PLEASE WRITE A TITLETitle of Your Project

Saman De Silva, Hans Elasri, Claire Yang

Introduction

China's Zero-COVID strategy relied on aggressive non-pharmaceutical interventions, including lockdowns, mass testing, mobility restrictions, and community-level quarantines that varied in intensity across provinces and over time. The policies imposed substantial and uneven constraints on daily activity, social interaction, and employment providing a natural setting to study how large-scale and persistent containment measures may have affected mental health.

Assessing the psychological consequences of Zero-COVID is complicated by the simultaneous presence of COVID-19 morbidity and mortality, which may independently influence mental health. In addition, individuals differ in baseline levels of depression and vulnerability to stress, making cross-sectional comparisons difficult to interpret. Longitudinal data are therefore critical for isolating changes in mental health associated with evolving local conditions rather than stable individual differences.

In this project, we combine longitudinal data from the China Family Panel Studies (CFPS) with province-level measures of COVID-19 policy stringency from the Oxford COVID-19 Government Response Tracker and COVID-19 mortality data. Depression is measured using the CESD-20 scale. Our primary research question is whether increases in policy stringency and COVID-19 deaths are associated with changes in depressive symptoms.

We study this question using multilevel longitudinal models that exploit within-individual and across-province variation over time while accounting for the hierarchical structure of the CFPS sample. This approach allows us to estimate how changes in province policy stringency and pandemic intensity relate to changes in depression, net of persistent individual characteristics and common time trends.

Data

China Family Panel Studies (CFPS)

Individual-level mental health and demographic data come from the China Family Panel Studies (CFPS), a nationally representative, biennial longitudinal survey administered by the Institute of Social Science Survey at Peking University. The CFPS tracks individuals, families, and communities across China and collects detailed information on health, economic conditions, education, and family structure.

We use four survey waves: 2016, 2018, 2020, and 2022. After cleaning and merging with province-level pandemic data, the final analytical sample consists of 51,224 person-wave observations covering 12,806 individuals, 7,770 families, and 3,318 communities across 33 provinces. Depression is measured using the CESD-20 scale, which captures the frequency of depressive symptoms over the past week.

The longitudinal structure of the CFPS allows us to observe changes in depressive symptoms within the same individuals before and during the COVID-19 period.

COVID-19 Mortality Data

Data on COVID-19 cases and deaths come from the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University, sourced from official Chinese reporting channels including the National Health Commission of China and the China CDC. The dataset provides daily province-level counts of confirmed cases and deaths. We aggregate these data to the monthly level to match the temporal resolution of the survey data.

Policy Stringency Data

Measures of COVID-19 containment policy severity are drawn from the Oxford COVID-19 Government Response Tracker (OxCGRT). The OxCGRT provides systematic, cross-province indicators of government responses, including stay-at-home requirements, restrictions on gatherings, workplace closures, and travel controls. We use the Oxford Stringency Index, which summarizes the intensity of these policies on a standardized scale (Hale et al., 2021). Monthly averages of the stringency index are constructed to align with the CFPS survey timing.

Data Construction

All datasets are merged at the province–month level and aligned to individual CFPS interview months COVID-19 cases, deaths, and stringency measures are aggregated to monthly frequencies since we do not have the exact date of the interview. The final analytical dataset spans 2016–2022 and links individual-level mental health outcomes to time-varying local pandemic conditions and policy responses.

Methods

To study how changes in COVID-19 policy stringency and pandemic severity relate to changes in depressive symptoms, we use multilevel longitudinal models that leverage repeated observations of the same individuals over time. This approach allows us to focus on within-person changes in mental health as local pandemic conditions evolve, rather than relying on cross-sectional comparisons that may be confounded by stable individual differences.

The CFPS data have a hierarchical structure, with individuals nested within families, communities, and provinces. Observations within these units are likely correlated due to shared environments, social conditions, and institutional factors. Multilevel modeling provides a principled way to account for this dependence by including random effects at each relevant level, avoiding underestimated uncertainty and biased inference.

We include random intercepts for individuals to account for correlation across waves within the same person. In addition, we cluster at the family and community levels to capture shared household conditions and local-level contextual factors that may shape mental health outcomes. We include province-level random intercepts to absorb persistent differences across provinces, such as baseline mental health environments, health system capacity, and reporting practices. We prefer community-level clustering to county-level clustering because communities are smaller and more socially cohesive units, where correlated mental health outcomes are more plausible, while counties are comparatively large and heterogeneous.

Survey wave fixed effects are included to control for common time shocks affecting all respondents, such as nationwide changes in economic conditions, information environments, the effect of COVID overall, and changing mental health resources over time. This ensures that identification comes from variation in local pandemic conditions relative to broader national trends.

Also, we include a small set of individual-level controls—age, self rated health, gender, education, marital status, and urban residence. These controls were chosen because they are well-established correlates of depression. Together, these modeling choices allow us to isolate the association between COVID-19 policy stringency, pandemic severity, and changes in depressive symptoms while accounting for the complex longitudinal and hierarchical structure of the CFPS data.

Our key treatment variables—policy stringency and COVID-19 deaths—are measured as three-month moving averages. We use moving averages for two reasons. First, both policy stringency and reported mortality can exhibit short-term volatility due to reporting noise and temporary policy adjustments; smoothing reduces sensitivity to small, transitory fluctuations that are unlikely to reflect meaningful changes in individual mental health. Second, we expect the psychological effects of prolonged containment measures and sustained pandemic severity to accumulate over time rather than operate instantaneously. A rolling average therefore better reflects sustained exposure to restrictive policies and local pandemic conditions than a single-month measure.

Model

Our data have survey wave w nested within person i, who belongs to family f in community c in province p. Let y_{wifcp} denote the CESD-20 depression score for individual i in family f, community c, province p at wave w. Our main province-by-time treatments are the three–month moving averages of the Oxford policy stringency index and COVID-19 deaths per capita, denoted Stringency_{pw} and Deaths_{pw}, respectively.

We specify the following five-level random-intercept model:

$$\text{Level 1 (wave within person):} \quad y_{wifcp} = \pi_{ifcp} + \beta_1 \, \text{Stringency}_{pw} + \beta_2 \, \text{Deaths}_{pw}$$
$$+ \, \boldsymbol{\beta}_Z^\top \mathbf{Z}_{wifcp} + \boldsymbol{\lambda}^\top \mathbf{W}_w + \varepsilon_{wifcp},$$
$$\varepsilon_{wifcp} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$$

$$\text{Level 2 (person):} \quad \pi_{ifcp} = \beta_{fcp} + \boldsymbol{\beta}_X^\top \mathbf{X}_{ifcp} + u_i,$$
$$u_i \sim \mathcal{N}(0, \sigma_i^2)$$

$$\text{Level 3 (family):} \quad \beta_{fcp} = \gamma_{cp} + u_f,$$
$$u_f \sim \mathcal{N}(0, \sigma_f^2)$$

$$\text{Level 4 (community):} \quad \gamma_{cp} = \alpha_p + u_c,$$
$$u_c \sim \mathcal{N}(0, \sigma_c^2)$$

$$\text{Level 5 (province):} \quad \alpha_p = \nu + u_p,$$
$$u_p \sim \mathcal{N}(0, \sigma_p^2),$$

where ν is the grand mean depression level.

At Level 1, $\mathbf{Z}_{wifcp}$ collects time-varying individual covariates (centered age and self-rated health), and $\mathbf{W}_w$ is a vector of survey-wave indicators (2016, 2018, 2020, 2022) that flexibly absorbs common

shocks over time. The coefficients β_1 and β_2 capture the average associations between province-level policy stringency, COVID-19 deaths, and depression, conditional on covariates and random intercepts.

At Level 2, $\mathbf{X}_{ifcp}$ contains time-invariant or slow-moving individual characteristics (gender, education, marital status, and urban residence), and u_i is a person-specific random intercept. The terms u_f, u_c, and u_p are random intercepts for families, communities, and provinces, respectively, allowing for clustering of depression outcomes within each higher level.

Results

Table 1 reports estimates from our baseline multilevel model relating province-level COVID-19 policy stringency and COVID-19 deaths per capita to individual depressive symptoms. After adjusting for survey wave fixed effects, individual-level controls, and random intercepts at the person, family, community, and province levels, neither policy stringency nor COVID-19 mortality is significantly associated with CESD-20 scores. The estimated coefficients on both treatment variables are small in magnitude and statistically indistinguishable from zero.

Table 1: Full Model

Term	Estimate	Std. Error	df	t-value	p-value
(Intercept)	14.01	0.33	265.8	43.09	<0.001
X3month_avg_StringencyIndex_Average	-0.00003	0.01	22080.0	-0.01	0.996
X3month_avg_new_deceased_per_capita	-2.10	1.34	24340.0	-1.56	0.119
wave2	0.73	0.09	23180.0	7.88	<0.001
wave3	1.13	0.32	22310.0	3.59	<0.001
wave4	1.81	0.37	22540.0	4.87	<0.001
age_c	-0.04	0.01	9762.0	-6.66	<0.001
health_c	1.46	0.04	30280.0	38.83	<0.001
urban1	-0.53	0.15	1624.5	-3.49	<0.001
gender1	-1.16	0.12	5396.8	-9.93	<0.001
education3	-1.07	0.21	10100.0	-5.21	<0.001
education4	-1.85	0.20	9737.0	-9.04	<0.001
education5	-2.21	0.23	10847.0	-9.69	<0.001
education6	-2.35	0.28	13363.0	-8.29	<0.001
education7	-2.60	0.33	14715.0	-7.99	<0.001
education8	-2.90	0.98	21296.0	-2.97	0.003
education9	-0.38	2.96	28693.0	-0.13	0.898
education10	0.24	0.36	28889.0	0.66	0.512
marital2	0.13	0.22	13854.0	0.61	0.542
marital3	0.92	0.74	29426.0	1.24	0.214
marital4	2.85	0.39	20547.0	7.26	<0.001
marital5	3.69	0.35	16864.0	10.53	<0.001

By contrast, survey wave indicators are strongly associated with depressive symptoms. Relative to the pre-pandemic 2016 baseline wave, depression scores significantly increase monotonically across

subsequent waves, with particularly large increases observed in the 2020 and 2022 survey waves. These patterns are consistent across all model specifications we consider.

To assess the robustness of the baseline results, we estimate several extensions of the model. First, we allow the effect of policy stringency to differ in the final survey wave by interacting stringency with a wave-4 indicator. This interaction term is not statistically significant, and the main effect of stringency remains close to zero. Second, we remove survey wave fixed effects; while the stringency coefficient becomes statistically significant in this specification, the estimate is no longer robust once wave indicators are reintroduced, indicating that this relationship is driven by aggregate time trends rather than within-wave variation. Third, we restrict identification to within-province variation in stringency after 2020 by demeaning stringency at the province level; the resulting estimate remains statistically insignificant. Finally, we allow the effect of stringency to vary across provinces by including a random slope for stringency at the province level. The average stringency effect remains statistically insignificant, and estimated cross-province heterogeneity in the stringency slope is limited.

Figure 1 illustrates mean CESD-20 scores by survey wave. Depression levels increase sharply between the pre-pandemic and post-pandemic waves, highlighting a substantial rise in reported depressive symptoms over time. Consistent with the regression results, this figure underscores that observed changes in depression are primarily associated with broad temporal shifts rather than differential exposure to local pandemic policies or mortality intensity.

```r
## figure 1: mean CESD-20 by survey wave after controls

library(dplyr)
library(ggplot2)

## Load data
cfps <- read.csv("data/model_ready_data.csv",
                 stringsAsFactors = FALSE)

## Keep analysis sample consistent with models
cfps_fig <- cfps %>%
  filter(
    !is.na(cesd20),
    !is.na(wave)
  ) %>%
  mutate(
    wave = factor(
      wave,
      levels = c(1, 2, 3, 4),
      labels = c("2016", "2018", "2020", "2022")
    )
  )

## Compute means and standard errors by wave
wave_summary <- cfps_fig %>%
  group_by(wave) %>%
  summarise(
```

```r
    mean_cesd = mean(cesd20, na.rm = TRUE),
    se_cesd   = sd(cesd20, na.rm = TRUE) / sqrt(n()),
    n         = n(),
    .groups = "drop"
  ) %>%
  mutate(
    ci_lower = mean_cesd - 1.96 * se_cesd,
    ci_upper = mean_cesd + 1.96 * se_cesd
  )

## Plot
ggplot(wave_summary, aes(x = wave, y = mean_cesd, group = 1)) +
  geom_line(color = "black", linewidth = 0.8) +
  geom_point(size = 3, shape = 21, fill = "white", color = "black") +
  geom_errorbar(
    aes(ymin = ci_lower, ymax = ci_upper),
    width = 0.15,
    linewidth = 0.6
  ) +
  labs(
    title = "Mean Depressive Symptoms by Survey Wave After Controls",
    x = "Survey Wave",
    y = "Mean CESD-20 Score"
  ) +
  theme_minimal(base_size = 12) +
  theme(
    plot.title = element_text(face = "bold"),
    panel.grid.minor = element_blank()
  )
```

Mean Depressive Symptoms by Survey Wave After Controls
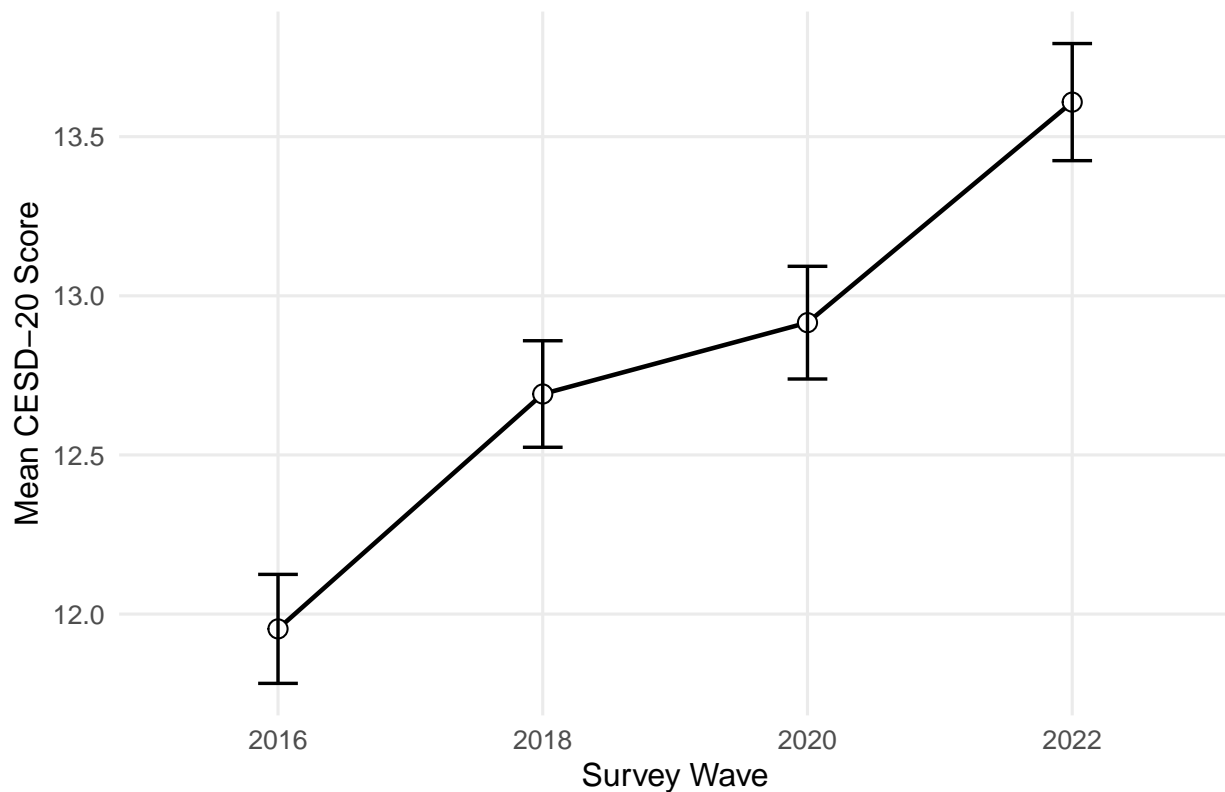
Taken together, the results indicate that, conditional on controls and fixed effects, we do not find evidence of a statistically significant association between province-level COVID-19 policy stringency or COVID-19 mortality and individual depressive symptoms. We defer discussion of potential explanations and implications of these findings to the conclusion.

```r
## ------------------------------------------------
## Figure 2: Actual vs Predicted Individual Trajectories
## ------------------------------------------------

library(dplyr)
library(ggplot2)
library(lme4)
library(tidyr)

set.seed(151)

## Reload data
cfps <- read.csv("data/model_ready_data.csv",
                 stringsAsFactors = FALSE)

pre_covid_zero_vars <- c(
  "X3month_avg_StringencyIndex_Average",
  "X3month_avg_new_deceased_per_capita"
)
```

```r
pre_covid_zero_vars <- intersect(pre_covid_zero_vars, names(cfps))

cfps <- cfps %>%
  mutate(
    across(
      all_of(pre_covid_zero_vars),
      ~ ifelse(wave %in% c(1, 2) & is.na(.), 0, .)
    )
  )


## Match model preprocessing
cfps <- cfps %>%
  mutate(
    pid   = factor(pid),
    fid   = factor(fid),
    cid   = factor(cid),
    provcd = factor(provcd),
    wave  = factor(wave),
    age_c   = scale(age, center = TRUE, scale = FALSE)[, 1],
    health_c = scale(health, center = TRUE, scale = FALSE)[, 1],
    gender   = factor(gender),
    education = factor(education),
    marital  = factor(marital),
    urban    = factor(urban)
  ) %>%
  filter(
    !is.na(cesd20),
    !is.na(X3month_avg_StringencyIndex_Average),
    !is.na(X3month_avg_new_deceased_per_capita),
    !is.na(wave)
  )



## Fit baseline model
model_full <- lmer(
  cesd20 ~
    X3month_avg_StringencyIndex_Average +
    X3month_avg_new_deceased_per_capita +
    wave +
    age_c +
    health_c +
    urban +
    gender +
    education +
    marital +
```

```r
    (1 | provcd) +
    (1 | cid) +
    (1 | fid) +
    (1 | pid),
  data = cfps,
  REML = TRUE
)


## -------------------------------------------------
## Sample individuals directly
## -------------------------------------------------

sample_ids <- sample(unique(cfps$pid), size = 20)

plot_data <- cfps %>%
  filter(pid %in% sample_ids) %>%
  mutate(
    wave_f = factor(
      wave,
      levels = c(1, 2, 3, 4),
      labels = c("2016", "2018", "2020", "2022")
    )
  )

plot_data <- plot_data %>%
  filter(
    !is.na(pid),
    !is.na(fid),
    !is.na(cid),
    !is.na(provcd)
  )

## Conditional predictions (include random intercepts)
plot_data$predicted <- predict(
  model_full,
  newdata = plot_data,
  re.form = NULL
)



## Long format
plot_long <- plot_data %>%
  select(pid, wave_f, cesd20, predicted) %>%
  pivot_longer(
    cols = c(cesd20, predicted),
    names_to = "series",
```

```r
    values_to = "value"
  ) %>%
  mutate(
    series = recode(series,
                    cesd20 = "Actual",
                    predicted = "Predicted")
  )

plot_long <- plot_long %>%
  filter(!is.na(value))

## Plot
ggplot(plot_long,
       aes(x = wave_f, y = value,
           group = series,
           linetype = series,
           color = series)) +
  geom_line(linewidth = 0.7) +
  geom_point(size = 1.6) +
  facet_wrap(~ pid, nrow = 4, ncol = 5) +
  scale_linetype_manual(
    values = c("Actual" = "solid",
               "Predicted" = "dashed")
  ) +
  scale_color_manual(
    values = c("Actual" = "red",
               "Predicted" = "blue")
  ) +
  labs(
    title = "Actual vs. Predicted Depression Trajectories",
    subtitle = "Random sample of 20 individuals",
    x = "Survey Wave",
    y = "CESD-20 Score",
    color = "",
    linetype = ""
  ) +
  theme_minimal(base_size = 11) +
  theme(
    plot.title = element_text(face = "bold"),
    axis.text.x = element_text(angle = 45, hjust = 1, size = 8),
    panel.grid.minor = element_blank(),
    strip.text = element_blank()   # ← removes person IDs
  )
```
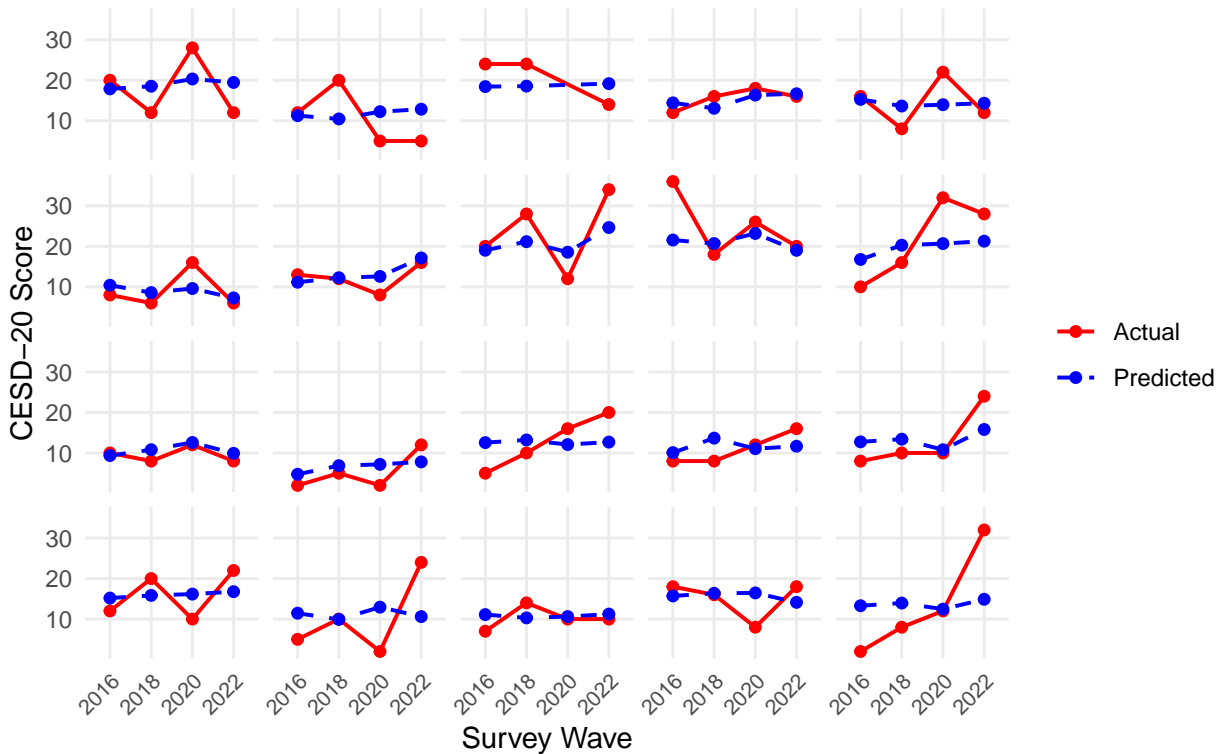
Actual vs. Predicted Depression Trajectories

Random sample of 20 individuals

Conclusion

Appendix A: Model Results (Required)

Model Call

```
# lmer(
#   outcome ~ predictors + (1 | group),
#   data = my_data
# )
```

Model Output

```
# library(arm)
# display(model)
```

References