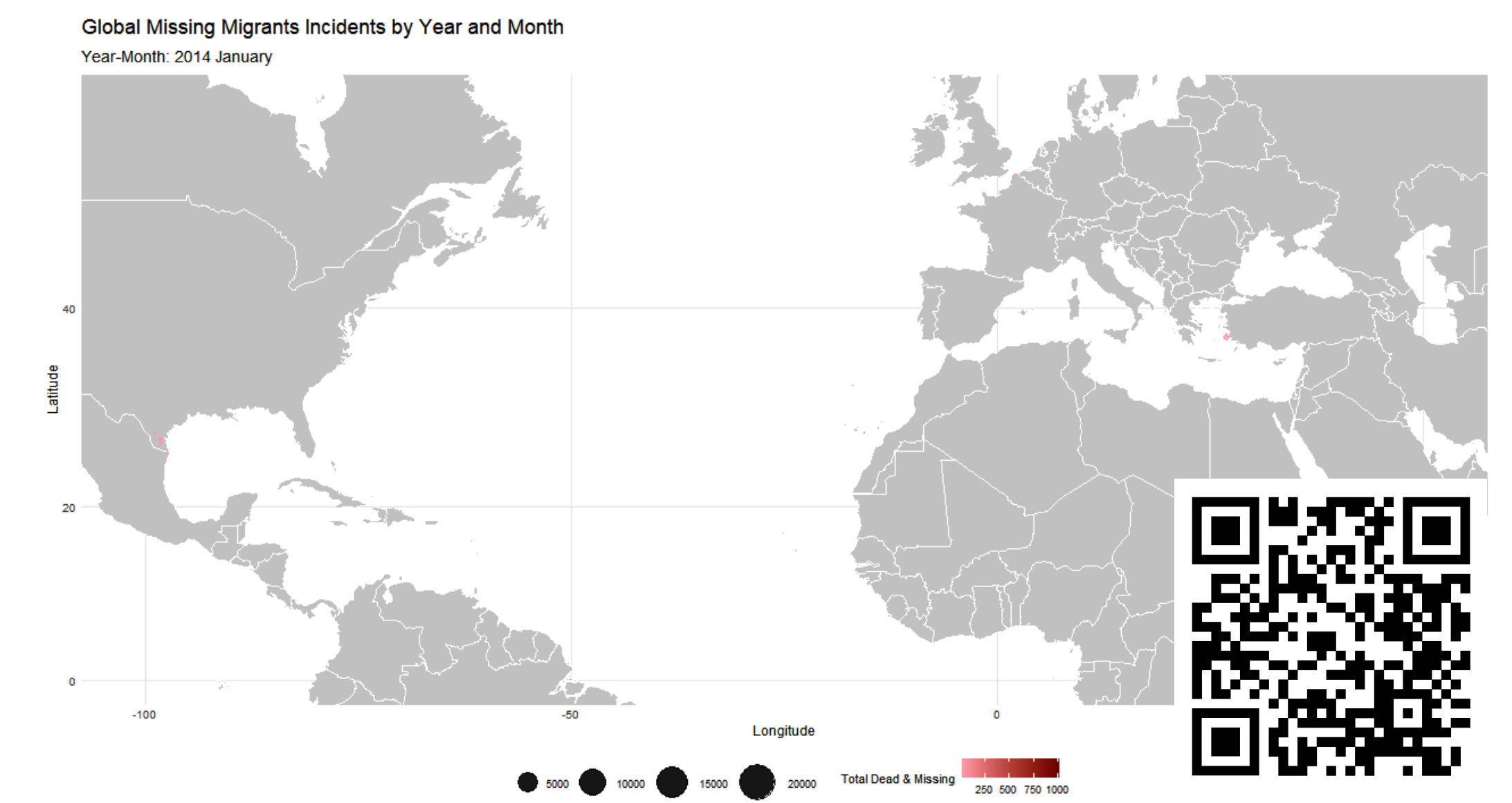


Figure 3.2: Animated Global Map of Missing Migrants Incidents by Year and Month.

4 Results

The panel data analysis investigates the factors influencing the number of deaths and missing children, controlling for continent and time. Table 4.1 presents the results of key statistical tests, and Table 4.2 displays the estimation results for Robust Random Effects.

Table 4.1: Results of Statistical Tests			
Test	p-value	H0	Conclusion
Breusch Pagan (POLS)	< 2.2e-16	Homosk.	Heterosk.
Breusch Pagan (FE)	< 2.2e-16	Homosk.	Heterosk.
Breusch Pagan (RE)	< 2.2e-16	Homosk.	Heterosk.
Robust Hausman	0.929	Random Effects	Random Effects
Lagrange Multiplier (Breusch Pagan)	< 2.2e-16	No Panel Effects	Panel Effects

The Breusch-Pagan tests, regardless of the type of model used, suggest the presence of heteroskedasticity (p -value $< \alpha = 5\%$). This gives statistical evidence that the errors are not homoscedastic, hence the need for the use of robust standard errors.

Lagrange Multiplier test (Breusch & Pagan, 1980) (p -value $< \alpha = 5\%$) supports that the Random Effects model is more appropriate than the Pooled OLS model.

For the Robust Hausman test performed, we obtained the p -value of 0.929. This result suggests that the Fixed Effects and Random Effects Models are consistent, but the only the latter is efficient.

Table 4.2: Random Effects Model With Robust Standard Errors				
	Estimate	Std. Error	t-value	Pr(> t)
(Intercept)	0.000	0.027	0.000	1.000
log(Number of Females)	0.275	0.062	4.428	0.000 ***
log(Number of Males)	0.069	0.026	2.650	0.008 **
log(Number of Survivors)	0.021	0.006	3.301	0.001 ***
Drowning	0.048	0.015	3.240	0.001 **

Table 4.2 presents the results from our Robust Random Effects model, which examines factors influencing child deaths or missing (D&M) in migrant incidents. Both log(Number of Females) and log(Number of Males) show positive coefficients, as an increase in overall D&M is expected to correlate with increased child D&M. However, the coefficient for log(Number of Females) is substantially larger (0.275) than that of log(Number of Males) (0.069), indicating a more pronounced effect. Specifically, a 1% increase in the number of female D&M is associated with a 0.275% increase in the number of children D&M. This suggests that women, often primary caregivers, may offer a higher degree of protection to children, and their deaths can subsequently elevate the risk for those children.

Additionally, the Drowning dummy variable has a positive coefficient of 0.048. This result indicates that, ceteris paribus, drowning incidents are associated with approximately a 4.92% $[(e^{0.048} - 1) \times 100]$ increase in the expected number of children D&M compared to other causes of death.

5 Conclusion

Our analysis highlights critical patterns in migrant deaths and disappearances, with high-risk areas including the Mediterranean, Central America, and the US-Mexico border. Peaks in incidents during 2016 and 2023, alongside rising survivor counts, suggest both increased risks and improved rescue efforts. However, regional data inconsistencies may obscure the full scope of the crisis.

The panel data analysis shows a strong link between female fatalities and higher child deaths, emphasizing the vulnerability of women and children. Drowning incidents further elevate these risks.

5.1 Next Steps

Future research should explore the influence of smuggling networks and conflicts in migrants' countries of origin, incorporating data on conflict levels (e.g., number of attacks or deaths in armed conflicts) and the percentage of unaccompanied minors among migrants.

A more detailed analysis of specific migration routes could help identify localized patterns, guiding tailored interventions to protect vulnerable groups, especially women and children. Improved data standardization across regions is also crucial for accurately assessing the crisis and informing effective policies.

References

Breusch, T. S., & Pagan, A. R. (1980). The lagrange multiplier test and its applications to model specification in econometrics. *Review of Econometric Studies*, 47, 239–253.

Council of Europe. (2024). *Missing migrants, refugees and asylum seekers – a call to clarify their fate*. <https://rm.coe.int/as-mig-2024-11-dra-report-missing-migrants-refugees-and-asylum-seeke/1680b090c3>

Global Migration Data Portal. (2024). *Child and young migrants*. <https://www.migrationdataportal.org/themes/child-and-young-migrants>

International Organization for Migration. (2024). *Missing migrants project*. <https://missingmigrants.iom.int/>

Migration Data Portal. (2024). *Migrant deaths and disappearances*. <https://www.migrationdataportal.org/themes/migrant-deaths-and-disappearances>

Thorne, B. (2021). *Posterdown: Poster presentations in r markdown*. <https://github.com/brenthorne/posterdown>