Economic Inequality and the Global Mental health Burden

Alona Sychevska, 2769408

June 5, 2025

Alona Sychevska

Tutorial lecturer's name: Chantal Schouwenaar, Jack Fitzgerald

1 Problem Motivation & Literature

Why It's a Recognized Social Problem

Mental health issues such as depression and anxiety are on the rise globally—not only in high-income countries like the Netherlands or the UK, but also across low- and middle-income countries (LMICs). According to the WHO, mental disorders are among the leading causes of disability worldwide, and their burden is projected to grow (World Health Organization, 2023).

A growing body of literature suggests that **economic inequality**, rather than just poverty or absolute income, is a key social determinant of mental health. Studies have found that countries with higher levels of income inequality tend to have higher rates of psychological distress, even after controlling for GDP per capita and other macroeconomic indicators (Wilkinson & Pickett, 2009).

A systematic review of 26 studies—primarily from high-income countries—found that greater income inequality is associated with a 19% increased risk of depression (Patel et al., 2018). While most of this literature focuses on wealthy countries, studies in LMICs have also confirmed that poverty and inequality significantly predict common mental disorders (Lund et al., 2010).

More recent cross-continental studies have demonstrated that income inequality not only exacerbates anxiety and depression in high-income countries, but also in emerging and developing economies, reinforcing the need for a more globally inclusive analysis (Patel et al., 2022).

Theoretical Framing & Mechanisms

Several mechanisms help explain how economic inequality may drive mental health problems:

- Social capital erosion: Inequality weakens social cohesion and trust, contributing to social exclusion and loneliness, which are known risk factors for mental illness (Kawachi & Berkman, 2000).
- Bidirectional poverty—mental health link: Mental illness can both result from and reinforce poverty, creating a feedback loop that is particularly severe in LMICs (Lund et al., 2010).

Gap in the Literature

While many studies link income inequality (often using the Gini index) to mental health, most focus on high-income countries and treat other economic indicators like GNI or GDP per capita as background controls rather than central variables.

Few studies explicitly examine Gini, GNI, and GDP per capita together to understand their combined impact on mental health—especially in low- and middle-income countries, where the mental health burden is growing fastest.

This project addresses that gap by analyzing how these core economic factors jointly relate to mental health outcomes across a diverse set of countries.

2 Data Sourcing & Description

Data Sources and Credibility

This project integrates datasets: the Global Burden of Disease (GBD) Study 2021 and a multi-sourced datasets compiled from the World Bank (WDI), OECD, and other internationally reputable institutions that includes socioeconomic variables.

The GBD Study is produced by the Institute for Health Metrics and Evaluation (IHME) at the University of Washington. It is considered the gold standard for global health statistics and is widely cited by academic institutions and policymakers worldwide (IHME, 2022).

The compilation of datasets drawing from sources like the World Bank's World Development Indicators (WDI), and World Population Review. These are recognized institutions known for robust methodologies and transparency. WDI, in particular, is frequently used in cross-country empirical studies due to its consistency and broad temporal coverage.

Metadata Overview

Go to ReadMe.txt

Complementarity of the Datasets

GBD offers high-quality health burden metrics (prevalence in numbers), crucial for understanding the impact of mental health disorders, while the additional datasets offer explanatory variables on inequality and economic performance. These datasets are complementary in structure and purpose: the GBD gives outcome measures, while the additional datasets give potential predictors. Their integration allows for robust modeling of social determinants of mental health at the national level.

While I initially collected more than two datasets, the final dataset has complementary strengths and ability to support the causality-oriented focus of this project: exploring how inequality and structural factors relate to mental health burdens across countries.

Relevance to the Topic

The project's focus is on the causal relationship between inequality and mental health in low-, middle-, and high-income countries. The selected datasets are ideal because: The GBD dataset provides mental health burden metrics (e.g., Prevalence of depression or anxiety), which are essential for measuring the societal impact of mental disorders across countries. The merged dataset allows for correlational and regression analyses of how inequality indicators (e.g., Gini index, top 10% wealth share), and economic performance relate to mental health prevalence. Datasets cover a wide range of countries, including those outside the high-income bracket, which aligns with the study's comparative angle.

Limitations of the Data

Despite their strengths, datasets have limitations: GBD data, though standardized, aggregate country-level estimates, potentially masking subnational disparities and cultural variation in diagnosis/reporting. The multi-sourced data suffers from missing data for certain indicators in low-income countries, which could bias regression results or reduce sample size. Differences in data collection years (some indicators are 2021, others 2022 or 2023) may introduce temporal misalignment. Mental health prevalence estimates from sources like World Population Review may lack the methodological rigor of epidemiological surveys.

2.1 Load in the data

2.2 Checking the structure

```
head(GBD %>% select(measure name, location name, cause name, metric name, year, val), 3)
## # A tibble: 3 x 6
##
    measure name location name cause name
                                                 metric name year
                                                                         val
##
     <chr>
                 <chr>
                                <chr>
                                                 <chr>
                                                             <dbl>
                                                                       <dbl>
## 1 Deaths
                 Israel
                                Mental disorders Number
                                                              2010 0.616
## 2 Deaths
                 Israel
                                Mental disorders Percent
                                                              2010 0.0000156
## 3 Deaths
                                Mental disorders Rate
                                                              2010 0.00786
                  Israel
```

2.3 Quantifying

Data cleaning

```
#new var: gini_bucket
economic_inequality_gini_index <- economic_inequality_gini_index %>%
  filter(Year >= 2012) %>%
  rename("Gini index" =
           "Gini coefficient (2017 prices) - Income or consumption consolidated") %>%
  select(-Code, -`1039568-annotations`)
# Creating gini buckets (new var)
gini_cats <- economic_inequality_gini_index %>%
  mutate(gini_bucket = case_when(
                                   ~ "Low",
    Gini_index < 0.25</pre>
    Gini_index >= 0.25 & Gini_index < 0.35 ~ "Moderate",</pre>
    Gini_index >= 0.35
                                   ~ "High",
    TRUE
                                 ~ NA_character_
 ))
```

To facilitate group-wise comparisons and enhance interpretability, we created a new categorical variable, gini_bucket, based on the continuous Gini index. Countries were grouped into three inequality categories: Low (< 0.25), Moderate (0.25-0.35), and High (> 0.35). This transformation allows for easier comparison across inequality levels, especially in visualizations and descriptive analyses. While some granularity is lost by binning continuous values, this approach improves clarity and supports the identification of broader patterns in mental health outcomes across different levels of income inequality.

```
# Select only the columns needed, no duplicates
gdp_data <- gdp_data %>%
  select(country, year, NY.GDP.PCAP.CD) %>%
  rename(gdp_per_capita = NY.GDP.PCAP.CD) %>%
  filter(year >= 2012)
```

```
gni_by_country_year <- gni_by_country_year %>%
    rename("2016" = "...31", "2017" = "...32", "2018" = "...33", "2019" = "...34") %>%
    pivot_longer(
        cols = matches("^\\d{4}$"),
        names_to = "year",
        values_to = "income_group"
) %>%
    mutate(year = as.integer(year)) %>%
    select(Country, year, income_group)
gni_by_country_year <- gni_by_country_year %>%
    mutate(income_group = ifelse(income_group == "..", NA, income_group)) %>%
    filter(year >= 2012)
```

I filtered all datasets to include only data from 2012 onward and removed unnecessary columns or duplicates. Variables were renamed for clarity, and formats were standardized (e.g., reshaping wide GNI data to long format, converting ".." to missing values). These steps ensured consistency across datasets and prepared the data for merging and analysis.

Creating new variables: pearcent and rate of depression by country

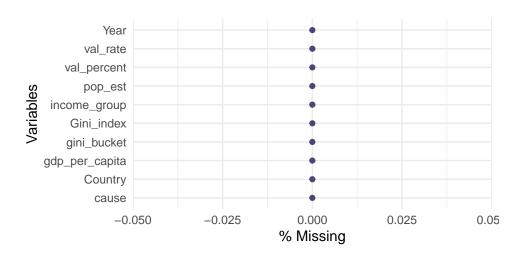
To standardize mental health prevalence data across countries with different population sizes, we created two new variables: val_rate, representing the number of cases per 100,000 population, and val_percent, indicating the proportion of the population affected by a disorder as a percentage. These transformations ensure comparability across countries and time periods, regardless of absolute population size. The per-100,000 rate (val_rate) is commonly used in public health reporting, while the percentage (val_percent) is helpful for interpreting the relative burden of disorders within a country's population.

2.4 Datasets merging

```
# Mental + gini_cats data
merged_data <- GBD %>%
  inner_join(gini_cats, by = c("Country" = "Entity", "Year" = "Year"))
```

To ensure consistency and minimize missing values, datasets were merged using inner_join. This method retains only rows with complete matches across key variables (Country and Year), which helps maintain data integrity for comparative analysis. While this approach may reduce the overall number of observations, it avoids issues introduced by incomplete records and ensures that all included data points are reliable and consistent across mental health, inequality, GDP, and income classification data.

gg_miss_var(merged_data, show_pct = TRUE)



2.5 Describe the type of variables included

```
describe_variables(merged_data)
```

```
## - Country: categorical (character), levels: China, Mongolia, Belarus, Kazakhstan, Armenia ...
## - cause: categorical (character), levels: Schizophrenia, Attention-deficit/hyperactivity disorder, I
## - Year: continuous (numeric), mean = 2016.37, sd = 2.82
## - val_rate: categorical (character), levels: 366.55, 366.01, 365.29, 364.93, 365.07 .
## - val_percent: continuous (numeric), mean = 2.78, sd = 4.25
## - pop_est: continuous (numeric), mean = 104628265.19, sd = 315184396.73
## - Gini_index: continuous (numeric), mean = 0.34, sd = 0.07
## - gini_bucket: categorical (character), levels: High, Moderate, Low
## - income_group: categorical (character), levels: UM, LM, H
```

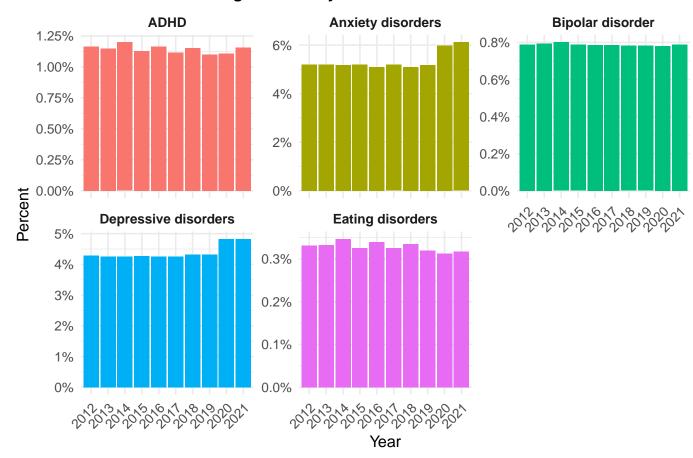
3.1 Visualizations

```
tendencies$cause <- gsub("Attention-deficit/hyperactivity disorder", "ADHD", tendencies$cause)
```

- gdp_per_capita: continuous (numeric), mean = 30698.76, sd = 24331.04

3.1.1 Temporal Variations

Mental Health Average Percent by Disorders Over Time



ca

The figure shows the average percentage prevalence of five major mental health disorders — Depressive

disorders, Anxiety disorders, Bipolar disorder, Eating disorders, and ADHD — over several years. Each disorder is presented in a separate facet for clearer comparison, with the y-axis scaled independently to reflect differences in prevalence magnitude.

Key observations include:

Depressive disorders consistently have the highest average prevalence, hovering between approximately 4% and 4.5%, with a slight upward trend over the years.

Anxiety disorders follow a similar pattern, also with mild fluctuations.

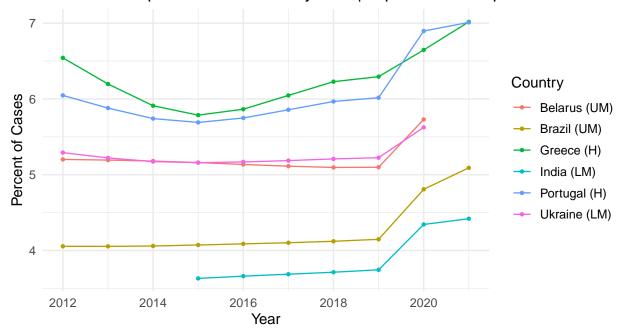
Bipolar disorder and ADHD exhibit much lower prevalence values (around 0.7% to 1%), remaining relatively stable over time without clear upward or downward trends.

Eating disorders show the lowest prevalence overall but display a slight increase in recent years.

Overall, the data suggest that while some disorders like depression and anxiety remain widespread and possibly increasing, others remain steady at lower prevalence levels. This temporal analysis highlights the importance of continuous monitoring to understand evolving mental health trends.

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

Percent of Depressive Disorders by Year | Top 2 Countries per Income Group



This line plot shows the percentage of depressive disorder cases over several years for the top two countries within each income group (Low, Lower-middle, Upper-middle, High). Each country is labeled along with its income group in the legend for clarity.

Key points:

High-income countries (e.g., Portugal, Greece) consistently report higher prevalence rates, generally between 5.6% and 7%, with some fluctuations.

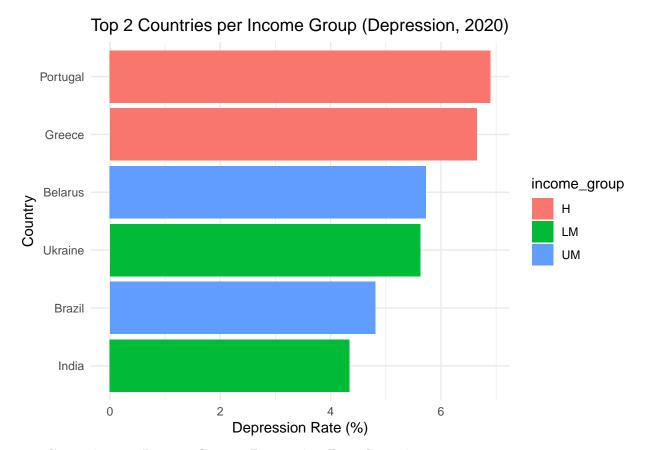
Upper-middle-income countries (e.g., Brazil) exhibit moderate prevalence around 4% to 5+, with mild variation over time, and a jump during the COVID time.

Lower-middle- and low-income countries (e.g., India) tend to have lower prevalence, though some show slight upward or downward trends.

The data reveal a clear pattern where higher-income countries report higher rates of depressive disorders, while lower-income groups show comparatively lower prevalence, emphasizing economic context's role in mental health.

This graph highlights important disparities in depressive disorder prevalence across economic groups, underscoring the need for income-sensitive mental health strategies.

```
dep_anx %>%
  filter(cause == "Depressive disorders", Year == 2020) %>%
  group_by(income_group) %>%
  slice_max(order_by = val_percent, n = 2, with_ties = FALSE) %>%
  ungroup() %>%
  ggplot(aes(x = reorder(Country, val_percent), y = val_percent, fill = income_group)) +
  geom_col() +
  coord_flip() +
  labs(
    title = "Top 2 Countries per Income Group (Depression, 2020)",
    x = "Country",
    y = "Depression Rate (%)"
  ) +
  theme_minimal()
```



Top Countries per Income Group: Depression Prevalence in 2020

To better understand the global distribution of mental health burdens, we analyzed the prevalence of depressive disorders across countries and income groups using 2020 data. The chart below displays the top two countries with the highest reported rates of depression within each World Bank income classification: High income (H), Upper-middle income (UM), and Lower-middle income (LM).

Methodology

The data was filtered from a larger dataset (dep_ann) to include only entries where the reported cause was "Depressive disorders" and the year was 2020. The data was grouped by income group, and the two countries with the highest depression prevalence percentages (val_percent) were selected within each group. A horizontal bar chart was produced using ggplot2 in R, with each income group represented by a distinct color. The depression rates were visualized to facilitate cross-group comparisons.

Findings

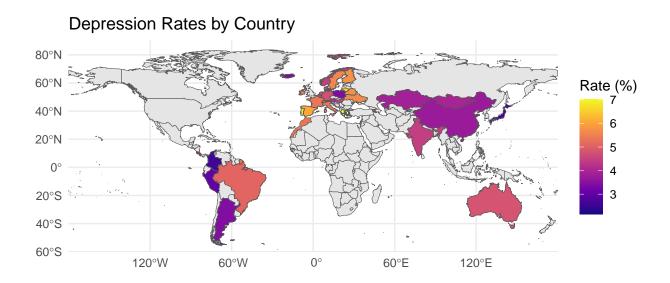
- High-Income Countries (H):
 - Portugal and Greece had the highest depression rates among high-income nations.
 - These findings may reflect the lingering socio-economic effects of the 2008–2012 European debt crisis, high youth unemployment, and austerity measures, particularly in Southern Europe.
- Upper-Middle-Income Countries (UM):
 - Belarus and Brazil were the most affected in this group.
 - In Brazil, rising urban violence, social inequality, and healthcare system strain could contribute to widespread mental health challenges.
 - Belarus may reflect mental health impacts of political instability, economic stress, and underreporting of health burdens in prior years.
- Lower-Middle-Income Countries (LM):
 - Ukraine and India had the highest reported rates.
 - Ukraine's high prevalence may be attributed to prolonged geopolitical conflict and economic insecurity, particularly following the 2014 crisis and ongoing tension in the region.
 - In India, the growing burden of depression may be associated with rapid urbanization, population density, and limited mental health infrastructure, though increased awareness and diagnostic reporting could also play a role.

This cross-income comparison reveals that depression is a global issue that affects both high-income and lower-income nations. Notably, the income level alone does not predict the prevalence of depressive disorders. Instead, country-specific socio-political factors, healthcare infrastructure, cultural stigma, and stress-related exposures appear to play a significant role.

The visualization helps highlight key target countries within each economic bracket for further investigation or policy prioritization. It underscores the need for context-sensitive mental health strategies and investment in mental health services regardless of national income level.

```
ggplot() +
  geom_sf(data = world, fill = "grey80", color = "white") +
  geom_sf(data = world_dep, aes(fill = val_percent)) +
  scale_fill_viridis_c(option = "plasma", na.value = "grey90") +
  coord_sf(xlim = c(-180, 180), ylim = c(-60, 90), expand = FALSE)+ # fixes centering
  theme_minimal() +
  labs(title = "Depression Rates by Country", fill = "Rate (%)")
```

3.1.2 Spatial Analysis (Maps)



Global Distribution of Depression Rates (2020)

To provide a spatial perspective on the global burden of depression, we created a choropleth map showing country-level depression prevalence for the year 2020. This visual representation helps identify regional trends and geographical disparities that might be missed in purely tabular or bar-chart formats.

Methodology The map was generated using the ggplot2 and sf packages in R. Depression data was filtered from the merged dataset (merged_data) to include only records where the cause was "Depressive disorders." This filtered dataset was joined with global shapefile data (world) using country names as the linking key. The val_percent variable, representing the percentage of the population affected by depression in 2020, was used to shade each country. The viridis color scale (option = "plasma") was applied to enhance visual contrast and ensure accessibility for color-blind users. Countries without data were shaded grey.

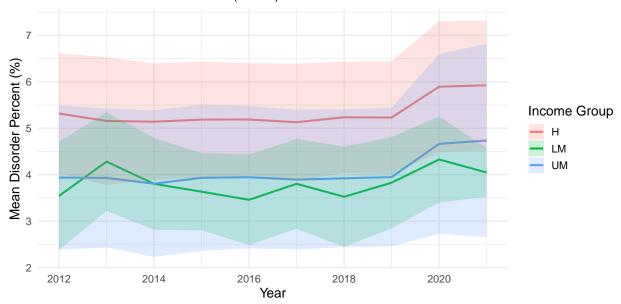
Findings * The highest prevalence rates (6–7%) appear concentrated in parts of Southern and Eastern Europe, including Portugal, Greece, and Ukraine. These regions may be affected by economic instability, post-conflict stressors, and limited access to mental health care. * South Asia (e.g., India) and parts of South America (e.g., Brazil) also show moderate to high depression rates. These findings suggest that mental health challenges are not confined to affluent countries but are also significant in emerging economies. * Lower rates (3–4%) were observed in much of Sub-Saharan Africa, Southeast Asia, and Scandinavia. However, these lower figures may reflect data limitations, underdiagnosis, or cultural stigma, rather than genuinely lower mental health burdens. * Data gaps are visible in some regions (e.g., parts of Africa and Central Asia), shown in grey. This highlights the need for improved data collection and reporting on mental health in low-resource settings.

Interpretation and Implications This spatial analysis emphasizes that depression is a global health issue with uneven geographic distribution. While higher-income countries tend to report more diagnoses—possibly due to better awareness and mental health infrastructure—countries with fewer resources may underreport due to stigma or systemic barriers. The map suggests that mental health policies must be adapted to regional needs, taking into account local economic, cultural, and healthcare contexts. It also underscores the necessity for more comprehensive global surveillance of mental health disorders to inform international health strategies and funding priorities.

3.1.3 Sub-population analysis

Depression & Anxiety Over Time by Income Group

With Standard Deviation Ribbon (±1 SD)



This graph depicts the average prevalence of depressive and anxiety disorders over time, broken down by income groups (Low, Lower-Middle, Upper-Middle, High). The colored lines show mean disorder rates, while the shaded ribbons represent ± 1 standard deviation, highlighting variability within each group.

Key observations:

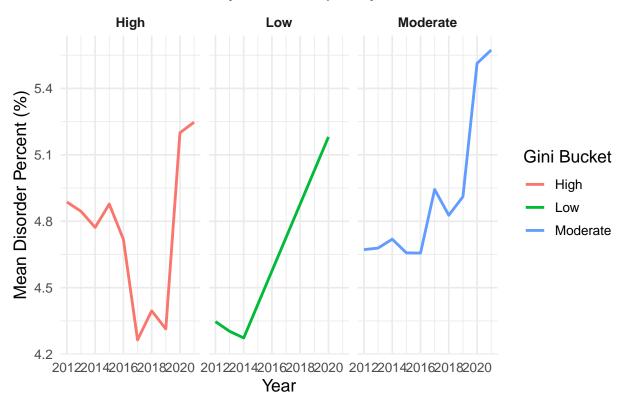
- High-income countries consistently show higher average disorder rates compared to lower-income groups.
- All income groups exhibit a gradual upward trend over the years.

The spread (ribbon width) indicates some variability within income groups but a generally steady increase across the board. This suggests that income level is associated with differences in mental health disorder prevalence, and the gradual increase points to a growing mental health challenge globally.

```
dep_anx %>%
  filter(!is.na(gini_bucket)) %>%
  filter(cause %in% c("Depressive disorders", "Anxiety disorders")) %>%
  group_by(Year, gini_bucket) %>%
  summarise(mean_disorder = mean(val_percent, na.rm = TRUE), .groups = "drop") %>%
  ggplot(aes(x = Year, y = mean_disorder, color = gini_bucket)) +
  geom_line(size = 1) +
```

```
facet_wrap(~ gini_bucket) +
labs(
   title = "Disorder Trends by Gini Inequality Level",
   x = "Year",
   y = "Mean Disorder Percent (%)",
   color = "Gini Bucket"
) +
theme_minimal(base_size = 13) +
theme(strip.text = element_text(face = "bold"))
```

Disorder Trends by Gini Inequality Level



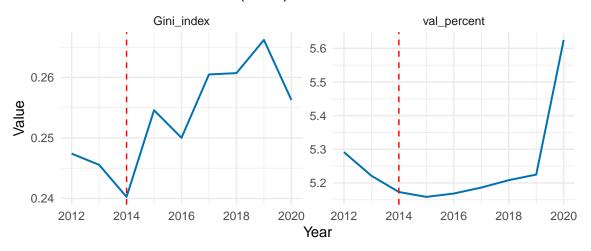
The graph shows the average prevalence of depressive and anxiety disorders from 2012 to 2021, categorized by low, moderate, and high income inequality (Gini index).

Countries with high inequality consistently exhibit higher disorder rates, while those with low inequality maintain lower rates. All groups show an upward trend over time, with a widening gap suggesting that greater economic inequality is linked to increased mental health burdens.

3.2 Event analysis

```
ggplot(ukraine_long, aes(x = Year, y = value)) +
  geom_line(color = "#0072B2", size = 0.7) +
  geom_vline(xintercept = 2014, linetype = "dashed", color = "red") +
  facet_wrap(~ variable, scales = "free_y") +
  labs(
```

Depressive Disorder Percent vs Inequality Trends in Ukraine Since Conflict (2014)



Event Analysis: Depression and Inequality in Ukraine since the 2014 Conflict

To examine the potential societal impacts of geopolitical conflict, this analysis explores Ukraine's trends in income inequality and depression prevalence, focusing on the years surrounding the 2014 conflict and subsequent instability.

Data and Visualization

This panel plot uses data from Ukraine between 2012 and 2020. Two key indicators were extracted:

- Gini index (a measure of income inequality)
- Prevalence of depressive disorders (val_percent) Data were transformed into long format to allow side-by-side trend visualization, with each variable plotted in its own facet. A vertical red dashed line marks the year 2014, used here as a start of conflict reference point.

Findings

• Income Inequality (Gini Index):

The Gini index shows a noticeable increase post-2014, rising from ~0.24 to more than ~0.26. While this may appear modest, the trend is directionally consistent with economic dislocation caused by political instability, currency devaluation, internal displacement, and loss of key economic regions (e.g., Donbas, Crimea).

• Depression Prevalence:

Depression rates were stable or slightly declining up to 2014 but began to rise significantly after 2015, reaching a local peak by 2020. This pattern is consistent with trauma-related mental health effects, economic insecurity, and reduced access to healthcare—all of which are documented consequences of conflict.

Interpretation and Context

The diverging post-2014 trajectories in both indicators—rising inequality and rising depression—support the hypothesis that armed conflict and socio-political shocks can exacerbate both economic and psychological distress. Notably, the delayed increase in depression suggests a lagged effect, where the full mental health consequences of conflict take time to materialize and be captured in surveys or health systems.

These findings align with studies showing that prolonged crises compound vulnerability, especially in health and economic systems already under strain. The data also emphasize the need for integrated policy responses—targeting both economic recovery and mental health support—in post-conflict contexts.

4 Discussion

4.1 Discuss your findings

Analysis of temporal, spatial, and subgroup patterns reveals significant disparities in mental health disorder prevalence across time, income levels, and geography.

Temporal Trends: Depression and anxiety consistently show the highest global burden, with slight upward trends over time. While ADHD, bipolar, and eating disorders remain less prevalent, their relative stability may mask localized spikes or diagnostic gaps.

Cross-Country and Income Comparisons: High-income countries like Portugal and Greece exhibit the highest rates of depression, while low- and lower-middle-income countries show lower averages. This could reflect better diagnosis and awareness in wealthier nations, or underreporting and stigma in resource-limited settings. Notably, countries such as Ukraine and India show unexpectedly high rates, likely driven by socio-political instability and rapid societal change.

Spatial Patterns: Choropleth maps highlight regional clusters—Southern and Eastern Europe, parts of South Asia, and South America—with above-average depression rates. In contrast, lower figures in Sub-Saharan Africa and Southeast Asia may indicate underdiagnosis rather than lower burden.

Subpopulation Insights: Income group analysis confirms that wealthier nations report higher average disorder rates, though all groups show a gradual increase over time. Similarly, a positive association is visible between inequality (Gini index) and disorder prevalence, reinforcing the connection between socioeconomic disparities and mental health outcomes.

5 Reproducibility

5.1 Github repository link

https://github.com/qgelena/Quantifying a social problem/tree/main

5.2 Reference list

- Kawachi, I., & Berkman, L. F. (2000). Social cohesion, social capital, and health. Social Epidemiology, 174–190.
- Lund, C., Breen, A., Flisher, A. J., Kakuma, R., Corrigall, J., Joska, J. A., ... & Patel, V. (2010). Poverty and common mental disorders in low and middle income countries: A systematic review. Social Science & Medicine, 71(3), 517–528. https://doi.org/10.1016/j.socscimed.2010.04.027
- Patel, V., Burns, J. K., Dhingra, M., Tarver, L., Kohrt, B. A., & Lund, C. (2018). Income inequality and depression: A systematic review and meta-analysis of the association and a scoping review of mechanisms. World Psychiatry, 17(1), 76–89. https://doi.org/10.1002/wps.20492

- Patel, V., Saxena, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., & UnÜtzer, J. (2022). The Lancet Commission on global mental health and sustainable development. *The Lancet*, 392(10157), 1553–1598. https://doi.org/10.1016/S0140-6736%2818%2931612-X
- Smith, K. E., Bambra, C., Hill, S. E., & Watt, R. G. (2012). Health inequalities and the social determinants of health: What works? *Journal of Public Health*, 34(4), 523–529. https://doi.org/10.1093/pubmed/fds052
- Wilkinson, R., & Pickett, K. (2009). The Spirit Level: Why more equal societies almost always do better. London: Allen Lane.
- World Health Organization. (2023). Mental health. https://www.who.int/health-topics/mental-health