Data Management & Analysis Final Project

Replication and Extention for Acemoglu, Naidu, Restrepo and Robinson (2019)

(Name:) Shoya Abe (University ID:) 31B24001
(Name:) Honoka Otani (University ID:) 31B24002

(Submission Due:) 2025/02/06

Contents

0.1 Setup

```
pacman::p_load(
  rmdformats,
  knitr,
  tinytex,
  haven,
  tidyverse,
  kableExtra,
  plm,
  texreg,
  PanelMatch,
  patchwork
)
options(max.print = "75")
opts_chunk$set(
  fig.align = "center",
  echo = TRUE,
  cache = TRUE,
  prompt = FALSE,
  tidy = FALSE,
  comment = NA,
  message = FALSE,
  warning = FALSE
)
opts_knit$set(width = 75)
```

1 About this Report

1.1 Project Type

In this report, we replicate and extend the previous paper. The paper we replicate is Acemoglu, D., Naidu, S., Restrepo, P., & Robinson, J. A. (2019). "Democracy Does Cause Growth." *Journal of Political Economy*, *127*(1), 47–100. https://doi.org/10.1086/700936.

We try to replicate Figure.1, Table.1, Table.2, Figure.4, Table.5 and Table.6 in the paper, which are especially critical results in the paper. We also try several extension approaches using the bootstrap method and the event study design. In appendix, we replicate Arellano Bond Estimation for Table.2 because we failed to replicate due to several limitations.

1.2 Summary of the Paper (Honoka Otani)

1.2.1 What the problem is

The authors attempted to provide a clear answer to the widely divergent topic of opinion on the causal relationship between democracy and economic growth. At the time of writing, there was a widely shared view that democracy has no relation to, or rather a negative effect on, economic growth.

On the other hand, there were empirical studies that showed a positive effect of democracy on economic growth, but they did not adequately address the endogeneity problem between political regimes and economic growth.

This paper points to four main challenges in estimating the causal relationship between democracy and economic growth. First, existing democracy indicators are subject to measurement error and changes in scores may not accurately reflect actual changes in political regimes. Second, there are institutional, historical and cultural differences between democracies and nondemocracies that also affect economic growth, which may introduce bias in the analysis. Third, democratization tends to occur after a temporary drop in GDP, which can bias estimates if not captured correctly in the model. Fourth, democratization and economic growth may be affected by common external factors, making it difficult to identify causality.

1.2.2 Why it is important

Demonstrating the causal relationship between democracy and economic growth has important implications for both political and economic development strategies. If democracy has a positive effect on economic growth, it provides an incentive to promote democratization across the world. It would also provide important hints to each country seeking to achieve economic growth. By providing empirical evidence, this study contributes to the competing debate on governance and economic growth.

1.2.3 How you solve the problem

To address the problem of measurement error in democracy indicators, the authors introduced a new democracy indicator by integrating several existing indicators and methods.

For other endogeneity problems, the authors employed three empirical strategies.

First, a dynamic (linear) panel model is used to control for country fixed effects and autoregressive GDP dynamics. By including lags of GDP per capita, this model accounts for the pre-democratization dip in GDP, ensuring that countries transitioning to democracy are not on a different GDP trend compared to other countries with similar past GDP levels.

Second, they adopted a propensity score reweighting strategy, one of semiparametric treatment effects framework, which democratization effects the distribution of potential GDP in all subsequent years. This method models the selection into democracy as a function of observable factors, particularly past GDP, without relying on a fully parametric GDP model. This approach increases flexibility in estimating how democracy effects GDP over time.

Third, they applied an instrumental variables (IV) method, using regional waves of democratization as an instrument for a country's transition to democracy. Since democratizations often occur in regional clusters, this method isolates exogenous variation in democracy that is not directly related to a country's own economic conditions. By leveraging this external source of variation, the IV approach strengthens the identification of the causal effect of democracy on GDP.

As for extension part, we undertake the following two tasks. First, in order to visualize the uncertainty of the long-term effect of democratization on economic growth, we estimate the confidence interval of the ATT estimate using the bootstrap method based on the event study in Figure 1. Second, we visualize the relationship between democracy and population using the event study design from the original paper.

1.2.4 What we find

The author found that democracy has a significant positive effect on GDP per capita. A country that transitions from nondemocracy to democracy experiences a long-run increase in GDP per capita of approximately 20–25% over the next 25 years. This effect is robust across different three strategies.

Furthermore, the analysis shows that the effect does not depend on a country's initial level of development, however, the effect is stronger in countries with higher levels of secondary education.

The authors also suggest several channels through which democracy promotes economic growth. They showed that democracy increases economic reforms, tax revenue (as a percentage of GDP). Enrollment in primary and secondary education and reduces child mortality rate. They also found the possibility that democracy promotes investment and open trade, and reduces social unrest.

Overall, the findings of this study strongly support the claim that democracy causes economic growth. This effect is primarily driven by democracy's ability to promote economic activity, improve human capital through education and healthcare, and strengthen governance structures, while may also be contributing to greater political stability and reduced social unrest. These results challenge the notion that democracy is a hindrance to economic growth and instead emphasize its role in promoting economic growth.

For the extension part, we discovered two things. First, based on the confidence intervals, we pointed out that the long-term effect of democratization on economic growth in the event study design is much more uncertain than stated in the paper. Second, we found that after democratisation, the proportion of children (0–14) initially declined before rebounding, while the working-age group (15–64) fell 5 years before democratization and recovered after 15 years of democratization. These shifts likely reflect temporary changes in fertility, migration, and economic conditions.

1.3 Data (Shoya Abe)

We use data obtained from the replication files available in the data archive on Professor Daron Acemoglu's homepage. This dataset consists of a large panel of 175 countries from 1960 to 2010. The sample size is 9,384, and the number of variables is 1,177. A list of variables is provided in the appendix.

```r
data <- read_dta("data/raw/DDCGdata_final.dta")

summarize_data <- function(data) {
  cat("Sample size (number of rows):", nrow(data), "\n")
  cat("Number of variables (columns):", ncol(data))
}

summarize_data(data)
```

```
Sample size (number of rows): 9384
Number of variables (columns): 1177
```

We replicate Figure 1 of Imai, et al. (2023) to check the transition of democratization in 175 countries. We used the dataset obtained from the replication file of Imai, et al. (2023), but the content is the same as the dataset of Acemoglu, et al. (2019). First, we will check the contents of this dataset.

```r
load("data/raw/Acemoglu.RData")

glimpse(d2)
```

```
Rows: 9,384
Columns: 12
$ wbcode2                   <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, ~
$ year                      <int> 1960, 1961, 1962, 1963, 1964, 1965, 1966, 196~
$ y                         <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ unrest                    <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, ~
$ dem                       <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
$ logpop                    <dbl> 1601.216, 1603.090, 1605.022, 1607.012, 1609.~
$ Populationages014oftotal <dbl> 42.17061, 42.47375, 42.64217, 42.73415, 42.82~
$ Populationages1564oftota <dbl> 55.03040, 54.71780, 54.55402, 54.48085, 54.42~
$ tradewb                   <dbl> NA, 11.47824, 12.97522, 18.52119, 25.75280, 2~
```

```
$ nfagdp                      <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ democ                       <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
$ rever                       <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
```

Next, we will use this dataset to check the transition of democratization in 175 countries.

```r
ADis <- DisplayTreatment(unit.id = "wbcode2",
                         time.id = "year",
                         xlab = "Years", ylab = "Countries",
                         legend.position = "bottom",
                         legend.labels = c("Autocracy",
                                           "Democracy"),
                         treatment = "dem", data = d2) +
  labs(title = NULL) +
  theme(axis.text.y = element_blank(),
        axis.ticks.y = element_blank(),
        axis.text.x = element_text(angle = 0, size = 6.5, vjust = 0.5)) +
  scale_x_discrete(breaks = c("1960", "1970", "1980", "1990", "2000", "2010"))

ggsave(file = "output/imai_figure_1.pdf",
       height = 8,
       width = 14,
       units = "cm",
       ADis)
```

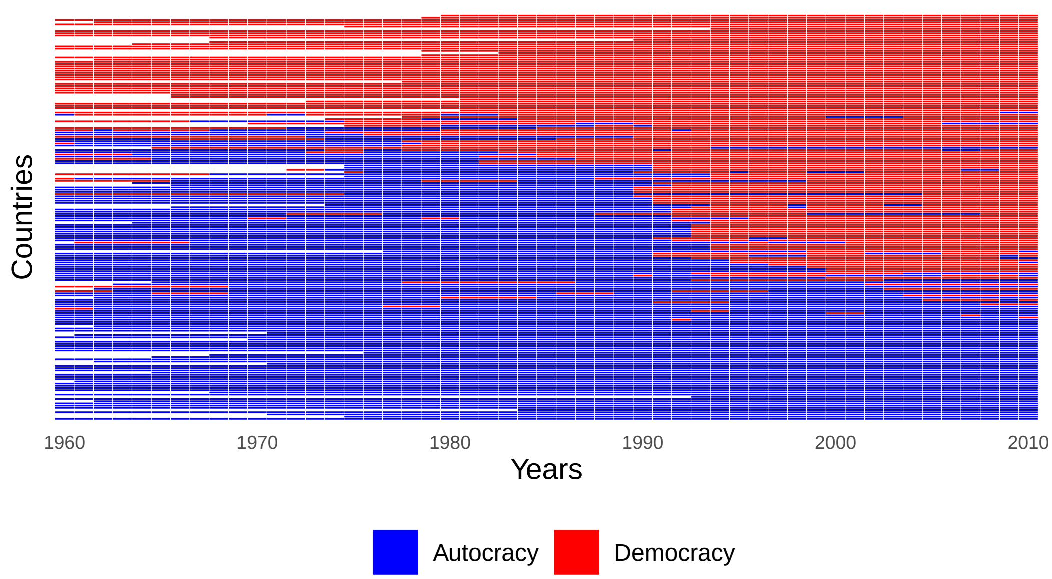

Figure 1: The transition of democratization in 175 countries

1.4 Empirical Methods (Shoya Abe)

We briefly explain the empirical methods we use for our replication. The original paper uses three main empirical strategies, in addition to visualization and descriptive statistics, to estimate the impact of democracy on economic growth. Here, we explain the four empirical methods used[1].

1.4.1 Event Study (Figure.1)

First, we conduct the event study. We estimate the average treatment effect on Treated (ATT) using the procedure described below.

Let T_c denote the year in which a given country experienced the democratization event. For any country c and year t, we define the relative year as

$$\tau_{c,t} = t - T_c. \tag{1}$$

Then, taking the outcome y in the year immediately preceding democratization (i.e., when $\tau = -1$) as the baseline, the outcome of interest is defined as

$$\text{gdpDiff}_{c,t} = y_{c,t} - y_{c,T_c-1}. \tag{2}$$

Next, we estimate the following regression model using the control group that did not experience democratization:

$$\text{gdpDiff}_{c,t} = \sum_{\tau=-15,\ \tau\neq-1}^{30} \beta_\tau \, \mathbf{1}\{\tau_{c,t} = \tau\} + \epsilon_{c,t}. \tag{3}$$

The estimated coefficient $\hat{\beta}_{\tau_{c,t}}$ from (3) can be interpreted as the counterfactual outcome for country c in year t in the absence of democratization. Therefore, the average difference between the observed outcome and this counterfactual outcome provides an estimate of the ATT for relative year τ, which is calculated as

$$\text{ATT}(\tau) = \frac{1}{N_\tau^{\text{treated}}} \sum_{\substack{(c,t)\in\text{treated} \\ \tau_{c,t}=\tau}} \Big(\text{gdpDiff}_{c,t} - \hat{\beta}_\tau\Big). \tag{4}$$

1.4.2 Dynamic Liner Panel Model (Table.2)

Second, we estimate the following dynamic linear panel model.

$$y_{c,t} = \beta D_{c,t} + \gamma_1 y_{c,t-1} + \alpha_c + \delta_t + \epsilon_{ct}, \tag{5}$$

$$y_{c,t} = \beta D_{c,t} + \sum_{j=1}^{2} \gamma_j y_{c,t-j} + \alpha_c + \delta_t + \epsilon_{c,t}, \tag{6}$$

$$y_{c,t} = \beta D_{c.t} + \sum_{j=1}^{4} \gamma_j y_{c,t-j} + \alpha_c + \delta_t + \epsilon_{c,t}, \tag{7}$$

$$y_{c,t} = \beta D_{c,t} + \sum_{j=1}^{8} \gamma_j y_{c,t-j} + \alpha_c + \delta_t + \epsilon_{c,t}, \tag{8}$$

where y_{ct} is the log of GDP per capita in country c at time t and D_{ct} is a dummy variable that takes the value 1 if country c is a democracy at time t and 0 otherwise. α_c is the country fixed effect and δ_t is the year fixed effect.

[1] We also worked on Arellano Bond estimation in table.2. However, it took an enormous amount of computation time and the results obtained were quite different from the original results. In other words, replication failed. However, in the belief that it is desirable to disclose the entire analysis process and results, we disclose the analysis code and results in the appendix.

1.4.3 Inverse-Propensity-Score Reweighting (Figure.4 and Table.5)

Third, we conduct inverse-propensity-score reweighting[2]. First, we estimate the following probit regression model and derive the propensity score $p(X)$ for the transition to democratization.

$$Pr(transition = 1|X) = \Phi\left(\gamma_0 + \sum_{j=1}^{4} \gamma_j y_{c,t-j} + \sum_{\tau} \beta_\tau \mathbf{1}\{\tau_{c,t} = \tau\}\right). \tag{9}$$

Next, based on the estimated propensity score $\hat{p}(X_i)$, we define the weight w_c for each observation as follows.

$$w_c = \begin{cases} 1, & \text{if}\, transition_c = 1, \\ \dfrac{\hat{P}(X_c)}{1 - \hat{P}(X_c)}, & \text{if}\, transition_c = 0. \end{cases} \tag{10}$$

Using this weight, we can estimate ATT as follows.

$$\hat{\text{ATT}} = \frac{1}{N_1} \sum_{c:transition_c=1} Y_i - \frac{\sum_{c:transition_c=0} w_c Y_c}{\sum_{c:transition_c=0} w_c}, \tag{11}$$

where N_1 is the sample size on treatment group. For the standard errors, we use the bootstrap method for estimation. This approach will be explained in the extension part.

1.4.4 Instrumental Variable (IV) Method (Table.6)

Fourth, we use the instrumental variable (IV) method in our analysis. The instrumental variable we employ is the regional waves of democratization. We formulate this as follows.

First, let D_{c,t_0} be a dummy variable indicating whether a country was democratic or non-democratic in 1960. Let R_c denote the geographic region to which country c belongs. Then, we define the set of countries that share a similar political history within the same region as $I_c = \{c^* : c^* \neq c, R_{c^*} = R_c, D_{c^*,t_0} = D_{c,t_0}\}$. The instrumental variable used in this analysis is given by:

$$Z_{c,t} = \frac{1}{|I_c|} \sum_{c^* \in I_c} D_{c^* t}. \tag{12}$$

This variable represents the proportion of countries that have undergone democratization among those with the same political history in the same region, thereby capturing the regional waves of democratization.

Using this instrumental variable, we conduct the following two-stage least squares (2SLS) estimation:

$$y_{c,t} = \beta D_{c,t} + \sum_{j=1}^{p} \gamma y_{c,t-j} + \alpha_c + \delta_t + \epsilon_{c,t}, \tag{13}$$

$$D_{c,t} = \sum_{j=1}^{q} \pi_j Z_{c,t-j} + \sum_{j=1}^{p} \phi_j y_{c,t-j} + \theta_c + \mu_t + v_{c,t}. \tag{14}$$

2 Replication

2.1 Figure.1 (Shoya Abe)

2.1.1 Preprocessing

First, we identify transitions in democratization from dataset, as well as the process of calculating time-series changes in GDP. Specifically, we calculate the differences in GDP from the past 15 years to 30 years in the

[2] This method is also known as Inverse Probability Weighting (IPW) estimation and is a representative approach in semi-parametric estimation.

future, and extract only data related to democratization transitions. In this way, the data is prepared for quantitative analysis of the effect of democratization on economic growth.

```r
#Prepare data
## rename ID column and calculate democracy transitions
data_f1 <- data |>
  rename(id = "_ID") |> # Rename column for consistency
  group_by(id) |>
  arrange(year) |>
  mutate(
    prev_dem = dplyr::lag(dem, 1),
      # Previous year's democracy status
    transition = case_when(
      dem == 1 & prev_dem == 0 ~ 1, # Transition to democracy
      dem == 0 & prev_dem == 0 ~ 0, # No transition
      TRUE ~ NA_real_
    ),
    lag1 = dplyr::lag(y, 1), # GDP per capita lag variables
    lag2 = dplyr::lag(y, 2),
    lag3 = dplyr::lag(y, 3),
    lag4 = dplyr::lag(y, 4)
  ) |>
  filter(
    !is.na(lag1) & !is.na(lag2) &
      !is.na(lag3) & !is.na(lag4) # Ensure complete lag data
  ) |>
  ungroup()

# Compute GDP differences for past years
for (t in -15:-2) {
  col_name <- paste0("gdpDiff_m", abs(t))
  data_f1 <- data_f1 |>
    group_by(id) |>
    arrange(year) |>
    mutate(!!col_name := dplyr::lag(y, abs(t)) - lag1) |>
    ungroup()
}

# Define GDP differences at t = -1 and t = 0
data_f1 <- data_f1 |>
  mutate(
    gdpDiff_m1 = 0,
    gdpDiff_0 = y - lag1
  )

# Compute GDP differences for future years
for (t in 1:30) {
  col_name <- paste0("gdpDiff_p", t)
  data_f1 <- data_f1 |>
    group_by(id) |>
    arrange(year) |>
    mutate(!!col_name := dplyr::lead(y, t) - lag1) |>
    ungroup()
}
```

```r
# Keep only observations relevant to democratization transitions
data_f1 <- data_f1 |>
  filter(!is.na(transition))
```

2.1.2 Estimation

Here, we define a function to estimate the ATT by comparing the treatment group (democratized countries) and the control group (non-democratized countries). Specifically, ATT is calculated by constructing a linear regression model using data from the control group, estimating the counterfactual (predicted value based on the control group) of GDP change for democratized countries, and finding the difference between the two. This is calculated for each relative time from the past 15 years to 30 years in the future, and the results are stored in a data frame.

```r
# Define function to estimate the Average Treatment Effect on the Treated (ATT)
estimateATT <- function(outcome_col) {
  sub_data <- data_f1 |>
    filter(!is.na(.data[[outcome_col]]), !is.na(transition))
  if (nrow(sub_data) == 0) return(NA)
  # Convert year to factor variable for regression
  year_levels <- sort(unique(sub_data$year))
  sub_data <- sub_data |>
    mutate(year_factor = factor(year, levels = year_levels))
  # Split data into control and treated groups
  control_data <- sub_data |>
    filter(transition == 0)
  treated_data <- sub_data |>
    filter(transition == 1)
  if (nrow(control_data) < 2 ||
      length(unique(control_data$year)) < 2) return(NA)
  # Estimate a linear model for control group
  model_formula <- as.formula(
    paste(outcome_col, "~ year_factor - 1")
  )
  control_model <- tryCatch(
    lm(model_formula, data = control_data),
    error = function(e) NULL
  )
  if (is.null(control_model)) return(NA)
  # Predict counterfactual outcomes for the treated group
  predicted_outcomes <- tryCatch(
    predict(control_model, newdata = treated_data),
    error = function(e) rep(NA, nrow(treated_data))
  )
  # Compute ATT as the difference between observed and predicted values
  treatment_effects <- treated_data[[outcome_col]] - predicted_outcomes
  mean(treatment_effects, na.rm = TRUE)
}

# Compute ATT estimates for each relative time period
relative_times <- c(seq(-15, -1), seq(0, 30))
atets <- numeric(length(relative_times))
```

```r
for (i in seq_along(relative_times)) {
  t_val <- relative_times[i]
  if (t_val < 0) {
    col_name <- paste0("gdpDiff_m", abs(t_val))
  } else {
    col_name <- if (t_val == 0) {
      "gdpDiff_0"
    } else {
      paste0("gdpDiff_p", t_val)
    }
  }
  atets[i] <- estimateATT(col_name)
}

# Store ATT estimates in a dataframe
results_df <- data.frame(
  RelativeTime = relative_times,
  ATT = atets
)
```

2.1.3 Plot

Plot ATT on Change in GDP per capita (log points) before and after democratization and save as a PDF.

```r
# Plot ATT estimates over time
figure_1 <- ggplot(results_df, aes(x = RelativeTime, y = ATT)) +
  geom_line(color = "black") +
  scale_x_continuous(breaks = seq(-15, 30, 5)) +
  labs(
    x = "Years around Democratization",
    y = "Change in GDP per capita (log points)"
  ) +
  theme_bw()

# Save the figure as a PDF
ggsave(
  "output/figure_1.pdf",
  width = 14,
  height = 8,
  units = "cm"
)
```

2.2 Table.1 (Honoka Otani)

2.2.1 Preprocessing

In order to compute summary statistics, the code defines the main variables related to the economy, education, politics, and health, and processes the mapping of variable names to their explanatory labels. We then create a data frame (data_sub) suitable for analysis by extracting from the original data (data) only the demogracy indicators (dem) and the variables specified.

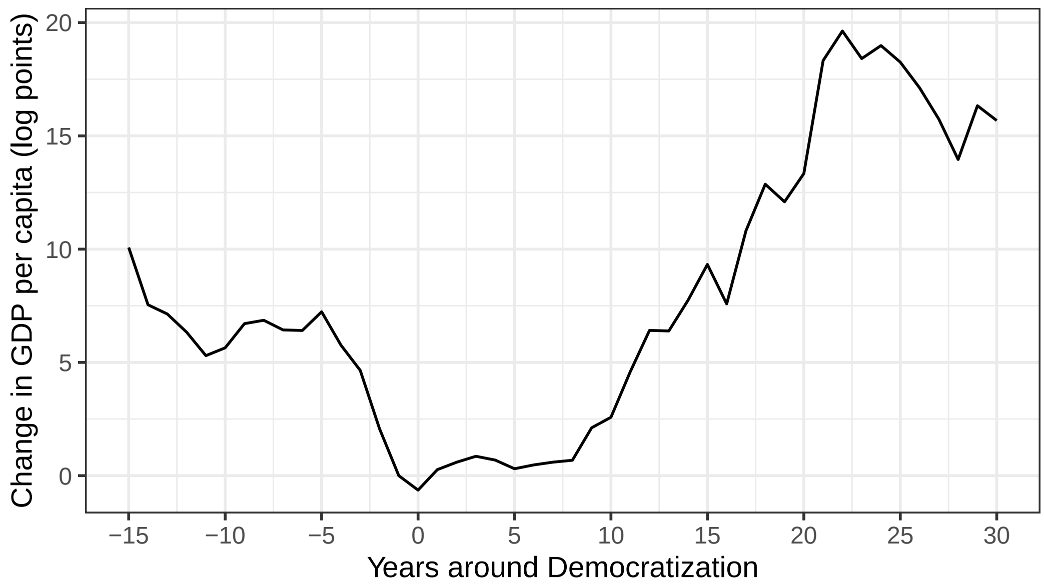

Figure 2: (Originally Figure.1) Change in GDP per capita (log points) before and after a democratization

```r
# Define variable names and labels for summary statistics
var_info <- tibble(
  var = c(
    "gdppercapitaconstant2000us",
    "loginvpc",
    "ltrade2",
    "lp_bl",
    "ls_bl",
    "lgov",
    "mortnew",
    "unrestn",
    "marketref"
  ),
  label = c(
    "GDP per capita",
    "Investment share of GDP",
    "Trade share of GDP",
    "Primary-school enrollment rate",
    "Secondary-school enrollment rate",
    "Tax revenue share of GDP",
    "Child mortality per 1,000 births",
    "Unrest rate",
    "Market reforms index (0-100)"
  )
)

# Select variables for summary statistics
data_sub <- data |>
  select(dem, all_of(var_info$var))
```

2.2.2 Calculation

For each variable, we extract data for democracies (dem == 1) and non-democracies (dem == 0), define a function (calc_stats) to calculate the number of observations, mean, and standard deviation for each, and apply it to all variables to create a summary statistics table (summary_table).

```r
# Compute summary statistics for each variable by democracy status
calc_stats <- function(variable) {
  non_demo <- data_sub |>
    filter(dem == 0) |>
    pull(.data[[variable]])
  non_demo <- non_demo[!is.na(non_demo)]
  demo <- data_sub |>
    filter(dem == 1) |>
    pull(.data[[variable]])
  demo <- demo[!is.na(demo)]
  tibble(
    var = variable,
    n_non_demo = length(non_demo),
    mean_non_demo = mean(non_demo),
    sd_non_demo = sd(non_demo),
    n_demo = length(demo),
    mean_demo = mean(demo),
    sd_demo = sd(demo)
  )
}

# Generate summary statistics table
summary_table <- map_dfr(var_info$var, calc_stats) |>
  left_join(var_info, by = "var") |>
  select(label, n_non_demo, mean_non_demo, sd_non_demo, n_demo, mean_demo, sd_demo)
```

2.2.3 Tabulation

We convert summary statistical tables into LaTeX format and saves them.

```r
# Convert summary statistics table to LaTeX format
latex_table <- summary_table |>
  kbl(
    caption = "Summary Statistics for the Main Variables Used in Our Analysis",
    format = "latex",
    booktabs = TRUE,
    digits = 2,
    col.names = c("", "N", "Mean", "SD", "N", "Mean", "SD")
  ) |>
  add_header_above(c(" " = 1, "Nondemocracies" = 3, "Democracies" = 3)) |>
  kable_styling(latex_options = c("HOLD_position", "striped"))

# Save the LaTeX table to a file
save_kable(latex_table, file = "output/table_1.tex")
```

Table 1: Summary Statistics for the Main Variables Used in Our Analysis

	Nondemocracies			Democracies		
	N	Mean	SD	N	Mean	SD
GDP per capita	3376	2074.46	3838.65	3558	8149.97	9334.83
Investment share of GDP	3222	297.18	50.15	3339	309.94	31.84
Trade share of GDP	3175	406.06	67.95	3485	419.29	58.74
Primary-school enrollment rate	817	32.14	19.56	689	38.10	20.05
Secondary-school enrollment rate	817	19.53	17.15	689	34.37	19.72
Tax revenue share of GDP	3122	-201.59	62.93	2564	-168.61	49.82
Child mortality per 1,000 births	4142	77.29	49.64	3615	33.26	32.65
Unrest rate	3739	28.70	45.24	3610	21.91	41.37
Market reforms index (0–100)	3476	21.89	23.26	2829	52.11	24.75

For summary statistics, the above code was used to generate the output, but figures other than GDP per Capita could not be replicated correctly. The same results were obtained when only the code for descriptive statistics was run in R, excluding all other code, or when the same code was derived for each variable.

The following points were confirmed to be correct, but the cause could not be identified. - The content of the original dataset used is correct. - The filtering of missing values in dem is functioning correctly. - All variables are in data_sub. - There is no significant difference between the contents of the dataset (number and composition of NAs) for GDP per capita and other variables.

Given that there are no major problems with the processing code, the reason why only GDP per capita fits and the other main control variables do not is that the pre-processing methods for these covariates may differ from those in the original paper. In the original paper, the calculations were done in Stata and the corresponding R codes were created this time, but the same results were not obtained. One possibility is that in the original paper, some special pre-processing is at work only for the covariates.

2.3 Table.2 (Honoka Otani)

2.3.1 Preprocessing

We select columns 1-30, arrange them in annual order by country, and then create lagged variables of 1 to 8 years for key variables such as GDP. This allows for analysis of the relationship with past values.

```r
# Select relevant variables and create lag variables
data_t2 <- data |>
  select(1:30) |>
  group_by(country_name) |>
  arrange(year) |>
  mutate(
    lag1 = dplyr::lag(y, 1),
    lag2 = dplyr::lag(y, 2),
    lag3 = dplyr::lag(y, 3),
    lag4 = dplyr::lag(y, 4),
    lag5 = dplyr::lag(y, 5),
    lag6 = dplyr::lag(y, 6),
    lag7 = dplyr::lag(y, 7),
    lag8 = dplyr::lag(y, 8)
  ) |>
  ungroup()
```

2.3.2 Estimation

We uses a Fixed Effects Model to estimate the effect of democratization on economic growth. Specifically, we estimates four panel regression models with different lagged variable specifications and performs the following analysis based on the coefficients of each:

- Calculation of the Long-Run Effect
- Calculation of the Persistence Effect
- Estimating the Effect after 25 years

Then, we calculate standard errors of estimated coefficients and format into LaTeX tables: store in override.coef and override.se.

```
# Estimate fixed effects models with different lag specifications
data_m1 <- data_t2 |>
  drop_na(y, dem, lag1) |>
  pdata.frame(index = c("country_name", "year"))
model_1 <- plm(
  y ~ dem + lag1,
  data = data_m1,
  model = "within",
  effect = "twoways"
)

data_m2 <- data_t2 |>
  drop_na(y, dem, lag1, lag2) |>
  pdata.frame(index = c("country_name", "year"))
model_2 <- plm(
  y ~ dem + lag1 + lag2,
  data = data_m2,
  model = "within",
  effect = "twoways"
)

data_m3 <- data_t2 |>
  drop_na(y, dem, lag1, lag2, lag3, lag4) |>
  pdata.frame(index = c("country_name", "year"))
model_3 <- plm(
  y ~ dem + lag1 + lag2 + lag3 + lag4,
  data = data_m3,
  model = "within",
  effect = "twoways"
)

data_m4 <- data_t2 |>
  drop_na(
    y, dem, lag1, lag2, lag3, lag4,
    lag5, lag6, lag7, lag8
  ) |>
  pdata.frame(index = c("country_name", "year"))
model_4 <- plm(
  y ~ dem + lag1 + lag2 + lag3 + lag4 +
    lag5 + lag6 + lag7 + lag8,
```

```r
  data = data_m4,
  model = "within",
  effect = "twoways"
)

# Compute long-run effects of democratization
beta_hat_1 <- coef(model_1)["dem"]
gamma_hat_1 <- coef(model_1)["lag1"]
long_run_effect_1 <- beta_hat_1 / (1 - sum(gamma_hat_1))

beta_hat_2 <- coef(model_2)["dem"]
gamma_hat_2 <- coef(model_2)[c("lag1", "lag2")]
long_run_effect_2 <- beta_hat_2 / (1 - sum(gamma_hat_2))

beta_hat_3 <- coef(model_3)["dem"]
gamma_hat_3 <- coef(model_3)[c("lag1", "lag2", "lag3", "lag4")]
long_run_effect_3 <- beta_hat_3 / (1 - sum(gamma_hat_3))

beta_hat_4 <- coef(model_4)["dem"]
gamma_hat_4 <- coef(model_4)[
  c("lag1", "lag2", "lag3", "lag4",
    "lag5", "lag6", "lag7", "lag8")
]
long_run_effect_4 <- beta_hat_4 / (1 - sum(gamma_hat_4))

lre <- round(
  c(long_run_effect_1, long_run_effect_2,
    long_run_effect_3, long_run_effect_4),
  3
)

# Compute persistence effects
pers1 <- sum(coef(model_1)[2])
pers2 <- sum(coef(model_2)[2:3])
pers3 <- sum(coef(model_3)[2:5])
pers4 <- sum(coef(model_4)[2:9])
pers <- round(c(pers1, pers2, pers3, pers4), 3)

# Compute effect after 25 years for each model
dem_shortrun <- coef(model_1)["dem"]
lag1_mod1 <- coef(model_1)[2]
effect1 <- dem_shortrun
effect2 <- (effect1 * lag1_mod1) + dem_shortrun
effects_mod1 <- c(effect1, effect2)
for (i in 3:30) {
  eff <- (effects_mod1[i - 1] * lag1_mod1) + dem_shortrun
  effects_mod1 <- c(effects_mod1, eff)
}
eff_25_1 <- effects_mod1[25]

dem_shortrun <- coef(model_2)["dem"]
lag1_mod2 <- coef(model_2)[2]
lag2_mod2 <- coef(model_2)[3]
```

```
effect1 <- dem_shortrun
effect2 <- (effect1 * lag1_mod2) + dem_shortrun
effect3 <- (effect2 * lag1_mod2) +
  (effect1 * lag2_mod2) + dem_shortrun
effects_mod2 <- c(effect1, effect2, effect3)
for (i in 4:30) {
  eff <- (effects_mod2[i - 1] * lag1_mod2) +
    (effects_mod2[i - 2] * lag2_mod2) +
    dem_shortrun
  effects_mod2 <- c(effects_mod2, eff)
}
eff_25_2 <- effects_mod2[25]

dem_shortrun <- coef(model_3)["dem"]
lag1_mod3 <- coef(model_3)[2]
lag2_mod3 <- coef(model_3)[3]
lag3_mod3 <- coef(model_3)[4]
lag4_mod3 <- coef(model_3)[5]
effect1 <- dem_shortrun
effect2 <- (effect1 * lag1_mod3) + dem_shortrun
effect3 <- (effect2 * lag1_mod3) +
  (effect1 * lag2_mod3) + dem_shortrun
effect4 <- (effect3 * lag1_mod3) +
  (effect2 * lag2_mod3) +
  (effect1 * lag3_mod3) + dem_shortrun
effects_mod3 <- c(effect1, effect2, effect3, effect4)
for (i in 5:30) {
  eff <- (effects_mod3[i - 1] * lag1_mod3) +
    (effects_mod3[i - 2] * lag2_mod3) +
    (effects_mod3[i - 3] * lag3_mod3) +
    (effects_mod3[i - 4] * lag4_mod3) +
    dem_shortrun
  effects_mod3 <- c(effects_mod3, eff)
}
eff_25_3 <- effects_mod3[25]

dem_shortrun <- coef(model_4)["dem"]
lag1_mod4 <- coef(model_4)[2]
lag2_mod4 <- coef(model_4)[3]
lag3_mod4 <- coef(model_4)[4]
lag4_mod4 <- coef(model_4)[5]
lag5_mod4 <- coef(model_4)[6]
lag6_mod4 <- coef(model_4)[7]
lag7_mod4 <- coef(model_4)[8]
lag8_mod4 <- coef(model_4)[9]
effect1 <- dem_shortrun
effect2 <- (effect1 * lag1_mod4) + dem_shortrun
effect3 <- (effect2 * lag1_mod4) +
  (effect1 * lag2_mod4) + dem_shortrun
effect4 <- (effect3 * lag1_mod4) +
  (effect2 * lag2_mod4) +
  (effect1 * lag3_mod4) + dem_shortrun
effect5 <- (effect4 * lag1_mod4) +
```

```r
  (effect3 * lag2_mod4) +
  (effect2 * lag3_mod4) +
  (effect1 * lag4_mod4) + dem_shortrun
effect6 <- (effect5 * lag1_mod4) +
  (effect4 * lag2_mod4) +
  (effect3 * lag3_mod4) +
  (effect2 * lag4_mod4) +
  (effect1 * lag5_mod4) + dem_shortrun
effect7 <- (effect6 * lag1_mod4) +
  (effect5 * lag2_mod4) +
  (effect4 * lag3_mod4) +
  (effect3 * lag4_mod4) +
  (effect2 * lag5_mod4) +
  (effect1 * lag6_mod4) + dem_shortrun
effect8 <- (effect7 * lag1_mod4) +
  (effect6 * lag2_mod4) +
  (effect5 * lag3_mod4) +
  (effect4 * lag4_mod4) +
  (effect3 * lag5_mod4) +
  (effect2 * lag6_mod4) +
  (effect1 * lag7_mod4) + dem_shortrun
effects_mod4 <- c(
  effect1, effect2, effect3, effect4,
  effect5, effect6, effect7, effect8
)
for (i in 9:30) {
  eff <- (effects_mod4[i - 1] * lag1_mod4) +
    (effects_mod4[i - 2] * lag2_mod4) +
    (effects_mod4[i - 3] * lag3_mod4) +
    (effects_mod4[i - 4] * lag4_mod4) +
    (effects_mod4[i - 5] * lag5_mod4) +
    (effects_mod4[i - 6] * lag6_mod4) +
    (effects_mod4[i - 7] * lag7_mod4) +
    (effects_mod4[i - 8] * lag8_mod4) +
    dem_shortrun
  effects_mod4 <- c(effects_mod4, eff)
}
eff_25_4 <- effects_mod4[25]

eff_25 <- round(
  c(eff_25_1, eff_25_2, eff_25_3, eff_25_4),
  3
)

# Compute standard errors for coefficients
se1 <- sqrt(diag(vcov(model_1)))
se2 <- sqrt(diag(vcov(model_2)))
se3 <- sqrt(diag(vcov(model_3)))
se4 <- sqrt(diag(vcov(model_4)))

# Override coefficients and standard errors for LaTeX table output
override.coef.1 <- c(
  coef(model_1)["dem"],
```

```
  coef(model_1)["lag1"],
  NA, NA, NA, NA, NA, NA, NA
)
override.se.1 <- c(
  se1["dem"],
  se1["lag1"],
  NA, NA, NA, NA, NA, NA, NA
)

override.coef.2 <- c(
  coef(model_2)["dem"],
  coef(model_2)["lag1"],
  coef(model_2)["lag2"],
  NA, NA, NA, NA, NA, NA
)
override.se.2 <- c(
  se2["dem"],
  se2["lag1"],
  se2["lag2"],
  NA, NA, NA, NA, NA, NA
)

override.coef.3 <- c(
  coef(model_3)["dem"],
  coef(model_3)["lag1"],
  coef(model_3)["lag2"],
  coef(model_3)["lag3"],
  coef(model_3)["lag4"],
  NA, NA, NA, NA
)
override.se.3 <- c(
  se3["dem"],
  se3["lag1"],
  se3["lag2"],
  se3["lag3"],
  se3["lag4"],
  NA, NA, NA, NA
)

override.coef.4 <- c(
  coef(model_4)["dem"],
  coef(model_4)["lag1"],
  coef(model_4)["lag2"],
  coef(model_4)["lag3"],
  coef(model_4)["lag4"],
  coef(model_4)["lag5"],
  coef(model_4)["lag6"],
  coef(model_4)["lag7"],
  coef(model_4)["lag8"]
)
override.se.4 <- c(
  se4["dem"],
  se4["lag1"],
```

```
  se4["lag2"],
  se4["lag3"],
  se4["lag4"],
  se4["lag5"],
  se4["lag6"],
  se4["lag7"],
  se4["lag8"]
)
```

2.3.3 Tabulation

We format the estimation results of the Dynamic Linear Panel Model in LaTeX format.

```
# Generate LaTeX table for regression results
models <- list(model_1, model_2, model_3, model_4)
texreg(
  models,
  override.coef = list(
    override.coef.1,
    override.coef.2,
    override.coef.3,
    override.coef.4
  ),
  override.se = list(
    override.se.1,
    override.se.2,
    override.se.3,
    override.se.4
  ),
  custom.model.names = c("(1)", "(2)", "(3)", "(4)"),
  custom.coef.names = c(
    "Democracy", "Lag 1", "Lag 2",
    "Lag 3", "Lag 4", "Lag 5",
    "Lag 6", "Lag 7", "Lag 8"
  ),
  custom.gof.rows = list(
    "Persistence" = pers,
    "Long run effect" = lre,
    "Effect after 25 years" = eff_25
  ),
  file = "output/table_2_FE.tex",
  caption = "Effect of Democracy on (Log) GDP per Capita"
)
```

2.4 Figure.4 (Honoka Otani)

2.4.1 Preprocessing

In replication of figure.4, we use the data set related to the estimated values contained in the replication file as is. We conduct the data shaping exercise to analyze the effect of democratization. First, the variable parm in the data is split into two variables (parm1 and parm2) separated by the letter “c”. Then, the line

	(1)	(2)	(3)	(4)
Democracy	0.97^{***}	0.65^{**}	0.79^{***}	0.89^{***}
	(0.24)	(0.23)	(0.23)	(0.24)
Lag 1	0.97^{***}	1.27^{***}	1.24^{***}	1.23^{***}
	(0.00)	(0.01)	(0.01)	(0.01)
Lag 2		-0.30^{***}	-0.21^{***}	-0.21^{***}
		(0.01)	(0.02)	(0.02)
Lag 3			-0.03	-0.02
			(0.02)	(0.02)
Lag 4			-0.04^{***}	-0.04
			(0.01)	(0.02)
Lag 5				-0.02
				(0.02)
Lag 6				0.01
				(0.02)
Lag 7				0.02
				(0.02)
Lag 8				-0.01
				(0.01)
Persistence	0.97	0.97	0.96	0.96
Long run effect	35.59	19.60	21.24	22.01
Effect after 25 years	17.79	13.80	16.90	17.72
R^2	0.96	0.96	0.96	0.96
Adj. R^2	0.96	0.96	0.96	0.96
Num. obs.	6790	6642	6336	5688

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 2: Effect of Democracy on (Log) GDP per Capita

with no value in parm2 is deleted and converted to be treated as a numerical value. Finally, a variable named "time" is created by subtracting 16 from parm2, and an indicator for the number of years before and after democratization is added.

```
# Load dataset and separate variable names
data_ipw <- read_dta("data/raw/impulse_ipw_alt.dta")

# Split 'parm' column into 'parm1' and 'parm2' using 'c' as separator
data_ipw <- data_ipw |>
  separate(parm,
           into = c("parm1", "parm2"),
           sep = "c",
           extra = "merge",
           fill = "right")

# Remove rows where 'parm2' is empty and convert 'parm2' to numeric
data_ipw <- data_ipw |>
  filter(parm2 != "") |>
  mutate(parm2 = as.numeric(parm2))

# Compute time relative to democratization
data_ipw <- data_ipw |>
  mutate(time = parm2 - 16)
```

2.4.2 Plot

We graphically summarizes the change in GDP in a time series. The x axis shows "years around democratization" and the y axis shows "change in GDP (log points)," with a black line depicting the change in GDP. In addition, a 95% confidence interval is added as a light blue band (ribbon).

```
# Plot impulse response function for GDP change around democratization
figure_4 <- ggplot(data_ipw, aes(x = time, y = estimate)) +
  geom_line(color = "black") +
  geom_ribbon(aes(ymin = min95, ymax = max95), fill = "skyblue", alpha = 0.3) +
  labs(x = "Years around democratization",
       y = "Change in GDP per capita (log points)") +
  scale_x_continuous(breaks = seq(-15, 30, 5)) +
  theme_bw()

# Save figure as a PDF
ggsave("output/figure_4.pdf",
       figure_4,
       width = 14,
       height = 8,
       units = "cm")
```

2.5 Table.5 (Shoya Abe)

2.5.1 Preprocessing

First, the data are sorted by year for each country so that we can check the democracy status of the previous year. Then, we record "if the country was non-democratic in the previous year but became democratic this

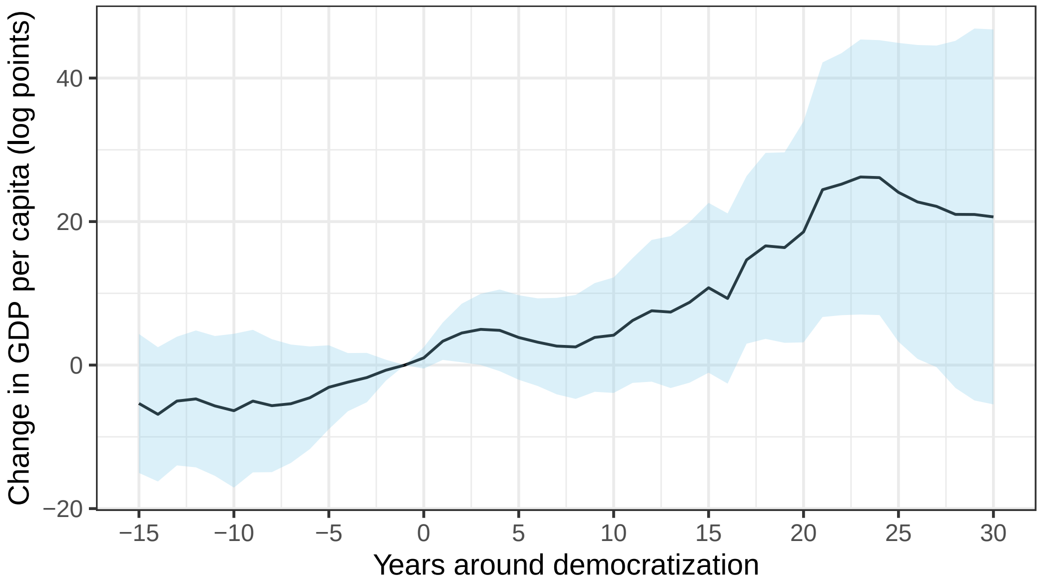

Figure 3: (Originally Figure.4) Semiparametric estimates of the over-time effects of democracy on the log of GDP, obtained with inverse-propensity-score reweighting

year" as a transition to democracy (transition = 1), and if the country did not become democratic, transition = 0. This clarifies the change in democratization for each country.

Next, we calculate GDP data from the previous year up to four years ago and keep only those data for which they are available. Furthermore, we calculate the change in GDP from 15 years ago to 2 years ago, and the year before democratization (t = -1) is set to 0 as the base value. In the year in which democratization occurred (t = 0), the difference from GDP in the previous year is calculated, and after democratization, the change in GDP is recorded from 1 year to 30 years later.

Finally, the data are narrowed down to only those countries for which democratization transition data are recorded, so that the effect of democratization can be accurately analyzed.

```
# Prepare dataset for ATT estimation
data_f1 <- data |>
  rename(id = "_ID") |>
  group_by(id) |>
  arrange(year) |>
  ungroup()

# Compute democratization transition variable
data_f1 <- data_f1 |>
  group_by(id) |>
  arrange(year) |>
  mutate(prev_dem = dplyr::lag(dem, 1)) |>
  ungroup() |>
  mutate(transition = case_when(
    dem == 1 & prev_dem == 0 ~ 1,
    dem == 0 & prev_dem == 0 ~ 0,
    TRUE ~ NA_real_
  ))

# Compute lag variables for GDP
```

```r
data_f1 <- data_f1 |>
  group_by(id) |>
  arrange(year) |>
  mutate(
    lag1 = dplyr::lag(y, 1),
    lag2 = dplyr::lag(y, 2),
    lag3 = dplyr::lag(y, 3),
    lag4 = dplyr::lag(y, 4)
  ) |>
  ungroup() |>
  filter(!is.na(lag1) & !is.na(lag2) & !is.na(lag3) & !is.na(lag4))

# Compute GDP differences for pre-democratization periods
for (t in -15:-2) {
  col_name <- paste0("gdpDiff_m", abs(t))
  data_f1 <- data_f1 |>
    group_by(id) |>
    arrange(year) |>
    mutate(!!col_name := dplyr::lag(y, abs(t)) - lag1) |>
    ungroup()
}

# Define GDP differences at t = -1 and t = 0
data_f1 <- data_f1 |>
  mutate(gdpDiff_m1 = 0)

data_f1 <- data_f1 |>
  group_by(id) |>
  arrange(year) |>
  mutate(gdpDiff_0 = y - lag1) |>
  ungroup()

# Compute GDP differences for post-democratization periods
for (t in 1:30) {
  col_name <- paste0("gdpDiff_p", t)
  data_f1 <- data_f1 |>
    group_by(id) |>
    arrange(year) |>
    mutate(!!col_name := dplyr::lead(y, t) - lag1) |>
    ungroup()
}

# Keep only relevant observations
data_f1 <- data_f1 |> filter(!is.na(transition))
```

2.5.2 Estimation

Here we estimate the ATT using Inverse Probability Weighting (IPW: Inverse Probability Weighting).

First, we estimate a propensity score that accounts for lagged GDP (lag1-lag4) and each year fixed effects so that countries with similar conditions can be compared between countries that did and did not democratize. The propensity score is then used to weight each country's data, and the ATT is calculated by estimating the change in GDP if the country had not democratized and finding the difference from the actual change in GDP.

To ensure the reliability of the estimates, we bootstrap (B = 200 resamplings) and estimate standard errors. This process is repeated for each year from 15 years before democratization to 30 years after democratization to assess the effect of GDP changes at each time point.

Furthermore, in order to summarize the obtained ATT estimates more clearly, we group them by time period and calculate the mean values and standard errors.

```r
# Compute ATT using inverse probability weighting (IPW)
compute_atet_ipw <- function(outcome_var, data) {
   # Filter relevant observations
  df <- data_f1 |>
    filter(!is.na(!!sym(outcome_var)),
           !is.na(transition),
           !is.na(lag1), !is.na(lag2), !is.na(lag3), !is.na(lag4),
           !is.na(year))
  # Estimate propensity score model
  prop_model <- glm(transition ~ lag1 + lag2 + lag3 + lag4 + factor(year),
                    data = df, family = binomial(link = "probit"))
  # Compute propensity scores and inverse probability weights
  df <- df |> mutate(ps = predict(prop_model, type = "response"))
  df <- df |> mutate(weight = ifelse(transition == 0, ps/(1 - ps), 1))
  # Extract treated and control group outcomes
  treated_outcome <- df |> filter(transition == 1) |> pull(!!sym(outcome_var))
  control_df <- df |> filter(transition == 0)
  control_outcome <- control_df[[outcome_var]]
  control_weight  <- control_df$weight
  # Compute ATT as the weighted mean difference
  att <- mean(treated_outcome) - (sum(control_outcome * control_weight) / sum(control_weight))
  return(att)
}

# Compute bootstrapped ATT estimates
compute_att_ipw_boot <- function(outcome_var, data, B = 200) {
  # Compute original ATT estimate
  att_est <- compute_atet_ipw(outcome_var, data)
  # Initialize bootstrap estimates
  n <- nrow(data)
  boot_est <- numeric(B)
  set.seed(123)
  # Bootstrap resampling
  for (b in 1:B) {
    boot_indices <- sample(1:n, size = n, replace = TRUE)
    boot_data <- data[boot_indices, ]
    boot_est[b] <- compute_atet_ipw(outcome_var, boot_data)
  }
  # Compute standard error from bootstrap estimates
  se_est <- sd(boot_est)
  return(list(att = att_est, se = se_est, boot = boot_est))
}

# Compute ATT estimates for each relative time period
outcome_vars <- c(
  paste0("gdpDiff_m", 15:2),
```

```r
  "gdpDiff_m1",
  "gdpDiff_0",
  paste0("gdpDiff_p", 1:30)
)

# Store results in a list
att_results <- list()
for (var in outcome_vars) {
  att_results[[var]] <- compute_att_ipw_boot(var, data_f1, B = 200)
}

# Aggregate ATT estimates into grouped time periods
group_definitions <- list(
  "-5 to -1" = c("gdpDiff_m5", "gdpDiff_m4", "gdpDiff_m3", "gdpDiff_m2", "gdpDiff_m1"),
  "0 to 4"   = c("gdpDiff_0", "gdpDiff_p1", "gdpDiff_p2", "gdpDiff_p3", "gdpDiff_p4"),
  "5 to 9"   = paste0("gdpDiff_p", 5:9),
  "10 to 14" = paste0("gdpDiff_p", 10:14),
  "15 to 19" = paste0("gdpDiff_p", 15:19),
  "20 to 24" = paste0("gdpDiff_p", 20:24),
  "25 to 29" = paste0("gdpDiff_p", 25:29)
)

# Compute mean ATT and standard error for each group
group_results <- list()
for (grp in names(group_definitions)) {
  vars_in_grp <- group_definitions[[grp]]
  att_vec <- sapply(vars_in_grp, function(x) att_results[[x]]$att)
  boot_mat <- sapply(vars_in_grp, function(x) att_results[[x]]$boot)
  grp_boot <- rowMeans(boot_mat)
  grp_att <- mean(att_vec)
  grp_se <- sd(grp_boot)
  group_results[[grp]] <- list(att = grp_att, se = grp_se)
}

# Prepare results for tabulation
group_names <- names(group_results)
table_values <- sapply(group_names, function(grp) {
  sprintf("%.3f", group_results[[grp]]$att)
})
table_ses <- sapply(group_names, function(grp) {
  sprintf("(%.3f)", group_results[[grp]]$se)
})
cell_text <- mapply(function(val, se) {
  paste0(val, "\n", se)
}, table_values, table_ses, SIMPLIFY = TRUE)

# Convert results to dataframe
results_df <- as.data.frame(t(cell_text))
colnames(results_df) <- group_names

# Rename column for clarity
results_df <- results_df |>
  mutate(years = "ATT on GDP (Log)")
```

```r
# Arrange columns in proper order
results_df <- results_df |>
  select(years, everything())
```

2.5.3 Tabulation

We format the results of the previous estimation into a LaTeX-formatted table.

```r
# Convert results dataframe to LaTeX format table
table_latex <- results_df |>
  kable(format = "latex",
        booktabs = TRUE,
        escape = FALSE,
        caption = "(Originally Table.5) Semiparametric Estimates of the Effect of Democratizations on (I
        label = "tab:table_5_ipw",
        digits = 3) |>
  add_header_above(c("Inverse propensity score reweighting" = ncol(results_df))) |>
  kable_styling(latex_options = c("hold_position", "scale_down"))

# Save LaTeX table output to file
writeLines(table_latex, con = "output/table_5_ipw.tex")
```

Table 3: (Originally Table.5) Semiparametric Estimates of the Effect of Democratizations on (Log) GDP per Capita

Inverse propensity score reweighting							
years	-5 to -1	0 to 4	5 to 9	10 to 14	15 to 19	20 to 24	25 to 29
ATT on GDP (Log)	-1.586 (1.069)	3.724 (1.921)	3.207 (2.849)	6.563 (3.466)	13.242 (4.192)	23.925 (5.697)	21.516 (7.729)

Our estimation results from the third column onward differ from those in the original paper. Two reasons may account for this discrepancy. First, there is a difference in the algorithms used for estimation. We created our own function in R, whereas the replication code in the original paper appears to utilize the standard functions available in Stata. This difference may have led to variations in bootstrap settings and, consequently, the estimation results. Second, there is a difference in the data preprocessing. There may have been issues in the preprocessing procedures between the original paper's replication code and our analysis code, which could have altered the covariates used for the propensity score. However, since the replication code of the original paper is written in Stata and we do not fully understand its preprocessing methods, we cannot provide a more detailed explanation.

2.6 Table.6 (Shoya Abe)

2.6.1 Preprocessing

We process data to conduct an instrument variable estimation in econometrics.

First, for each country, the data are sorted by year (arrange(year)) and the historical value of GDP (y) (lag variable) is calculated. Specifically, we add the values of GDP from 1 to 8 years ago (lag1 to lag8).

The data are then ungrouped (ungroup()) and converted to panel data format (pdata.frame), which makes it easier to apply operational variable estimation.

```r
# Prepare dataset for IV estimation
data_t6 <- data  |>
  group_by(country_name) |> # Group data by country
  arrange(year) |> # Arrange data by year in ascending order
  mutate(
    # Generate lagged variables for GDP up to 8 years
    lag1 = dplyr::lag(y, 1),
    lag2 = dplyr::lag(y, 2),
    lag3 = dplyr::lag(y, 3),
    lag4 = dplyr::lag(y, 4),
    lag5 = dplyr::lag(y, 5),
    lag6 = dplyr::lag(y, 6),
    lag7 = dplyr::lag(y, 7),
    lag8 = dplyr::lag(y, 8)
  ) |>
  ungroup() |> # Remove grouping
  pdata.frame(index = c("country_name", "year"))
```

2.6.2 Estimation

The code implements a panel data regression analysis using instrumental variable (IV) estimation to more accurately estimate the causal effect of democratization on GDP.

First, we estimate four different IV models. The basic model (model_iv_1) uses democratization (dem) and historical values of GDP (lagged variables for up to four years) as explanatory variables and the previous year's democratization status (lagged demreg) as the instrumental variable. In the next model (model_iv_2), the instrument variable of democratization is not only lagged by one year, but is estimated more robustly using four years of lagged variables. Furthermore, in model_iv_3, we add dummy variables (sov1 to sov4) indicating former Soviet Union countries to account for region-specific effects. Finally, in model_iv_4, we control for region-specific growth trends (rtrend2 to rtrend7) to provide more refined estimates.

Next, we compute the long-run effect of democratization.We then simulate the short-run effect of democratization as its effect on GDP after 25 years. Finally, we calculate the persistence of GDP. This is the sum of the effect of past values of GDP on current GDP.

```r
# Estimate instrumental variables (IV) models
model_iv_1 <- plm(
  y ~ dem + plm::lag(y, 1:4) | # Democracy and up to 4 lags of GDP as regressors
    plm::lag(demreg, 1) + plm::lag(y, 1:4),
    # Instrument: lagged democracy variable
  data = data_t6,
  effect = "twoways"
)
model_iv_2 <- plm(
  y ~ dem + plm::lag(y, 1:4) |
    plm::lag(demreg, 1:4) + plm::lag(y, 1:4),
    # Use additional lags of democracy as instruments
  data = data_t6,
  effect = "twoways"
)
model_iv_3 <- plm(
  y ~ dem + plm::lag(y, 1:4) + sov1 + sov2 + sov3 + sov4 |
    # Include Soviet region dummies
    plm::lag(demreg, 1:4) + plm::lag(y, 1:4) +
```

```
    sov1 + sov2 + sov3 + sov4,
  data = data_t6,
  effect = "twoways"
)
model_iv_4 <- plm(
  y ~ dem + plm::lag(y, 1:4) +
    rtrend2 + rtrend3 + rtrend4 + rtrend5 + rtrend6 + rtrend7 |
     # Include regional trends
    plm::lag(demreg, 1:4) + plm::lag(y, 1:4) +
    rtrend2 + rtrend3 + rtrend4 + rtrend5 + rtrend6 + rtrend7,
  data = data_t6,
  effect = "twoways",
  model = "within"
)

# Compute long-run effects of democratization
beta_hat_1 <- coef(model_iv_1)["dem"]
gamma_hat_1 <- coef(model_iv_1)[2:5]
long_run_effect_1 <- beta_hat_1 / (1 - sum(gamma_hat_1))

beta_hat_2 <- coef(model_iv_2)["dem"]
gamma_hat_2 <- coef(model_iv_2)[2:5]
long_run_effect_2 <- beta_hat_2 / (1 - sum(gamma_hat_2))

beta_hat_3 <- coef(model_iv_3)["dem"]
gamma_hat_3 <- coef(model_iv_3)[2:5]
long_run_effect_3 <- beta_hat_3 / (1 - sum(gamma_hat_3))

beta_hat_4 <- coef(model_iv_4)["dem"]
gamma_hat_4 <- coef(model_iv_4)[2:5]
long_run_effect_4 <- beta_hat_4 / (1 - sum(gamma_hat_4))

# Round the results for clarity
lre <- round(
  c(long_run_effect_1, long_run_effect_2,
    long_run_effect_3, long_run_effect_4),
  3
)

# Compute short-run effects at year 25
sre <- c()

dem_shortrun <- coef(model_iv_1)["dem"]
lag1 <- coef(model_iv_1)[2]
lag2 <- coef(model_iv_1)[3]
lag3 <- coef(model_iv_1)[4]
lag4 <- coef(model_iv_1)[5]
effect1 <- dem_shortrun
effect2 <- effect1 * lag1 + dem_shortrun
effect3 <- effect2 * lag1 + effect1 * lag2 + dem_shortrun
effect4 <- effect3 * lag1 + effect2 * lag2 +
  effect1 * lag3 + dem_shortrun
effects <- c(effect1, effect2, effect3, effect4)
```

```r
for (i in 5:30) {
  eff <- effects[i - 1] * lag1 +
    effects[i - 2] * lag2 +
    effects[i - 3] * lag3 +
    effects[i - 4] * lag4 + dem_shortrun
  effects <- c(effects, eff)
}
# Round short-run effect estimates for clarity
sre <- c(sre, effects[25])

dem_shortrun <- coef(model_iv_2)["dem"]
lag1 <- coef(model_iv_2)[2]
lag2 <- coef(model_iv_2)[3]
lag3 <- coef(model_iv_2)[4]
lag4 <- coef(model_iv_2)[5]
effect1 <- dem_shortrun
effect2 <- effect1 * lag1 + dem_shortrun
effect3 <- effect2 * lag1 + effect1 * lag2 + dem_shortrun
effect4 <- effect3 * lag1 + effect2 * lag2 +
  effect1 * lag3 + dem_shortrun
effects <- c(effect1, effect2, effect3, effect4)
for (i in 5:30) {
  eff <- effects[i - 1] * lag1 +
    effects[i - 2] * lag2 +
    effects[i - 3] * lag3 +
    effects[i - 4] * lag4 + dem_shortrun
  effects <- c(effects, eff)
}
sre <- c(sre, effects[25])

dem_shortrun <- coef(model_iv_3)["dem"]
lag1 <- coef(model_iv_3)[2]
lag2 <- coef(model_iv_3)[3]
lag3 <- coef(model_iv_3)[4]
lag4 <- coef(model_iv_3)[5]
effect1 <- dem_shortrun
effect2 <- effect1 * lag1 + dem_shortrun
effect3 <- effect2 * lag1 + effect1 * lag2 + dem_shortrun
effect4 <- effect3 * lag1 + effect2 * lag2 +
  effect1 * lag3 + dem_shortrun
effects <- c(effect1, effect2, effect3, effect4)
for (i in 5:30) {
  eff <- effects[i - 1] * lag1 +
    effects[i - 2] * lag2 +
    effects[i - 3] * lag3 +
    effects[i - 4] * lag4 + dem_shortrun
  effects <- c(effects, eff)
}
sre <- c(sre, effects[25])

dem_shortrun <- coef(model_iv_4)["dem"]
lag1 <- coef(model_iv_4)[2]
lag2 <- coef(model_iv_4)[3]
```

```
lag3 <- coef(model_iv_4)[4]
lag4 <- coef(model_iv_4)[5]
effect1 <- dem_shortrun
effect2 <- effect1 * lag1 + dem_shortrun
effect3 <- effect2 * lag1 + effect1 * lag2 + dem_shortrun
effect4 <- effect3 * lag1 + effect2 * lag2 +
  effect1 * lag3 + dem_shortrun
effects <- c(effect1, effect2, effect3, effect4)
for (i in 5:30) {
  eff <- effects[i - 1] * lag1 +
    effects[i - 2] * lag2 +
    effects[i - 3] * lag3 +
    effects[i - 4] * lag4 + dem_shortrun
  effects <- c(effects, eff)
}
sre <- c(sre, effects[25])

sre <- round(sre, 3)

# Compute the persistence of GDP
pers1 <- sum(coef(model_iv_1)[2:5])
pers2 <- sum(coef(model_iv_2)[2:5])
pers3 <- sum(coef(model_iv_3)[2:5])
pers4 <- sum(coef(model_iv_4)[2:5])
pers <- round(c(pers1, pers2, pers3, pers4), 3)
```

2.6.3 Tabulation

We format the results of the IV estimation into a LaTeX-formatted table.

First, the estimated coefficients and standard errors of democratization (dem) for each model are obtained to clarify the values to be displayed in the table. Then, a table is created that summarizes the results of the different estimation models so that the effect of democratization can be compared more clearly.

```
# Override coefficients for LaTeX table
override.coef.1 <- coef(model_iv_1)["dem", drop = FALSE]
override.coef.2 <- coef(model_iv_2)["dem", drop = FALSE]
override.coef.3 <- coef(model_iv_3)["dem", drop = FALSE]
override.coef.4 <- coef(model_iv_4)["dem", drop = FALSE]
override.se.1 <- sqrt(diag(vcov(model_iv_1)))["dem"]
override.se.2 <- sqrt(diag(vcov(model_iv_2)))["dem"]
override.se.3 <- sqrt(diag(vcov(model_iv_3)))["dem"]
override.se.4 <- sqrt(diag(vcov(model_iv_4)))["dem"]
models <- list(model_iv_1, model_iv_2, model_iv_3, model_iv_4)

# Generate LaTeX table
texreg(
  models,
  override.coef = list(
    override.coef.1,
    override.coef.2,
    override.coef.3,
    override.coef.4
```

	1 Lag	4 Lags	Soviet Dummies	Regional Trends
Democracy	0.97	1.15	1.29	1.70*
	(0.61)	(0.61)	(0.67)	(0.78)
Persistence	0.96	0.96	0.96	0.95
Long run effect	26.32	31.52	35.72	36.79
Effect after 25 years	20.84	24.87	27.93	32.05
Num. obs.	6312	6309	6309	6309

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 4: (Originally Table.6) IV Estimate of the Effect of Democracy on (Log) GDP per Capita

```
  ),
  override.se = list(
    override.se.1,
    override.se.2,
    override.se.3,
    override.se.4
  ),
  custom.model.names = c(
    "1 Lag", "4 Lags",
    "Soviet Dummies",
    "Regional Trends"
  ),
  custom.coef.map = list(dem = "Democracy"),
  custom.gof.rows = list(
    "Persistence" = pers,
    "Long run effect" = lre,
    "Effect after 25 years" = sre
  ),
  file = "output/table_6_iv.tex",
  caption = "(Originally Table.6) IV Estimate of the Effect of Democracy on (Log) GDP per Capita",
  include.rsquared = FALSE,
  include.adjrs = FALSE,
  include.fstat = FALSE
)
```

3 Extention

3.1 Confidence Interval by the Bootstrap Method (Shoya Abe)

In Figure 1 of the original paper, confidence intervals are not presented. We employ the bootstrap method to derive the confidence interval for the estimated ATT. This allows us to visualize the uncertainty associated with the estimated ATT.

3.1.1 Bootstrap Method

We explain the bootstrap method used in our analysis. The bootstrap method is a computational simulation technique that allows us to estimate the distribution of a statistic in a finite sample. The procedure for deriving a confidence interval using the bootstrap method is as follows:

1. Randomly draw n observations with replacement from the original sample to generate n bootstrap samples.
2. Estimate the ATT for each bootstrap sample.
3. Compute the standard error of the ATT estimates obtained from the bootstrap samples.
4. Use this standard error to estimate the confidence interval.

Here, we derive the confidence interval using two different methods. The first method assumes that the distribution of the estimated ATT follows a normal distribution and estimates the confidence interval using the 2.5% and 97.5% percentiles. This corresponds to the light blue-shaded interval in Figure 3. The second method estimates the confidence interval using the 2.5% and 97.5% percentiles of the bootstrap distribution. This corresponds to the pink-shaded interval in Figure 3.

3.1.2 Estimation

We estimate the confidence interval by executing the following code. The number of bootstrap replications is 200.

```r
# Function to comute ATT estimates for a given bootstrap sample
compute_atets <- function(data_boot) {
  original_data <- data_f1
   # Store the originial dataset to restore later
  data_f1 <<- data_boot
   # Teporarily replace the dataset with the bootstrap sample
  out <- numeric(length(relative_times))
  for (i in seq_along(relative_times)) {
    # Loop through each time period relative to democratization
    t_val <- relative_times[i]
    # Define column names for GDP differences based on the relative time
    if (t_val < 0) {
      col_name <- paste0("gdpDiff_m", abs(t_val))
    } else {
      col_name <- if (t_val == 0) "gdpDiff_0" else paste0("gdpDiff_p", t_val)
    }
    out[i] <- estimateATT(col_name)
     # Compute ATT for the given time period
  }
  data_f1 <<- original_data # Restore the original dataset
  out
}

B <- 200
set.seed(123)

# Initialize a matrix to store bootstrap estimates
boot_mat <- matrix(NA, nrow = B, ncol = length(relative_times))

# Extract unique country IDs for bootstrap resampling
unique_ids <- unique(data_f1$id)

# Perform bootstrap resampling
for (b in seq_len(B)) {# Iterate over B bootstrap samples
```

```r
  sampled_ids <- sample(unique_ids, size = length(unique_ids), replace = TRUE) # Sample country IDs wit
  # Create a bootstrap sample by selecting data corresponding to the sampled IDs
  bs_data <- lapply(sampled_ids, function(x) {
    data_f1[data_f1$id == x, ]
  }) |> bind_rows()
  # Compute ATT estimates for the bootstrap sample
  boot_mat[b, ] <- compute_atets(bs_data)
}

# Compute standard errors from bootstrap samples
boot_se <- apply(boot_mat, 2, sd, na.rm = TRUE)

# Compute normal-based confidence intervals (mean ± 1.96 * standard error)
ci_lower_normal <- atets - 1.96 * boot_se
ci_upper_normal <- atets + 1.96 * boot_se

# Compute percentile-based confidence intervals (2.5% and 97.5% quantiles)
ci_lower_perc <- apply(boot_mat, 2, quantile, probs = 0.025, na.rm = TRUE)
ci_upper_perc <- apply(boot_mat, 2, quantile, probs = 0.975, na.rm = TRUE)

# Create a dataframe containing ATT estimates and confidence intervals
results_with_ci <- data.frame(
  RelativeTime = relative_times,
  ATT = atets,
  ciL_normal = ci_lower_normal,
  ciU_normal = ci_upper_normal,
  ciL_perc = ci_lower_perc,
  ciU_perc = ci_upper_perc
)
```

3.1.3 Plot

```r
# Create a plot showing the ATT estimates with confidence intervals
figure_1_withCI <- ggplot(results_with_ci, aes(x = RelativeTime, y = ATT)) +
  geom_line(color = "black") +
  geom_ribbon(aes(ymin = ciL_perc, ymax = ciU_perc), fill = "pink", alpha = 0.3) +
  geom_ribbon(aes(ymin = ciL_normal, ymax = ciU_normal), fill = "skyblue", alpha = 0.3) +
  scale_x_continuous(breaks = seq(-15, 30, 5)) +
  labs(
    x = "Years around Democratization",
    y = "Change in GDP per capita (log points)"
  ) +
  theme_bw()

# Save the plot as a PDF
ggsave("output/figure_1_withCI.pdf",
       figure_1_withCI,
       width = 14,
       height = 8,
       units = "cm")
```

Figure 1 appears to strongly support the claim that "Democracy does cause growth". However, when we

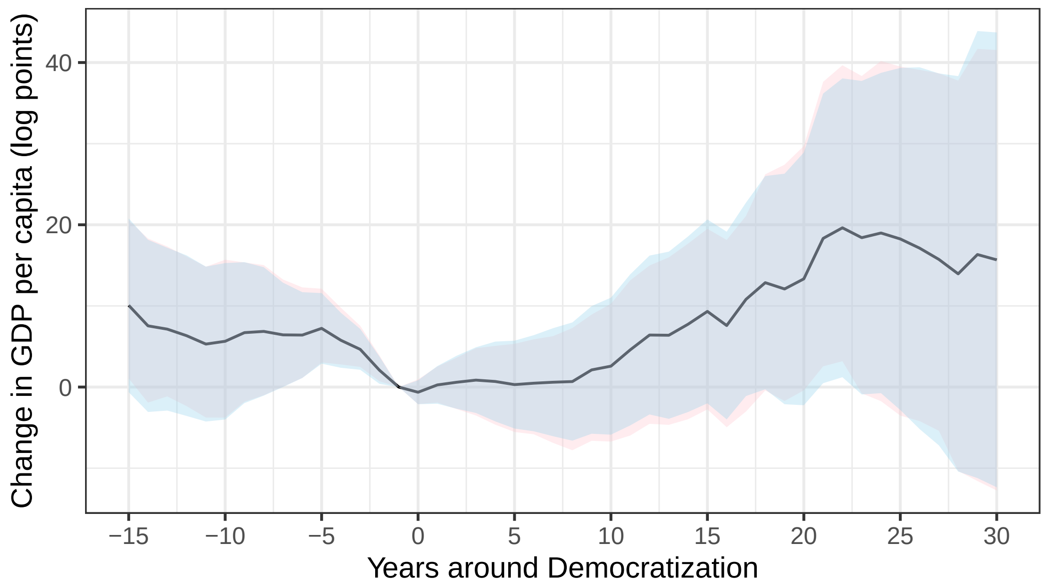

Figure 4: The Long-Term Impact of Democratization on Economic Growth (with the confidence interval)

look at Figure 3, which includes confidence intervals, the picture changes completely. While we do not deny that democratization has a positive effect on economic growth, it becomes clear that the long-term effects of democratization on economic growth are highly uncertain. Perhaps the authors chose not to display the confidence intervals, even if unintentionally, in a way that emphasized the claim that "Democracy does cause growth."

3.2 The relationship between democratization and population (Honoka Otani)

Additional event study analysis was conducted using demographic data that was in the dataset for replication.

Demographics was used because it is considered to be a variable that is strongly related to both democratisation and economic activity. Although the analysis in this paper uses GDP per capita as a variable, changes in demographic composition, for example, cannot be seen by simply dividing by the total population. Therefore, we have decided to look at changes in the proportion of the population aged 14 and under and the proportion of the population aged 15-64.

3.2.1 Change in Population

The event study calculation method is the same as for Replication.

```
data_fx <- data %>%
  rename(id = "_ID") %>%
  group_by(id) %>%
  arrange(year) %>%
  ungroup()

data_fx <- data_fx %>%
  group_by(id) %>%
  arrange(year) %>%
  mutate(prev_dem = dplyr::lag(dem, 1)) %>%
```

```r
  ungroup() %>%
  mutate(transition = case_when(
    dem == 1 & prev_dem == 0 ~ 1,
    dem == 0 & prev_dem == 0 ~ 0,
    TRUE ~ NA_real_
  ))

data_fx <- data_fx %>%
  mutate(pop_log = log(PopulationtotalSPPOPTOTL))

data_fx <- data_fx %>%
  group_by(id) %>%
  arrange(year) %>%
  mutate(
    lag1 = dplyr::lag(pop_log, 1),
    lag2 = dplyr::lag(pop_log, 2),
    lag3 = dplyr::lag(pop_log, 3),
    lag4 = dplyr::lag(pop_log, 4)
  ) %>%
  ungroup() %>%
  filter(!is.na(lag1) & !is.na(lag2) & !is.na(lag3) & !is.na(lag4))

for (t in -15:-2) {
  col_name <- paste0("popDiff_m", abs(t))
  data_fx <- data_fx %>%
    group_by(id) %>%
    arrange(year) %>%
    mutate(!!col_name := dplyr::lag(pop_log, abs(t)) - lag1) %>%
    ungroup()
}

data_fx <- data_fx %>%
  mutate(popDiff_m1 = 0)

data_fx <- data_fx %>%
  group_by(id) %>%
  arrange(year) %>%
  mutate(popDiff_0 = pop_log - lag1) %>%
  ungroup()

for (t in 1:30) {
  col_name <- paste0("popDiff_p", t)
  data_fx <- data_fx %>%
    group_by(id) %>%
    arrange(year) %>%
    mutate(!!col_name := dplyr::lead(pop_log, t) - lag1) %>%
    ungroup()
}

data_fx <- data_fx %>% filter(!is.na(transition))

estimateATT <- function(dataset, outcome_col) {
  sub_data <- dataset %>%
```

```r
    filter(!is.na(.data[[outcome_col]]), !is.na(transition))

  if(nrow(sub_data) == 0) return(NA)

  year_levels <- sort(unique(sub_data$year))
  sub_data <- sub_data %>%
    mutate(year_factor = factor(year, levels = year_levels))

  control_data <- sub_data %>% filter(transition == 0)
  treated_data <- sub_data %>% filter(transition == 1)

  if(nrow(control_data) < 2 || length(unique(control_data$year)) < 2) return(NA)

  model_formula <- as.formula(paste(outcome_col, "~ year_factor - 1"))
  control_model <- tryCatch(lm(model_formula, data = control_data),
                            error = function(e) NULL)
  if(is.null(control_model)) return(NA)

  predicted_outcomes <- tryCatch(predict(control_model, newdata = treated_data),
                                 error = function(e) rep(NA, nrow(treated_data)))

  treatment_effects <- treated_data[[outcome_col]] - predicted_outcomes
  mean(treatment_effects, na.rm = TRUE)
}

relative_times <- c(seq(-15, -1), seq(0, 30))
atts <- numeric(length(relative_times))

for (i in seq_along(relative_times)) {
  t_val <- relative_times[i]
  if(t_val < 0) {
    col_name <- paste0("popDiff_m", abs(t_val))
  } else {
    col_name <- if(t_val == 0) "popDiff_0" else paste0("popDiff_p", t_val)
  }
  atts[i] <- estimateATT(data_fx, col_name)
}

results_df <- data.frame(RelativeTime = relative_times, ATT = atts)

figure_population <- ggplot(results_df, aes(x = RelativeTime, y = ATT)) +
  geom_line(color = "black") +
  scale_x_continuous(breaks = seq(-15, 30, 5)) +
  labs(x = "Years around Democratization",
       y = "Change in Population (log points)") +
  theme_bw()

ggsave("output/figure_population.pdf",
       figure_population,
       width = 14,
       height = 8,
       units = "cm")
```

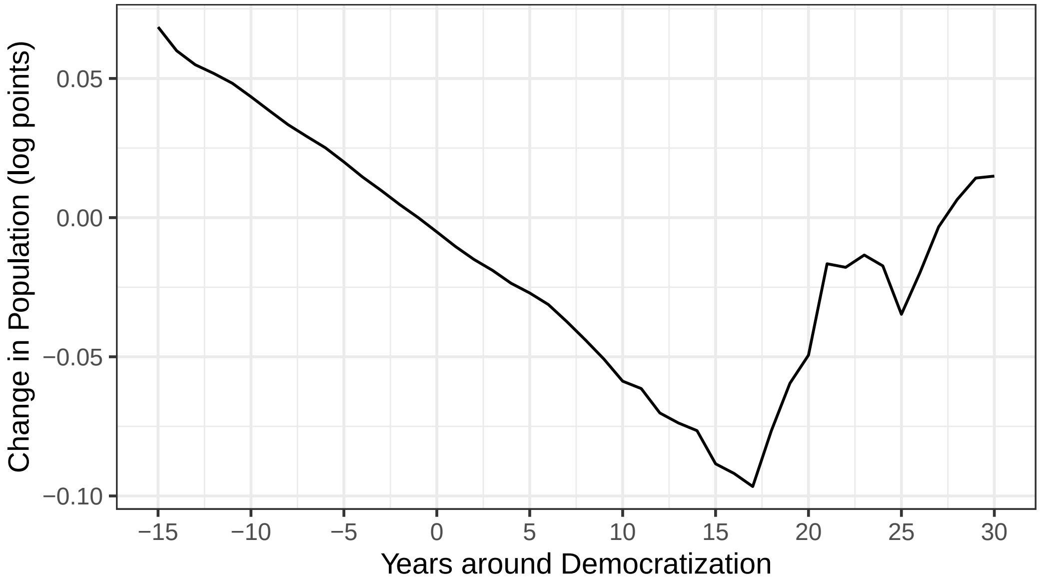

Figure 5: The relationship between democratization and population growth

3.2.2 Population structure

The event study calculation method is the same as for Replication.

```
data_fx_1 <- data %>%
  rename(id = "_ID") %>%
  group_by(id) %>%
  arrange(year) %>%
  ungroup()

data_fx_1 <- data_fx_1 %>%
  group_by(id) %>%
  arrange(year) %>%
  mutate(prev_dem = dplyr::lag(dem, 1)) %>%
  ungroup() %>%
  mutate(transition = case_when(
    dem == 1 & prev_dem == 0 ~ 1,
    dem == 0 & prev_dem == 0 ~ 0,
    TRUE ~ NA_real_
  ))

data_fx_1 <- data_fx_1 %>%
  group_by(id) %>%
  arrange(year) %>%
  mutate(
    lag1 = dplyr::lag(Populationages1564oftota, 1),
    lag2 = dplyr::lag(Populationages1564oftota, 2),
    lag3 = dplyr::lag(Populationages1564oftota, 3),
    lag4 = dplyr::lag(Populationages1564oftota, 4)
  ) %>%
  ungroup() %>%
```

```r
  filter(!is.na(lag1) & !is.na(lag2) & !is.na(lag3) & !is.na(lag4))

for (t in -15:-2) {
  col_name <- paste0("age1564Diff_m", abs(t))
  data_fx_1 <- data_fx_1 %>%
    group_by(id) %>%
    arrange(year) %>%
    mutate(!!col_name := dplyr::lag(Populationages1564oftota, abs(t)) - lag1) %>%
    ungroup()
}

data_fx_1 <- data_fx_1 %>%
  mutate(age1564Diff_m1 = 0)

data_fx_1 <- data_fx_1 %>%
  group_by(id) %>%
  arrange(year) %>%
  mutate(age1564Diff_0 = Populationages1564oftota - lag1) %>%
  ungroup()

for (t in 1:30) {
  col_name <- paste0("age1564Diff_p", t)
  data_fx_1 <- data_fx_1 %>%
    group_by(id) %>%
    arrange(year) %>%
    mutate(!!col_name := dplyr::lead(Populationages1564oftota, t) - lag1) %>%
    ungroup()
}

data_fx_1 <- data_fx_1 %>% filter(!is.na(transition))

estimateATT <- function(dataset, outcome_col) {
  sub_data <- dataset %>%
    filter(!is.na(.data[[outcome_col]]), !is.na(transition))
  if (nrow(sub_data) == 0) return(NA)
  year_levels <- sort(unique(sub_data$year))
  sub_data <- sub_data %>%
    mutate(year_factor = factor(year, levels = year_levels))
  control_data <- sub_data %>% filter(transition == 0)
  treated_data <- sub_data %>% filter(transition == 1)
  if(nrow(control_data) < 2 || length(unique(control_data$year)) < 2) return(NA)
  model_formula <- as.formula(paste(outcome_col, "~ year_factor - 1"))
  control_model <- tryCatch(lm(model_formula, data = control_data),
                            error = function(e) NULL)
  if (is.null(control_model)) return(NA)
  predicted_outcomes <- tryCatch(predict(control_model, newdata = treated_data),
                                 error = function(e) rep(NA, nrow(treated_data)))
  treatment_effects <- treated_data[[outcome_col]] - predicted_outcomes
  mean(treatment_effects, na.rm = TRUE)
}

relative_times <- c(seq(-15, -1), seq(0, 30))
atts_1 <- numeric(length(relative_times))
```

```
for (i in seq_along(relative_times)) {
  t_val <- relative_times[i]
  if (t_val < 0) {
    col_name <- paste0("age1564Diff_m", abs(t_val))
  } else {
    col_name <- if(t_val == 0) "age1564Diff_0" else paste0("age1564Diff_p", t_val)
  }
  atts_1[i] <- estimateATT(data_fx_1, col_name)
}

results_df_1 <- data.frame(RelativeTime = relative_times, ATT = atts_1)

figure_populationages1564oftota <- ggplot(results_df_1, aes(x = RelativeTime, y = ATT)) +
  geom_line(color = "black") +
  scale_x_continuous(breaks = seq(-15, 30, 5)) +
  labs(x = "Years around Democratization",
       y = "Change in Population 15-64 Age") +
  theme_bw()

data_fx_2 <- data %>%
  rename(id = "_ID") %>%
  group_by(id) %>%
  arrange(year) %>%
  ungroup()

data_fx_2 <- data_fx_2 %>%
  group_by(id) %>%
  arrange(year) %>%
  mutate(prev_dem = dplyr::lag(dem, 1)) %>%
  ungroup() %>%
  mutate(transition = case_when(
    dem == 1 & prev_dem == 0 ~ 1,
    dem == 0 & prev_dem == 0 ~ 0,
    TRUE ~ NA_real_
  ))

data_fx_2 <- data_fx_2 %>%
  group_by(id) %>%
  arrange(year) %>%
  mutate(
    lag1 = dplyr::lag(Populationages014oftotal, 1),
    lag2 = dplyr::lag(Populationages014oftotal, 2),
    lag3 = dplyr::lag(Populationages014oftotal, 3),
    lag4 = dplyr::lag(Populationages014oftotal, 4)
  ) %>%
  ungroup() %>%
  filter(!is.na(lag1) & !is.na(lag2) & !is.na(lag3) & !is.na(lag4))

for (t in -15:-2) {
  col_name <- paste0("ageDiff_m", abs(t))
  data_fx_2 <- data_fx_2 %>%
    group_by(id) %>%
    arrange(year) %>%
```

```r
    mutate(!!col_name := dplyr::lag(Populationages014oftotal, abs(t)) - lag1) %>%
    ungroup()
}

data_fx_2 <- data_fx_2 %>%
  mutate(ageDiff_m1 = 0)

data_fx_2 <- data_fx_2 %>%
  group_by(id) %>%
  arrange(year) %>%
  mutate(ageDiff_0 = Populationages014oftotal - lag1) %>%
  ungroup()

for (t in 1:30) {
  col_name <- paste0("ageDiff_p", t)
  data_fx_2 <- data_fx_2 %>%
    group_by(id) %>%
    arrange(year) %>%
    mutate(!!col_name := dplyr::lead(Populationages014oftotal, t) - lag1) %>%
    ungroup()
}

data_fx_2 <- data_fx_2 %>% filter(!is.na(transition))

atts_2 <- numeric(length(relative_times))

for (i in seq_along(relative_times)) {
  t_val <- relative_times[i]
  if (t_val < 0) {
    col_name <- paste0("ageDiff_m", abs(t_val))
  } else {
    col_name <- if(t_val == 0) "ageDiff_0" else paste0("ageDiff_p", t_val)
  }
  atts_2[i] <- estimateATT(data_fx_2, col_name)
}

results_df_2 <- data.frame(RelativeTime = relative_times, ATT = atts_2)

figure_populationages014oftotal <- ggplot(results_df_2, aes(x = RelativeTime, y = ATT)) +
  geom_line(color = "black") +
  scale_x_continuous(breaks = seq(-15, 30, 5)) +
  labs(x = "Years around Democratization",
       y = "Change in Population 0-14 Age ") +
  theme_bw()

figure_combined <- figure_populationages1564oftota + figure_populationages014oftotal

ggsave("output/figure_combined.pdf",
       figure_combined,
       width = 14,
       height = 8,
       units = "cm")
```

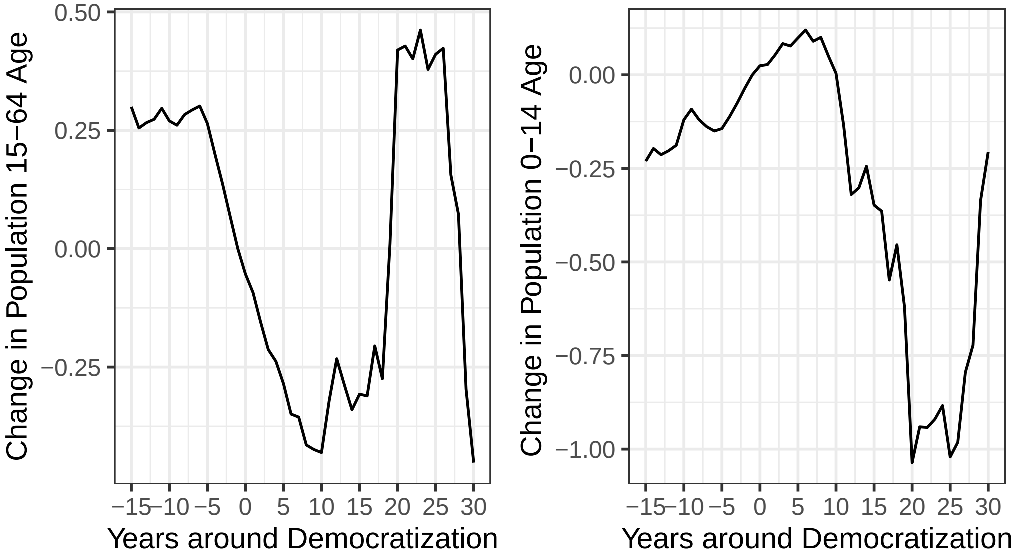

Figure 6: The relationship between democratization and population growth

The percentage of the population aged 14 years and under and the percentage of the population aged 15-64 years show that there are long-term contrasting changes before and after democratisation. Due to the time-consuming nature of the bootstrap method calculations, confidence intervals were not estimated.

The decline in the proportion of the population aged 0-14 from five years after democratisation to 25 years after democratisation suggests that the fertility rate may have declined in the immediate post-democratisation period. This is often attributed to the economic instability and urbanisation associated with pre- and post-democratisation, as well as the entry of women into the labour market. The proportion increases again after 30 years, which may be influenced by long-term political and economic stability due to the time that has passed since democratisation.

In addition, the proportion of the working-age population aged 15-64 decreased from five to ten years before democratisation and has since been recovering. The decline in the working-age population ratio from five years before to ten years after democratisation may have been influenced by the decline in the fertility rate due to political and economic instability prior to democratisation, the outflow of young people from the country and the ageing of the population. Subsequently, the proportion of the working-age population may have recovered due to economic stabilisation, a recovery in the fertility rate and the return of immigrants.

For a more detailed interpretation, fertility data and changes in the proportion of older people (65+) before and after democratisation could be added. Also, by looking at the heterogeneity of effects across countries, it is possible to ascertain whether this is a universal trend.

In any case, the results are likely to be strongly correlated with the effect of democratization on GDP per capita analyzed in this paper, providing modest support for causal effects and the mechanisms examined within the paper.

4 References

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A Robinson. 2019. "Democracy Does Cause Growth." *Journal of Political Economy* 127 (1): 47–100. https://doi.org/10.1086/700936.

Hansen, Bruce. 2022. *Econometrics*. Princeton University Press.

Imai, Kosuke, In Song Kim, and Erik H Wang. 2023. "Matching Methods for Causal Inference with Time-Series Cross-Sectional Data." *American Journal of Political Science* 67 (3): 587–605. https://doi.org/10.1111/ajps.12685.

5 Appendix

5.1 List of Variables (Shoya Abe)

```r
# Extract variable labels from the dataset
var_labels <- sapply(data, function(x) attr(x, "label"))

# Create a tibble containing variable names and their labels
list_var <- tibble(
  variable = names(var_labels),
  label = var_labels
)

# Generate a LaTeX-formatted table listing all variables
kable(
  list_var,
  format = "latex",
  booktabs = TRUE,
  longtable = TRUE,
  caption = "List of Variables"
) |>
  kable_styling(latex_options = "repeat_header")
```

Table 5: List of Variables

variable	label
country_name	Country name
wbcode	World Bank country code
year	Year (from 1960 to 2010)
gdppercapitaconstant2000us	GDP per capita (constant 2000 US$, from World Bank)
lp_bl	Percentage of population with at most primary (Barro-Lee)
ls_bl	Percentage of population with at most secondary (Barro-Lee)
lh_bl	Percentage of population with tertiary education (Barro-Lee)
taxratio	Tax revenue as a share of GDP (from Hendrix)
region	Geographical region
wbcode2	Generated numeric country code
demCGV	Democracy measure by CGV
demBMR	Democracy measure by BMR
yeardem	Identifier for a democratization during this year
yearrev	Identifier for a reversal to autocracy during this year
secenr	Secondary enrollment from World bank
prienr	Primary enrollment from World Bank
tradewb	Exports plus Imports as a share of GDP from World Bank
mortnew	Child mortality per 1000 births from World Bank
ginv	Gross investment as a share of GDP
rtfpna	TFP at constant national prices (2005=1) from PWT
y	log of GDP per capita in 2000 constant dollars (multiplied by a 100)
dem	Democracy measure by ANRR
yy1	year== 1960.0000
yy2	year== 1961.0000

Table 5: List of Variables *(continued)*

variable	label
yy3	year== 1962.0000
yy4	year== 1963.0000
yy5	year== 1964.0000
yy6	year== 1965.0000
yy7	year== 1966.0000
yy8	year== 1967.0000
yy9	year== 1968.0000
yy10	year== 1969.0000
yy11	year== 1970.0000
yy12	year== 1971.0000
yy13	year== 1972.0000
yy14	year== 1973.0000
yy15	year== 1974.0000
yy16	year== 1975.0000
yy17	year== 1976.0000
yy18	year== 1977.0000
yy19	year== 1978.0000
yy20	year== 1979.0000
yy21	year== 1980.0000
yy22	year== 1981.0000
yy23	year== 1982.0000
yy24	year== 1983.0000
yy25	year== 1984.0000
yy26	year== 1985.0000
yy27	year== 1986.0000
yy28	year== 1987.0000
yy29	year== 1988.0000
yy30	year== 1989.0000
yy31	year== 1990.0000
yy32	year== 1991.0000
yy33	year== 1992.0000
yy34	year== 1993.0000
yy35	year== 1994.0000
yy36	year== 1995.0000
yy37	year== 1996.0000
yy38	year== 1997.0000
yy39	year== 1998.0000
yy40	year== 1999.0000
yy41	year== 2000.0000
yy42	year== 2001.0000
yy43	year== 2002.0000
yy44	year== 2003.0000
yy45	year== 2004.0000
yy46	year== 2005.0000
yy47	year== 2006.0000
yy48	year== 2007.0000

Table 5: List of Variables *(continued)*

variable	label
yy49	year== 2008.0000
yy50	year== 2009.0000
yy51	year== 2010.0000
InitReg	Democratic status after independence or in 1960
unrest	Occurrence of events of unrest (from Banks CNTS)
loginvpc	log investment (multiplied by 100)
ltfp	log TFP (multiplied by 100)
ltrade2	lof of trade (multiplied by 100)
lprienr	lof of primary enrollment (multiplied by 100)
lsecenr	log of secondary enrollment (multiplied by 100)
lgov	log of taxes to GDP (multiplied by a 100)
lmort	log of child mortality rate (multiplied by a 100)
unrestn	Likelihood of unrest (0-100 scale)
demFH	democracy measure based on Freedom House
demPOL	democracy measure based on Polity IV
demPS	democracy measure by PS
demPOL__xconst	dummy for constraints on executive (based on polity)
demPOL__parcomp	dummy for competitiveness of participation (based on polity)
demPOL__exrec	dummy for quality of executive recruitment process (based on Polity)
demFH__pr	Dummy for political rights (based on Freedom House)
demFH__cl	Dummy for civil liberties (based on Freedom House)
demevent	Event of democratization
revevent	Event of reversal to autocracy
democ	Cummulative number of democratizations
rever	Cummulative number of reversals
demext	Democratic status at beginning of sample
regionINITREG	Region/Initial regime at start of sample cells
demreg	Average democracy in the region*initial regime (leaving own country out)
tradewbreg	Regional trade
unrestreg	Regional unrest
yreg	Regional GDP per capita
rtrend1	Region 1 trend
rtrend2	Region 2 trend
rtrend3	Region 3 trend
rtrend4	Region 4 trend
rtrend5	Region 5 trend
rtrend6	Region trend 6
rtrend7	region trend 7
region60	Region/Democratic in 1960 cells
regionDA	Region/Always democratic cells
regionREG	Region/Detailed regime in 1960 cells
demreg60	Average democracy in the region*initial regim (using regime in 1960, jackniffed)
demregDA	Average democracy in the region*initial regim (using always democracy, jackniffe
demregREGIME	Average democracy in the region*initial regime (detailed regimes, jackniffed)
d60__1	region60==AFR__dem
d60__2	region60==AFR__nd

Table 5: List of Variables *(continued)*

variable	label
d60_3	region60==EAP_dem
d60_4	region60==EAP_nd
d60_5	region60==ECA_nd
d60_6	region60==INL_dem
d60_7	region60==INL_nd
d60_8	region60==LAC_dem
d60_9	region60==LAC_nd
d60_10	region60==MNA_dem
d60_11	region60==MNA_nd
d60_12	region60==SAS_dem
d60_13	region60==SAS_nd
dDA_1	regionDA==AFR_dem
dDA_2	regionDA==AFR_nd
dDA_3	regionDA==EAP_dem
dDA_4	regionDA==EAP_nd
dDA_5	regionDA==ECA_nd
dDA_6	regionDA==INL_dem
dDA_7	regionDA==INL_nd
dDA_8	regionDA==LAC_dem
dDA_9	regionDA==LAC_nd
dDA_10	regionDA==MNA_nd
dDA_11	regionDA==SAS_dem
dDA_12	regionDA==SAS_nd
dREG_1	regionREG==AFRBritishColony
dREG_2	regionREG==AFRCivilDictator
dREG_3	regionREG==AFRFrenchColony
dREG_4	regionREG==AFRMilitaryDictator
dREG_5	regionREG==AFRParlamentaryDemocracy
dREG_6	regionREG==AFRRoyalDictator
dREG_7	regionREG==AFRSocialistRegime
dREG_8	regionREG==EAPBritishColony
dREG_9	regionREG==EAPCivilDictator
dREG_10	regionREG==EAPMilitaryDictator
dREG_11	regionREG==EAPMixedAndPresidentialDemocracy
dREG_12	regionREG==EAPRoyalDictator
dREG_13	regionREG==EAPSocialistRegime
dREG_14	regionREG==ECAMilitaryDictator
dREG_15	regionREG==ECASocialistRegime
dREG_16	regionREG==INLCivilDictator
dREG_17	regionREG==INLFrenchColony
dREG_18	regionREG==INLMilitaryDictator
dREG_19	regionREG==INLMixedAndPresidentialDemocracy
dREG_20	regionREG==INLParlamentaryDemocracy
dREG_21	regionREG==LACBritishColony
dREG_22	regionREG==LACFrenchColony
dREG_23	regionREG==LACMilitaryDictator
dREG_24	regionREG==LACMixedAndPresidentialDemocracy

Table 5: List of Variables *(continued)*

variable	label
dREG__25	regionREG==LACSocialistRegime
dREG__26	regionREG==MNABritishColony
dREG__27	regionREG==MNACivilDictator
dREG__28	regionREG==MNAFrenchColony
dREG__29	regionREG==MNAMilitaryDictator
dREG__30	regionREG==MNAParlamentaryDemocracy
dREG__31	regionREG==MNARoyalDictator
dREG__32	regionREG==SASBritishColony
dREG__33	regionREG==SASMilitaryDictator
dREG__34	regionREG==SASParlamentaryDemocracy
dREG__35	regionREG==SASRoyalDictator
gdp1960	GDP per capita in 1960 from Madisson
region__initreg__year	Region/Initial regime/year cells
incomequint50s__year	Income quintiles in 50s/year cells
sov1	Soviets post 89
sov2	Soviets post 90
sov3	Soviets post 91
sov4	Soviets post 92
marketref	Index of market reforms
regdum1	region__initreg__year==AFR__dem1960
regdum2	region__initreg__year==AFR__dem1961
regdum3	region__initreg__year==AFR__dem1962
regdum4	region__initreg__year==AFR__dem1963
regdum5	region__initreg__year==AFR__dem1964
regdum6	region__initreg__year==AFR__dem1965
regdum7	region__initreg__year==AFR__dem1966
regdum8	region__initreg__year==AFR__dem1967
regdum9	region__initreg__year==AFR__dem1968
regdum10	region__initreg__year==AFR__dem1969
regdum11	region__initreg__year==AFR__dem1970
regdum12	region__initreg__year==AFR__dem1971
regdum13	region__initreg__year==AFR__dem1972
regdum14	region__initreg__year==AFR__dem1973
regdum15	region__initreg__year==AFR__dem1974
regdum16	region__initreg__year==AFR__dem1975
regdum17	region__initreg__year==AFR__dem1976
regdum18	region__initreg__year==AFR__dem1977
regdum19	region__initreg__year==AFR__dem1978
regdum20	region__initreg__year==AFR__dem1979
regdum21	region__initreg__year==AFR__dem1980
regdum22	region__initreg__year==AFR__dem1981
regdum23	region__initreg__year==AFR__dem1982
regdum24	region__initreg__year==AFR__dem1983
regdum25	region__initreg__year==AFR__dem1984
regdum26	region__initreg__year==AFR__dem1985
regdum27	region__initreg__year==AFR__dem1986
regdum28	region__initreg__year==AFR__dem1987

Table 5: List of Variables *(continued)*

variable	label
regdum29	region_initreg_year==AFR_dem1988
regdum30	region_initreg_year==AFR_dem1989
regdum31	region_initreg_year==AFR_dem1990
regdum32	region_initreg_year==AFR_dem1991
regdum33	region_initreg_year==AFR_dem1992
regdum34	region_initreg_year==AFR_dem1993
regdum35	region_initreg_year==AFR_dem1994
regdum36	region_initreg_year==AFR_dem1995
regdum37	region_initreg_year==AFR_dem1996
regdum38	region_initreg_year==AFR_dem1997
regdum39	region_initreg_year==AFR_dem1998
regdum40	region_initreg_year==AFR_dem1999
regdum41	region_initreg_year==AFR_dem2000
regdum42	region_initreg_year==AFR_dem2001
regdum43	region_initreg_year==AFR_dem2002
regdum44	region_initreg_year==AFR_dem2003
regdum45	region_initreg_year==AFR_dem2004
regdum46	region_initreg_year==AFR_dem2005
regdum47	region_initreg_year==AFR_dem2006
regdum48	region_initreg_year==AFR_dem2007
regdum49	region_initreg_year==AFR_dem2008
regdum50	region_initreg_year==AFR_dem2009
regdum51	region_initreg_year==AFR_dem2010
regdum52	region_initreg_year==AFR_nd1960
regdum53	region_initreg_year==AFR_nd1961
regdum54	region_initreg_year==AFR_nd1962
regdum55	region_initreg_year==AFR_nd1963
regdum56	region_initreg_year==AFR_nd1964
regdum57	region_initreg_year==AFR_nd1965
regdum58	region_initreg_year==AFR_nd1966
regdum59	region_initreg_year==AFR_nd1967
regdum60	region_initreg_year==AFR_nd1968
regdum61	region_initreg_year==AFR_nd1969
regdum62	region_initreg_year==AFR_nd1970
regdum63	region_initreg_year==AFR_nd1971
regdum64	region_initreg_year==AFR_nd1972
regdum65	region_initreg_year==AFR_nd1973
regdum66	region_initreg_year==AFR_nd1974
regdum67	region_initreg_year==AFR_nd1975
regdum68	region_initreg_year==AFR_nd1976
regdum69	region_initreg_year==AFR_nd1977
regdum70	region_initreg_year==AFR_nd1978
regdum71	region_initreg_year==AFR_nd1979
regdum72	region_initreg_year==AFR_nd1980
regdum73	region_initreg_year==AFR_nd1981
regdum74	region_initreg_year==AFR_nd1982

Table 5: List of Variables *(continued)*

variable	label
regdum75	region__initreg__year==AFR__nd1983
regdum76	region__initreg__year==AFR__nd1984
regdum77	region__initreg__year==AFR__nd1985
regdum78	region__initreg__year==AFR__nd1986
regdum79	region__initreg__year==AFR__nd1987
regdum80	region__initreg__year==AFR__nd1988
regdum81	region__initreg__year==AFR__nd1989
regdum82	region__initreg__year==AFR__nd1990
regdum83	region__initreg__year==AFR__nd1991
regdum84	region__initreg__year==AFR__nd1992
regdum85	region__initreg__year==AFR__nd1993
regdum86	region__initreg__year==AFR__nd1994
regdum87	region__initreg__year==AFR__nd1995
regdum88	region__initreg__year==AFR__nd1996
regdum89	region__initreg__year==AFR__nd1997
regdum90	region__initreg__year==AFR__nd1998
regdum91	region__initreg__year==AFR__nd1999
regdum92	region__initreg__year==AFR__nd2000
regdum93	region__initreg__year==AFR__nd2001
regdum94	region__initreg__year==AFR__nd2002
regdum95	region__initreg__year==AFR__nd2003
regdum96	region__initreg__year==AFR__nd2004
regdum97	region__initreg__year==AFR__nd2005
regdum98	region__initreg__year==AFR__nd2006
regdum99	region__initreg__year==AFR__nd2007
regdum100	region__initreg__year==AFR__nd2008
regdum101	region__initreg__year==AFR__nd2009
regdum102	region__initreg__year==AFR__nd2010
regdum103	region__initreg__year==EAP__dem1960
regdum104	region__initreg__year==EAP__dem1961
regdum105	region__initreg__year==EAP__dem1962
regdum106	region__initreg__year==EAP__dem1963
regdum107	region__initreg__year==EAP__dem1964
regdum108	region__initreg__year==EAP__dem1965
regdum109	region__initreg__year==EAP__dem1966
regdum110	region__initreg__year==EAP__dem1967
regdum111	region__initreg__year==EAP__dem1968
regdum112	region__initreg__year==EAP__dem1969
regdum113	region__initreg__year==EAP__dem1970
regdum114	region__initreg__year==EAP__dem1971
regdum115	region__initreg__year==EAP__dem1972
regdum116	region__initreg__year==EAP__dem1973
regdum117	region__initreg__year==EAP__dem1974
regdum118	region__initreg__year==EAP__dem1975
regdum119	region__initreg__year==EAP__dem1976
regdum120	region__initreg__year==EAP__dem1977
regdum121	region__initreg__year==EAP__dem1978

Table 5: List of Variables *(continued)*

variable	label
regdum122	region_initreg_year==EAP_dem1979
regdum123	region_initreg_year==EAP_dem1980
regdum124	region_initreg_year==EAP_dem1981
regdum125	region_initreg_year==EAP_dem1982
regdum126	region_initreg_year==EAP_dem1983
regdum127	region_initreg_year==EAP_dem1984
regdum128	region_initreg_year==EAP_dem1985
regdum129	region_initreg_year==EAP_dem1986
regdum130	region_initreg_year==EAP_dem1987
regdum131	region_initreg_year==EAP_dem1988
regdum132	region_initreg_year==EAP_dem1989
regdum133	region_initreg_year==EAP_dem1990
regdum134	region_initreg_year==EAP_dem1991
regdum135	region_initreg_year==EAP_dem1992
regdum136	region_initreg_year==EAP_dem1993
regdum137	region_initreg_year==EAP_dem1994
regdum138	region_initreg_year==EAP_dem1995
regdum139	region_initreg_year==EAP_dem1996
regdum140	region_initreg_year==EAP_dem1997
regdum141	region_initreg_year==EAP_dem1998
regdum142	region_initreg_year==EAP_dem1999
regdum143	region_initreg_year==EAP_dem2000
regdum144	region_initreg_year==EAP_dem2001
regdum145	region_initreg_year==EAP_dem2002
regdum146	region_initreg_year==EAP_dem2003
regdum147	region_initreg_year==EAP_dem2004
regdum148	region_initreg_year==EAP_dem2005
regdum149	region_initreg_year==EAP_dem2006
regdum150	region_initreg_year==EAP_dem2007
regdum151	region_initreg_year==EAP_dem2008
regdum152	region_initreg_year==EAP_dem2009
regdum153	region_initreg_year==EAP_dem2010
regdum154	region_initreg_year==EAP_nd1960
regdum155	region_initreg_year==EAP_nd1961
regdum156	region_initreg_year==EAP_nd1962
regdum157	region_initreg_year==EAP_nd1963
regdum158	region_initreg_year==EAP_nd1964
regdum159	region_initreg_year==EAP_nd1965
regdum160	region_initreg_year==EAP_nd1966
regdum161	region_initreg_year==EAP_nd1967
regdum162	region_initreg_year==EAP_nd1968
regdum163	region_initreg_year==EAP_nd1969
regdum164	region_initreg_year==EAP_nd1970
regdum165	region_initreg_year==EAP_nd1971
regdum166	region_initreg_year==EAP_nd1972
regdum167	region_initreg_year==EAP_nd1973
regdum168	region_initreg_year==EAP_nd1974

Table 5: List of Variables *(continued)*

variable	label
regdum169	region__initreg__year==EAP__nd1975
regdum170	region__initreg__year==EAP__nd1976
regdum171	region__initreg__year==EAP__nd1977
regdum172	region__initreg__year==EAP__nd1978
regdum173	region__initreg__year==EAP__nd1979
regdum174	region__initreg__year==EAP__nd1980
regdum175	region__initreg__year==EAP__nd1981
regdum176	region__initreg__year==EAP__nd1982
regdum177	region__initreg__year==EAP__nd1983
regdum178	region__initreg__year==EAP__nd1984
regdum179	region__initreg__year==EAP__nd1985
regdum180	region__initreg__year==EAP__nd1986
regdum181	region__initreg__year==EAP__nd1987
regdum182	region__initreg__year==EAP__nd1988
regdum183	region__initreg__year==EAP__nd1989
regdum184	region__initreg__year==EAP__nd1990
regdum185	region__initreg__year==EAP__nd1991
regdum186	region__initreg__year==EAP__nd1992
regdum187	region__initreg__year==EAP__nd1993
regdum188	region__initreg__year==EAP__nd1994
regdum189	region__initreg__year==EAP__nd1995
regdum190	region__initreg__year==EAP__nd1996
regdum191	region__initreg__year==EAP__nd1997
regdum192	region__initreg__year==EAP__nd1998
regdum193	region__initreg__year==EAP__nd1999
regdum194	region__initreg__year==EAP__nd2000
regdum195	region__initreg__year==EAP__nd2001
regdum196	region__initreg__year==EAP__nd2002
regdum197	region__initreg__year==EAP__nd2003
regdum198	region__initreg__year==EAP__nd2004
regdum199	region__initreg__year==EAP__nd2005
regdum200	region__initreg__year==EAP__nd2006
regdum201	region__initreg__year==EAP__nd2007
regdum202	region__initreg__year==EAP__nd2008
regdum203	region__initreg__year==EAP__nd2009
regdum204	region__initreg__year==EAP__nd2010
regdum205	region__initreg__year==ECA__nd1960
regdum206	region__initreg__year==ECA__nd1961
regdum207	region__initreg__year==ECA__nd1962
regdum208	region__initreg__year==ECA__nd1963
regdum209	region__initreg__year==ECA__nd1964
regdum210	region__initreg__year==ECA__nd1965
regdum211	region__initreg__year==ECA__nd1966
regdum212	region__initreg__year==ECA__nd1967
regdum213	region__initreg__year==ECA__nd1968
regdum214	region__initreg__year==ECA__nd1969

Table 5: List of Variables *(continued)*

variable	label
regdum215	region_initreg_year==ECA_nd1970
regdum216	region_initreg_year==ECA_nd1971
regdum217	region_initreg_year==ECA_nd1972
regdum218	region_initreg_year==ECA_nd1973
regdum219	region_initreg_year==ECA_nd1974
regdum220	region_initreg_year==ECA_nd1975
regdum221	region_initreg_year==ECA_nd1976
regdum222	region_initreg_year==ECA_nd1977
regdum223	region_initreg_year==ECA_nd1978
regdum224	region_initreg_year==ECA_nd1979
regdum225	region_initreg_year==ECA_nd1980
regdum226	region_initreg_year==ECA_nd1981
regdum227	region_initreg_year==ECA_nd1982
regdum228	region_initreg_year==ECA_nd1983
regdum229	region_initreg_year==ECA_nd1984
regdum230	region_initreg_year==ECA_nd1985
regdum231	region_initreg_year==ECA_nd1986
regdum232	region_initreg_year==ECA_nd1987
regdum233	region_initreg_year==ECA_nd1988
regdum234	region_initreg_year==ECA_nd1989
regdum235	region_initreg_year==ECA_nd1990
regdum236	region_initreg_year==ECA_nd1991
regdum237	region_initreg_year==ECA_nd1992
regdum238	region_initreg_year==ECA_nd1993
regdum239	region_initreg_year==ECA_nd1994
regdum240	region_initreg_year==ECA_nd1995
regdum241	region_initreg_year==ECA_nd1996
regdum242	region_initreg_year==ECA_nd1997
regdum243	region_initreg_year==ECA_nd1998
regdum244	region_initreg_year==ECA_nd1999
regdum245	region_initreg_year==ECA_nd2000
regdum246	region_initreg_year==ECA_nd2001
regdum247	region_initreg_year==ECA_nd2002
regdum248	region_initreg_year==ECA_nd2003
regdum249	region_initreg_year==ECA_nd2004
regdum250	region_initreg_year==ECA_nd2005
regdum251	region_initreg_year==ECA_nd2006
regdum252	region_initreg_year==ECA_nd2007
regdum253	region_initreg_year==ECA_nd2008
regdum254	region_initreg_year==ECA_nd2009
regdum255	region_initreg_year==ECA_nd2010
regdum256	region_initreg_year==INL_dem1960
regdum257	region_initreg_year==INL_dem1961
regdum258	region_initreg_year==INL_dem1962
regdum259	region_initreg_year==INL_dem1963
regdum260	region_initreg_year==INL_dem1964
regdum261	region_initreg_year==INL_dem1965

Table 5: List of Variables *(continued)*

variable	label
regdum262	region__initreg__year==INL__dem1966
regdum263	region__initreg__year==INL__dem1967
regdum264	region__initreg__year==INL__dem1968
regdum265	region__initreg__year==INL__dem1969
regdum266	region__initreg__year==INL__dem1970
regdum267	region__initreg__year==INL__dem1971
regdum268	region__initreg__year==INL__dem1972
regdum269	region__initreg__year==INL__dem1973
regdum270	region__initreg__year==INL__dem1974
regdum271	region__initreg__year==INL__dem1975
regdum272	region__initreg__year==INL__dem1976
regdum273	region__initreg__year==INL__dem1977
regdum274	region__initreg__year==INL__dem1978
regdum275	region__initreg__year==INL__dem1979
regdum276	region__initreg__year==INL__dem1980
regdum277	region__initreg__year==INL__dem1981
regdum278	region__initreg__year==INL__dem1982
regdum279	region__initreg__year==INL__dem1983
regdum280	region__initreg__year==INL__dem1984
regdum281	region__initreg__year==INL__dem1985
regdum282	region__initreg__year==INL__dem1986
regdum283	region__initreg__year==INL__dem1987
regdum284	region__initreg__year==INL__dem1988
regdum285	region__initreg__year==INL__dem1989
regdum286	region__initreg__year==INL__dem1990
regdum287	region__initreg__year==INL__dem1991
regdum288	region__initreg__year==INL__dem1992
regdum289	region__initreg__year==INL__dem1993
regdum290	region__initreg__year==INL__dem1994
regdum291	region__initreg__year==INL__dem1995
regdum292	region__initreg__year==INL__dem1996
regdum293	region__initreg__year==INL__dem1997
regdum294	region__initreg__year==INL__dem1998
regdum295	region__initreg__year==INL__dem1999
regdum296	region__initreg__year==INL__dem2000
regdum297	region__initreg__year==INL__dem2001
regdum298	region__initreg__year==INL__dem2002
regdum299	region__initreg__year==INL__dem2003
regdum300	region__initreg__year==INL__dem2004
regdum301	region__initreg__year==INL__dem2005
regdum302	region__initreg__year==INL__dem2006
regdum303	region__initreg__year==INL__dem2007
regdum304	region__initreg__year==INL__dem2008
regdum305	region__initreg__year==INL__dem2009
regdum306	region__initreg__year==INL__dem2010
regdum307	region__initreg__year==INL__nd1960
regdum308	region__initreg__year==INL__nd1961

Table 5: List of Variables *(continued)*

variable	label
regdum309	region__initreg__year==INL__nd1962
regdum310	region__initreg__year==INL__nd1963
regdum311	region__initreg__year==INL__nd1964
regdum312	region__initreg__year==INL__nd1965
regdum313	region__initreg__year==INL__nd1966
regdum314	region__initreg__year==INL__nd1967
regdum315	region__initreg__year==INL__nd1968
regdum316	region__initreg__year==INL__nd1969
regdum317	region__initreg__year==INL__nd1970
regdum318	region__initreg__year==INL__nd1971
regdum319	region__initreg__year==INL__nd1972
regdum320	region__initreg__year==INL__nd1973
regdum321	region__initreg__year==INL__nd1974
regdum322	region__initreg__year==INL__nd1975
regdum323	region__initreg__year==INL__nd1976
regdum324	region__initreg__year==INL__nd1977
regdum325	region__initreg__year==INL__nd1978
regdum326	region__initreg__year==INL__nd1979
regdum327	region__initreg__year==INL__nd1980
regdum328	region__initreg__year==INL__nd1981
regdum329	region__initreg__year==INL__nd1982
regdum330	region__initreg__year==INL__nd1983
regdum331	region__initreg__year==INL__nd1984
regdum332	region__initreg__year==INL__nd1985
regdum333	region__initreg__year==INL__nd1986
regdum334	region__initreg__year==INL__nd1987
regdum335	region__initreg__year==INL__nd1988
regdum336	region__initreg__year==INL__nd1989
regdum337	region__initreg__year==INL__nd1990
regdum338	region__initreg__year==INL__nd1991
regdum339	region__initreg__year==INL__nd1992
regdum340	region__initreg__year==INL__nd1993
regdum341	region__initreg__year==INL__nd1994
regdum342	region__initreg__year==INL__nd1995
regdum343	region__initreg__year==INL__nd1996
regdum344	region__initreg__year==INL__nd1997
regdum345	region__initreg__year==INL__nd1998
regdum346	region__initreg__year==INL__nd1999
regdum347	region__initreg__year==INL__nd2000
regdum348	region__initreg__year==INL__nd2001
regdum349	region__initreg__year==INL__nd2002
regdum350	region__initreg__year==INL__nd2003
regdum351	region__initreg__year==INL__nd2004
regdum352	region__initreg__year==INL__nd2005
regdum353	region__initreg__year==INL__nd2006
regdum354	region__initreg__year==INL__nd2007

Table 5: List of Variables *(continued)*

variable	label
regdum355	region__initreg__year==INL__nd2008
regdum356	region__initreg__year==INL__nd2009
regdum357	region__initreg__year==INL__nd2010
regdum358	region__initreg__year==LAC__dem1960
regdum359	region__initreg__year==LAC__dem1961
regdum360	region__initreg__year==LAC__dem1962
regdum361	region__initreg__year==LAC__dem1963
regdum362	region__initreg__year==LAC__dem1964
regdum363	region__initreg__year==LAC__dem1965
regdum364	region__initreg__year==LAC__dem1966
regdum365	region__initreg__year==LAC__dem1967
regdum366	region__initreg__year==LAC__dem1968
regdum367	region__initreg__year==LAC__dem1969
regdum368	region__initreg__year==LAC__dem1970
regdum369	region__initreg__year==LAC__dem1971
regdum370	region__initreg__year==LAC__dem1972
regdum371	region__initreg__year==LAC__dem1973
regdum372	region__initreg__year==LAC__dem1974
regdum373	region__initreg__year==LAC__dem1975
regdum374	region__initreg__year==LAC__dem1976
regdum375	region__initreg__year==LAC__dem1977
regdum376	region__initreg__year==LAC__dem1978
regdum377	region__initreg__year==LAC__dem1979
regdum378	region__initreg__year==LAC__dem1980
regdum379	region__initreg__year==LAC__dem1981
regdum380	region__initreg__year==LAC__dem1982
regdum381	region__initreg__year==LAC__dem1983
regdum382	region__initreg__year==LAC__dem1984
regdum383	region__initreg__year==LAC__dem1985
regdum384	region__initreg__year==LAC__dem1986
regdum385	region__initreg__year==LAC__dem1987
regdum386	region__initreg__year==LAC__dem1988
regdum387	region__initreg__year==LAC__dem1989
regdum388	region__initreg__year==LAC__dem1990
regdum389	region__initreg__year==LAC__dem1991
regdum390	region__initreg__year==LAC__dem1992
regdum391	region__initreg__year==LAC__dem1993
regdum392	region__initreg__year==LAC__dem1994
regdum393	region__initreg__year==LAC__dem1995
regdum394	region__initreg__year==LAC__dem1996
regdum395	region__initreg__year==LAC__dem1997
regdum396	region__initreg__year==LAC__dem1998
regdum397	region__initreg__year==LAC__dem1999
regdum398	region__initreg__year==LAC__dem2000
regdum399	region__initreg__year==LAC__dem2001
regdum400	region__initreg__year==LAC__dem2002
regdum401	region__initreg__year==LAC__dem2003

Table 5: List of Variables *(continued)*

variable	label
regdum402	region_initreg_year==LAC_dem2004
regdum403	region_initreg_year==LAC_dem2005
regdum404	region_initreg_year==LAC_dem2006
regdum405	region_initreg_year==LAC_dem2007
regdum406	region_initreg_year==LAC_dem2008
regdum407	region_initreg_year==LAC_dem2009
regdum408	region_initreg_year==LAC_dem2010
regdum409	region_initreg_year==LAC_nd1960
regdum410	region_initreg_year==LAC_nd1961
regdum411	region_initreg_year==LAC_nd1962
regdum412	region_initreg_year==LAC_nd1963
regdum413	region_initreg_year==LAC_nd1964
regdum414	region_initreg_year==LAC_nd1965
regdum415	region_initreg_year==LAC_nd1966
regdum416	region_initreg_year==LAC_nd1967
regdum417	region_initreg_year==LAC_nd1968
regdum418	region_initreg_year==LAC_nd1969
regdum419	region_initreg_year==LAC_nd1970
regdum420	region_initreg_year==LAC_nd1971
regdum421	region_initreg_year==LAC_nd1972
regdum422	region_initreg_year==LAC_nd1973
regdum423	region_initreg_year==LAC_nd1974
regdum424	region_initreg_year==LAC_nd1975
regdum425	region_initreg_year==LAC_nd1976
regdum426	region_initreg_year==LAC_nd1977
regdum427	region_initreg_year==LAC_nd1978
regdum428	region_initreg_year==LAC_nd1979
regdum429	region_initreg_year==LAC_nd1980
regdum430	region_initreg_year==LAC_nd1981
regdum431	region_initreg_year==LAC_nd1982
regdum432	region_initreg_year==LAC_nd1983
regdum433	region_initreg_year==LAC_nd1984
regdum434	region_initreg_year==LAC_nd1985
regdum435	region_initreg_year==LAC_nd1986
regdum436	region_initreg_year==LAC_nd1987
regdum437	region_initreg_year==LAC_nd1988
regdum438	region_initreg_year==LAC_nd1989
regdum439	region_initreg_year==LAC_nd1990
regdum440	region_initreg_year==LAC_nd1991
regdum441	region_initreg_year==LAC_nd1992
regdum442	region_initreg_year==LAC_nd1993
regdum443	region_initreg_year==LAC_nd1994
regdum444	region_initreg_year==LAC_nd1995
regdum445	region_initreg_year==LAC_nd1996
regdum446	region_initreg_year==LAC_nd1997
regdum447	region_initreg_year==LAC_nd1998
regdum448	region_initreg_year==LAC_nd1999

Table 5: List of Variables *(continued)*

variable	label
regdum449	region__initreg__year==LAC__nd2000
regdum450	region__initreg__year==LAC__nd2001
regdum451	region__initreg__year==LAC__nd2002
regdum452	region__initreg__year==LAC__nd2003
regdum453	region__initreg__year==LAC__nd2004
regdum454	region__initreg__year==LAC__nd2005
regdum455	region__initreg__year==LAC__nd2006
regdum456	region__initreg__year==LAC__nd2007
regdum457	region__initreg__year==LAC__nd2008
regdum458	region__initreg__year==LAC__nd2009
regdum459	region__initreg__year==LAC__nd2010
regdum460	region__initreg__year==MNA__dem1960
regdum461	region__initreg__year==MNA__dem1961
regdum462	region__initreg__year==MNA__dem1962
regdum463	region__initreg__year==MNA__dem1963
regdum464	region__initreg__year==MNA__dem1964
regdum465	region__initreg__year==MNA__dem1965
regdum466	region__initreg__year==MNA__dem1966
regdum467	region__initreg__year==MNA__dem1967
regdum468	region__initreg__year==MNA__dem1968
regdum469	region__initreg__year==MNA__dem1969
regdum470	region__initreg__year==MNA__dem1970
regdum471	region__initreg__year==MNA__dem1971
regdum472	region__initreg__year==MNA__dem1972
regdum473	region__initreg__year==MNA__dem1973
regdum474	region__initreg__year==MNA__dem1974
regdum475	region__initreg__year==MNA__dem1975
regdum476	region__initreg__year==MNA__dem1976
regdum477	region__initreg__year==MNA__dem1977
regdum478	region__initreg__year==MNA__dem1978
regdum479	region__initreg__year==MNA__dem1979
regdum480	region__initreg__year==MNA__dem1980
regdum481	region__initreg__year==MNA__dem1981
regdum482	region__initreg__year==MNA__dem1982
regdum483	region__initreg__year==MNA__dem1983
regdum484	region__initreg__year==MNA__dem1984
regdum485	region__initreg__year==MNA__dem1985
regdum486	region__initreg__year==MNA__dem1986
regdum487	region__initreg__year==MNA__dem1987
regdum488	region__initreg__year==MNA__dem1988
regdum489	region__initreg__year==MNA__dem1989
regdum490	region__initreg__year==MNA__dem1990
regdum491	region__initreg__year==MNA__dem1991
regdum492	region__initreg__year==MNA__dem1992
regdum493	region__initreg__year==MNA__dem1993
regdum494	region__initreg__year==MNA__dem1994

Table 5: List of Variables *(continued)*

variable	label
regdum495	region_initreg_year==MNA_dem1995
regdum496	region_initreg_year==MNA_dem1996
regdum497	region_initreg_year==MNA_dem1997
regdum498	region_initreg_year==MNA_dem1998
regdum499	region_initreg_year==MNA_dem1999
regdum500	region_initreg_year==MNA_dem2000
regdum501	region_initreg_year==MNA_dem2001
regdum502	region_initreg_year==MNA_dem2002
regdum503	region_initreg_year==MNA_dem2003
regdum504	region_initreg_year==MNA_dem2004
regdum505	region_initreg_year==MNA_dem2005
regdum506	region_initreg_year==MNA_dem2006
regdum507	region_initreg_year==MNA_dem2007
regdum508	region_initreg_year==MNA_dem2008
regdum509	region_initreg_year==MNA_dem2009
regdum510	region_initreg_year==MNA_dem2010
regdum511	region_initreg_year==MNA_nd1960
regdum512	region_initreg_year==MNA_nd1961
regdum513	region_initreg_year==MNA_nd1962
regdum514	region_initreg_year==MNA_nd1963
regdum515	region_initreg_year==MNA_nd1964
regdum516	region_initreg_year==MNA_nd1965
regdum517	region_initreg_year==MNA_nd1966
regdum518	region_initreg_year==MNA_nd1967
regdum519	region_initreg_year==MNA_nd1968
regdum520	region_initreg_year==MNA_nd1969
regdum521	region_initreg_year==MNA_nd1970
regdum522	region_initreg_year==MNA_nd1971
regdum523	region_initreg_year==MNA_nd1972
regdum524	region_initreg_year==MNA_nd1973
regdum525	region_initreg_year==MNA_nd1974
regdum526	region_initreg_year==MNA_nd1975
regdum527	region_initreg_year==MNA_nd1976
regdum528	region_initreg_year==MNA_nd1977
regdum529	region_initreg_year==MNA_nd1978
regdum530	region_initreg_year==MNA_nd1979
regdum531	region_initreg_year==MNA_nd1980
regdum532	region_initreg_year==MNA_nd1981
regdum533	region_initreg_year==MNA_nd1982
regdum534	region_initreg_year==MNA_nd1983
regdum535	region_initreg_year==MNA_nd1984
regdum536	region_initreg_year==MNA_nd1985
regdum537	region_initreg_year==MNA_nd1986
regdum538	region_initreg_year==MNA_nd1987
regdum539	region_initreg_year==MNA_nd1988
regdum540	region_initreg_year==MNA_nd1989
regdum541	region_initreg_year==MNA_nd1990

Table 5: List of Variables *(continued)*

variable	label
regdum542	region__initreg__year==MNA__nd1991
regdum543	region__initreg__year==MNA__nd1992
regdum544	region__initreg__year==MNA__nd1993
regdum545	region__initreg__year==MNA__nd1994
regdum546	region__initreg__year==MNA__nd1995
regdum547	region__initreg__year==MNA__nd1996
regdum548	region__initreg__year==MNA__nd1997
regdum549	region__initreg__year==MNA__nd1998
regdum550	region__initreg__year==MNA__nd1999
regdum551	region__initreg__year==MNA__nd2000
regdum552	region__initreg__year==MNA__nd2001
regdum553	region__initreg__year==MNA__nd2002
regdum554	region__initreg__year==MNA__nd2003
regdum555	region__initreg__year==MNA__nd2004
regdum556	region__initreg__year==MNA__nd2005
regdum557	region__initreg__year==MNA__nd2006
regdum558	region__initreg__year==MNA__nd2007
regdum559	region__initreg__year==MNA__nd2008
regdum560	region__initreg__year==MNA__nd2009
regdum561	region__initreg__year==MNA__nd2010
regdum562	region__initreg__year==SAS__dem1960
regdum563	region__initreg__year==SAS__dem1961
regdum564	region__initreg__year==SAS__dem1962
regdum565	region__initreg__year==SAS__dem1963
regdum566	region__initreg__year==SAS__dem1964
regdum567	region__initreg__year==SAS__dem1965
regdum568	region__initreg__year==SAS__dem1966
regdum569	region__initreg__year==SAS__dem1967
regdum570	region__initreg__year==SAS__dem1968
regdum571	region__initreg__year==SAS__dem1969
regdum572	region__initreg__year==SAS__dem1970
regdum573	region__initreg__year==SAS__dem1971
regdum574	region__initreg__year==SAS__dem1972
regdum575	region__initreg__year==SAS__dem1973
regdum576	region__initreg__year==SAS__dem1974
regdum577	region__initreg__year==SAS__dem1975
regdum578	region__initreg__year==SAS__dem1976
regdum579	region__initreg__year==SAS__dem1977
regdum580	region__initreg__year==SAS__dem1978
regdum581	region__initreg__year==SAS__dem1979
regdum582	region__initreg__year==SAS__dem1980
regdum583	region__initreg__year==SAS__dem1981
regdum584	region__initreg__year==SAS__dem1982
regdum585	region__initreg__year==SAS__dem1983
regdum586	region__initreg__year==SAS__dem1984
regdum587	region__initreg__year==SAS__dem1985
regdum588	region__initreg__year==SAS__dem1986

Table 5: List of Variables *(continued)*

variable	label
regdum589	region_initreg_year==SAS_dem1987
regdum590	region_initreg_year==SAS_dem1988
regdum591	region_initreg_year==SAS_dem1989
regdum592	region_initreg_year==SAS_dem1990
regdum593	region_initreg_year==SAS_dem1991
regdum594	region_initreg_year==SAS_dem1992
regdum595	region_initreg_year==SAS_dem1993
regdum596	region_initreg_year==SAS_dem1994
regdum597	region_initreg_year==SAS_dem1995
regdum598	region_initreg_year==SAS_dem1996
regdum599	region_initreg_year==SAS_dem1997
regdum600	region_initreg_year==SAS_dem1998
regdum601	region_initreg_year==SAS_dem1999
regdum602	region_initreg_year==SAS_dem2000
regdum603	region_initreg_year==SAS_dem2001
regdum604	region_initreg_year==SAS_dem2002
regdum605	region_initreg_year==SAS_dem2003
regdum606	region_initreg_year==SAS_dem2004
regdum607	region_initreg_year==SAS_dem2005
regdum608	region_initreg_year==SAS_dem2006
regdum609	region_initreg_year==SAS_dem2007
regdum610	region_initreg_year==SAS_dem2008
regdum611	region_initreg_year==SAS_dem2009
regdum612	region_initreg_year==SAS_dem2010
regdum613	region_initreg_year==SAS_nd1960
regdum614	region_initreg_year==SAS_nd1961
regdum615	region_initreg_year==SAS_nd1962
regdum616	region_initreg_year==SAS_nd1963
regdum617	region_initreg_year==SAS_nd1964
regdum618	region_initreg_year==SAS_nd1965
regdum619	region_initreg_year==SAS_nd1966
regdum620	region_initreg_year==SAS_nd1967
regdum621	region_initreg_year==SAS_nd1968
regdum622	region_initreg_year==SAS_nd1969
regdum623	region_initreg_year==SAS_nd1970
regdum624	region_initreg_year==SAS_nd1971
regdum625	region_initreg_year==SAS_nd1972
regdum626	region_initreg_year==SAS_nd1973
regdum627	region_initreg_year==SAS_nd1974
regdum628	region_initreg_year==SAS_nd1975
regdum629	region_initreg_year==SAS_nd1976
regdum630	region_initreg_year==SAS_nd1977
regdum631	region_initreg_year==SAS_nd1978
regdum632	region_initreg_year==SAS_nd1979
regdum633	region_initreg_year==SAS_nd1980
regdum634	region_initreg_year==SAS_nd1981

Table 5: List of Variables *(continued)*

variable	label
regdum635	region_initreg_year==SAS_nd1982
regdum636	region_initreg_year==SAS_nd1983
regdum637	region_initreg_year==SAS_nd1984
regdum638	region_initreg_year==SAS_nd1985
regdum639	region_initreg_year==SAS_nd1986
regdum640	region_initreg_year==SAS_nd1987
regdum641	region_initreg_year==SAS_nd1988
regdum642	region_initreg_year==SAS_nd1989
regdum643	region_initreg_year==SAS_nd1990
regdum644	region_initreg_year==SAS_nd1991
regdum645	region_initreg_year==SAS_nd1992
regdum646	region_initreg_year==SAS_nd1993
regdum647	region_initreg_year==SAS_nd1994
regdum648	region_initreg_year==SAS_nd1995
regdum649	region_initreg_year==SAS_nd1996
regdum650	region_initreg_year==SAS_nd1997
regdum651	region_initreg_year==SAS_nd1998
regdum652	region_initreg_year==SAS_nd1999
regdum653	region_initreg_year==SAS_nd2000
regdum654	region_initreg_year==SAS_nd2001
regdum655	region_initreg_year==SAS_nd2002
regdum656	region_initreg_year==SAS_nd2003
regdum657	region_initreg_year==SAS_nd2004
regdum658	region_initreg_year==SAS_nd2005
regdum659	region_initreg_year==SAS_nd2006
regdum660	region_initreg_year==SAS_nd2007
regdum661	region_initreg_year==SAS_nd2008
regdum662	region_initreg_year==SAS_nd2009
regdum663	region_initreg_year==SAS_nd2010
dFY_1	regionINITREG==AFR_dem
dFY_2	regionINITREG==AFR_nd
dFY_3	regionINITREG==EAP_dem
dFY_4	regionINITREG==EAP_nd
dFY_5	regionINITREG==ECA_nd
dFY_6	regionINITREG==INL_dem
dFY_7	regionINITREG==INL_nd
dFY_8	regionINITREG==LAC_dem
dFY_9	regionINITREG==LAC_nd
dFY_10	regionINITREG==MNA_dem
dFY_11	regionINITREG==MNA_nd
dFY_12	regionINITREG==SAS_dem
dFY_13	regionINITREG==SAS_nd
gfa	(sum) gfa
nfa	(sum) nfa
totalassets	(sum) totalassets
totalliabilities	(sum) totalliabilities
nfagdp	(mean) nfagdp

Table 5: List of Variables *(continued)*

variable	label
nfagdpreg	NULL
incomequint50s_year1	NULL
incomequint50s_year2	NULL
quintile50s	NULL
dquint1	quintile50s== 1.0000
dquint2	quintile50s== 2.0000
dquint3	quintile50s== 3.0000
dquint4	quintile50s== 4.0000
dquint5	quintile50s== 5.0000
interfull_yy1_quintile1	NULL
interfull_yy1_quintile2	NULL
interfull_yy1_quintile3	NULL
interfull_yy1_quintile4	NULL
interfull_yy1_quintile5	NULL
interfull_yy2_quintile1	NULL
interfull_yy2_quintile2	NULL
interfull_yy2_quintile3	NULL
interfull_yy2_quintile4	NULL
interfull_yy2_quintile5	NULL
interfull_yy3_quintile1	NULL
interfull_yy3_quintile2	NULL
interfull_yy3_quintile3	NULL
interfull_yy3_quintile4	NULL
interfull_yy3_quintile5	NULL
interfull_yy4_quintile1	NULL
interfull_yy4_quintile2	NULL
interfull_yy4_quintile3	NULL
interfull_yy4_quintile4	NULL
interfull_yy4_quintile5	NULL
interfull_yy5_quintile1	NULL
interfull_yy5_quintile2	NULL
interfull_yy5_quintile3	NULL
interfull_yy5_quintile4	NULL
interfull_yy5_quintile5	NULL
interfull_yy6_quintile1	NULL
interfull_yy6_quintile2	NULL
interfull_yy6_quintile3	NULL
interfull_yy6_quintile4	NULL
interfull_yy6_quintile5	NULL
interfull_yy7_quintile1	NULL
interfull_yy7_quintile2	NULL
interfull_yy7_quintile3	NULL
interfull_yy7_quintile4	NULL
interfull_yy7_quintile5	NULL
interfull_yy8_quintile1	NULL
interfull_yy8_quintile2	NULL
interfull_yy8_quintile3	NULL

Table 5: List of Variables *(continued)*

variable	label
interfull__yy8__quintile4	NULL
interfull__yy8__quintile5	NULL
interfull__yy9__quintile1	NULL
interfull__yy9__quintile2	NULL
interfull__yy9__quintile3	NULL
interfull__yy9__quintile4	NULL
interfull__yy9__quintile5	NULL
interfull__yy10__quintile1	NULL
interfull__yy10__quintile2	NULL
interfull__yy10__quintile3	NULL
interfull__yy10__quintile4	NULL
interfull__yy10__quintile5	NULL
interfull__yy11__quintile1	NULL
interfull__yy11__quintile2	NULL
interfull__yy11__quintile3	NULL
interfull__yy11__quintile4	NULL
interfull__yy11__quintile5	NULL
interfull__yy12__quintile1	NULL
interfull__yy12__quintile2	NULL
interfull__yy12__quintile3	NULL
interfull__yy12__quintile4	NULL
interfull__yy12__quintile5	NULL
interfull__yy13__quintile1	NULL
interfull__yy13__quintile2	NULL
interfull__yy13__quintile3	NULL
interfull__yy13__quintile4	NULL
interfull__yy13__quintile5	NULL
interfull__yy14__quintile1	NULL
interfull__yy14__quintile2	NULL
interfull__yy14__quintile3	NULL
interfull__yy14__quintile4	NULL
interfull__yy14__quintile5	NULL
interfull__yy15__quintile1	NULL
interfull__yy15__quintile2	NULL
interfull__yy15__quintile3	NULL
interfull__yy15__quintile4	NULL
interfull__yy15__quintile5	NULL
interfull__yy16__quintile1	NULL
interfull__yy16__quintile2	NULL
interfull__yy16__quintile3	NULL
interfull__yy16__quintile4	NULL
interfull__yy16__quintile5	NULL
interfull__yy17__quintile1	NULL
interfull__yy17__quintile2	NULL
interfull__yy17__quintile3	NULL
interfull__yy17__quintile4	NULL

Table 5: List of Variables *(continued)*

variable	label
interfull__yy17__quintile5	NULL
interfull__yy18__quintile1	NULL
interfull__yy18__quintile2	NULL
interfull__yy18__quintile3	NULL
interfull__yy18__quintile4	NULL
interfull__yy18__quintile5	NULL
interfull__yy19__quintile1	NULL
interfull__yy19__quintile2	NULL
interfull__yy19__quintile3	NULL
interfull__yy19__quintile4	NULL
interfull__yy19__quintile5	NULL
interfull__yy20__quintile1	NULL
interfull__yy20__quintile2	NULL
interfull__yy20__quintile3	NULL
interfull__yy20__quintile4	NULL
interfull__yy20__quintile5	NULL
interfull__yy21__quintile1	NULL
interfull__yy21__quintile2	NULL
interfull__yy21__quintile3	NULL
interfull__yy21__quintile4	NULL
interfull__yy21__quintile5	NULL
interfull__yy22__quintile1	NULL
interfull__yy22__quintile2	NULL
interfull__yy22__quintile3	NULL
interfull__yy22__quintile4	NULL
interfull__yy22__quintile5	NULL
interfull__yy23__quintile1	NULL
interfull__yy23__quintile2	NULL
interfull__yy23__quintile3	NULL
interfull__yy23__quintile4	NULL
interfull__yy23__quintile5	NULL
interfull__yy24__quintile1	NULL
interfull__yy24__quintile2	NULL
interfull__yy24__quintile3	NULL
interfull__yy24__quintile4	NULL
interfull__yy24__quintile5	NULL
interfull__yy25__quintile1	NULL
interfull__yy25__quintile2	NULL
interfull__yy25__quintile3	NULL
interfull__yy25__quintile4	NULL
interfull__yy25__quintile5	NULL
interfull__yy26__quintile1	NULL
interfull__yy26__quintile2	NULL
interfull__yy26__quintile3	NULL
interfull__yy26__quintile4	NULL
interfull__yy26__quintile5	NULL
interfull__yy27__quintile1	NULL

Table 5: List of Variables *(continued)*

variable	label
interfull__yy27__quintile2	NULL
interfull__yy27__quintile3	NULL
interfull__yy27__quintile4	NULL
interfull__yy27__quintile5	NULL
interfull__yy28__quintile1	NULL
interfull__yy28__quintile2	NULL
interfull__yy28__quintile3	NULL
interfull__yy28__quintile4	NULL
interfull__yy28__quintile5	NULL
interfull__yy29__quintile1	NULL
interfull__yy29__quintile2	NULL
interfull__yy29__quintile3	NULL
interfull__yy29__quintile4	NULL
interfull__yy29__quintile5	NULL
interfull__yy30__quintile1	NULL
interfull__yy30__quintile2	NULL
interfull__yy30__quintile3	NULL
interfull__yy30__quintile4	NULL
interfull__yy30__quintile5	NULL
interfull__yy31__quintile1	NULL
interfull__yy31__quintile2	NULL
interfull__yy31__quintile3	NULL
interfull__yy31__quintile4	NULL
interfull__yy31__quintile5	NULL
interfull__yy32__quintile1	NULL
interfull__yy32__quintile2	NULL
interfull__yy32__quintile3	NULL
interfull__yy32__quintile4	NULL
interfull__yy32__quintile5	NULL
interfull__yy33__quintile1	NULL
interfull__yy33__quintile2	NULL
interfull__yy33__quintile3	NULL
interfull__yy33__quintile4	NULL
interfull__yy33__quintile5	NULL
interfull__yy34__quintile1	NULL
interfull__yy34__quintile2	NULL
interfull__yy34__quintile3	NULL
interfull__yy34__quintile4	NULL
interfull__yy34__quintile5	NULL
interfull__yy35__quintile1	NULL
interfull__yy35__quintile2	NULL
interfull__yy35__quintile3	NULL
interfull__yy35__quintile4	NULL
interfull__yy35__quintile5	NULL
interfull__yy36__quintile1	NULL
interfull__yy36__quintile2	NULL
interfull__yy36__quintile3	NULL

Table 5: List of Variables *(continued)*

variable	label
interfull__yy36__quintile4	NULL
interfull__yy36__quintile5	NULL
interfull__yy37__quintile1	NULL
interfull__yy37__quintile2	NULL
interfull__yy37__quintile3	NULL
interfull__yy37__quintile4	NULL
interfull__yy37__quintile5	NULL
interfull__yy38__quintile1	NULL
interfull__yy38__quintile2	NULL
interfull__yy38__quintile3	NULL
interfull__yy38__quintile4	NULL
interfull__yy38__quintile5	NULL
interfull__yy39__quintile1	NULL
interfull__yy39__quintile2	NULL
interfull__yy39__quintile3	NULL
interfull__yy39__quintile4	NULL
interfull__yy39__quintile5	NULL
interfull__yy40__quintile1	NULL
interfull__yy40__quintile2	NULL
interfull__yy40__quintile3	NULL
interfull__yy40__quintile4	NULL
interfull__yy40__quintile5	NULL
interfull__yy41__quintile1	NULL
interfull__yy41__quintile2	NULL
interfull__yy41__quintile3	NULL
interfull__yy41__quintile4	NULL
interfull__yy41__quintile5	NULL
interfull__yy42__quintile1	NULL
interfull__yy42__quintile2	NULL
interfull__yy42__quintile3	NULL
interfull__yy42__quintile4	NULL
interfull__yy42__quintile5	NULL
interfull__yy43__quintile1	NULL
interfull__yy43__quintile2	NULL
interfull__yy43__quintile3	NULL
interfull__yy43__quintile4	NULL
interfull__yy43__quintile5	NULL
interfull__yy44__quintile1	NULL
interfull__yy44__quintile2	NULL
interfull__yy44__quintile3	NULL
interfull__yy44__quintile4	NULL
interfull__yy44__quintile5	NULL
interfull__yy45__quintile1	NULL
interfull__yy45__quintile2	NULL
interfull__yy45__quintile3	NULL
interfull__yy45__quintile4	NULL

Table 5: List of Variables *(continued)*

variable	label
interfull__yy45__quintile5	NULL
interfull__yy46__quintile1	NULL
interfull__yy46__quintile2	NULL
interfull__yy46__quintile3	NULL
interfull__yy46__quintile4	NULL
interfull__yy46__quintile5	NULL
interfull__yy47__quintile1	NULL
interfull__yy47__quintile2	NULL
interfull__yy47__quintile3	NULL
interfull__yy47__quintile4	NULL
interfull__yy47__quintile5	NULL
interfull__yy48__quintile1	NULL
interfull__yy48__quintile2	NULL
interfull__yy48__quintile3	NULL
interfull__yy48__quintile4	NULL
interfull__yy48__quintile5	NULL
interfull__yy49__quintile1	NULL
interfull__yy49__quintile2	NULL
interfull__yy49__quintile3	NULL
interfull__yy49__quintile4	NULL
interfull__yy49__quintile5	NULL
interfull__yy50__quintile1	NULL
interfull__yy50__quintile2	NULL
interfull__yy50__quintile3	NULL
interfull__yy50__quintile4	NULL
interfull__yy50__quintile5	NULL
interfull__yy51__quintile1	NULL
interfull__yy51__quintile2	NULL
interfull__yy51__quintile3	NULL
interfull__yy51__quintile4	NULL
interfull__yy51__quintile5	NULL
country	Country Name
areakm2	Area in km2
cen__lat	latitude of country centroid
cen__lon	longitude of country centroid
elev	mean m above sea level
distcr	mean distance to coast or river
distc	mean distance to coast
distr	mean distance to river
tropicar	% land area in geographical tropics
troppop	%pop ('95) in geographical tropics
lc100km	%area 100km from icefree coast
lcr100km	%area 100km from icefree coast or sea-nav. river
pop95	1995 pop (from GPWv2)
pdenpavg	typical pop density experienced
pop100km	%pop ('95) 100km from icefree coast
pop100cr	%pop ('95) 100km from icefree coast or sea-nav. river

Table 5: List of Variables *(continued)*

variable	label
cen_c	dist centroid to coast(km)
cen_cr	dist centroid to coast/riv (km)
polity	NULL
xrreg	NULL
xrcomp	NULL
xropen	NULL
xconst	NULL
parreg	NULL
parcomp	NULL
exrec	NULL
exconst	NULL
polcomp	NULL
polity2_aug	NULL
independent	NULL
transition	NULL
interruption	NULL
interregnum	NULL
pr	NULL
cl	NULL
pr_aug	NULL
cl_aug	NULL
demt	NULL
polity2	NULL
status	NULL
NAME	NAME
LON	LON
LAT	LAT
_ID	NULL
GDPpercapitaconstantLCUN	GDP per capita (constant LCU) [NY.GDP.PCAP.KN]
rgdpl2	NULL
rgdpna_full	NULL
PopulationtotalSPPOPTOTL	Population, total [SP.POP.TOTL]
Populationages014oftotal	Population ages 0-14 (% of total) [SP.POP.0014.TO.ZS]
Populationages1564oftota	Population ages 15-64 (% of total) [SP.POP.1564.TO.ZS]

5.2 Arellano Bond Estimation for Table.2 (Shoya Abe)

5.2.1 Arellano Bond Estimation

We will explain the Arellano Bond Estimation that we tried. This estimation is a type of generalized method of moments (GMM). To simplify the explanation, consider estimating the following equation.

$$y_{c,t} = \mu_c + \alpha y_{c,t-1} + \epsilon_{c,t}. \tag{15}$$

For $s \geq 2$, the following moment condition holds:

$$E\Big(y_{c,t-2}(\epsilon_{c,t} - \epsilon_{c,t-1})\Big) = E\Big(y_{c,t-2}(y_{c,t} - y_{c,t-1} - \alpha(y_{c,t-1} - y_{c,t-2}))\Big) = 0. \tag{16}$$

The Arellano Bond estimator is a GMM estimator that uses all of these moment conditions.

5.2.2 Preprocessing

```
# Select the first 30 columns of the dataset and prepare panel data
data_t2 <- data |>
  select(1:30) |> # Select relevant columns
  group_by(country_name) |>
  arrange(year) |> # Arrange by year in ascending order
  mutate(
    lag1 = dplyr::lag(y, 1),
    lag2 = dplyr::lag(y, 2),
    lag3 = dplyr::lag(y, 3),
    lag4 = dplyr::lag(y, 4),
    lag5 = dplyr::lag(y, 5),
    lag6 = dplyr::lag(y, 6),
    lag7 = dplyr::lag(y, 7),
    lag8 = dplyr::lag(y, 8)
  ) |>
  ungroup()

# Prepare panel data structures for different models with varying lags
data_m1 <- data_t2 |>
  drop_na(y, dem, lag1) |>
  pdata.frame(index = c("country_name", "year"))
data_m2 <- data_t2 |>
  drop_na(y, dem, lag1, lag2) |>
  pdata.frame(index = c("country_name", "year"))
data_m3 <- data_t2 |>
  drop_na(y, dem, lag1, lag2, lag3, lag4) |>
  pdata.frame(index = c("country_name", "year"))
data_m4 <- data_t2 |>
  drop_na(
    y, dem, lag1, lag2, lag3, lag4,
    lag5, lag6, lag7, lag8
  ) |>
  pdata.frame(index = c("country_name", "year"))

# Maximum lag to be used for instruments
maxlag <- 49
```

5.2.3 Estimation

```
# Estimate Arellano-Bond GMM models with different lag structures
model_1_gmm <- pgmm(
  y ~ dem + lag(y, 1) |
    lag(y, 2:maxlag) + lag(dem, 1:maxlag),
    # Use higher lags as instruments
  data = data_m1,
  effect = "twoways",
```

```r
  model = "twosteps",
  transformation = "d"
)

model_2_gmm <- pgmm(
  y ~ dem + lag(y, 1) + lag(y, 2) |# Include two lags of GDP
    lag(y, 2:maxlag) + lag(dem, 1:maxlag),
  data = data_m2,
  effect = "twoways",
  model = "twosteps",
  transformation = "d"
)

model_3_gmm <- pgmm(
  y ~ dem + lag(y, 1) + lag(y, 2) +
    lag(y, 3) + lag(y, 4) | # Include four lags
    lag(y, 2:maxlag) + lag(dem, 1:maxlag),
  data = data_m3,
  effect = "twoways",
  model = "twosteps",
  transformation = "d"
)

model_4_gmm <- pgmm(
  y ~ dem + lag(y, 1) + lag(y, 2) +
    lag(y, 3) + lag(y, 4) +
    lag(y, 5) + lag(y, 6) +
    lag(y, 7) + lag(y, 8) | # Include eight lags
    lag(y, 2:maxlag) + lag(dem, 1:maxlag),
  data = data_m4,
  effect = "twoways",
  model = "twosteps",
  transformation = "d"
)

# Function to compute cumulative effects over a given time horizon
compute_dynamic_effect <- function(dem_coef, lag_coefs, n_periods) {
  effects <- numeric(n_periods) # Initialize vector to store dynamic effects
  effects[1] <- dem_coef # Initial effect of democratization
  k <- length(lag_coefs) # Number of lags considered
  if (n_periods > 1) {
    for (i in 2:n_periods) {
      eff <- dem_coef # Start with the direct effect of democracy
      for (j in 1:min(i - 1, k)) {
        eff <- eff + effects[i - j] * lag_coefs[j]
        # Accumulate effects over time
      }
      effects[i] <- eff
    }
  }
  effects[n_periods]
}
```

```r
# Extract estimated coefficients for each model
coef_1 <- coef(model_1_gmm)
dem_coef_1 <- coef_1["dem"]
lag1_1 <- coef_1["lag(y, 1)"]

# Compute long-run and short-run effects for each model
lre1 <- dem_coef_1 / (1 - lag1_1)

# Compute persistence (sum of lag coefficients)
pers1 <- lag1_1

# Compute the effect after 25 years for each model
eff_25_1 <- compute_dynamic_effect(
  dem_coef_1, c(lag1_1), 25
)

coef_2 <- coef(model_2_gmm)
dem_coef_2 <- coef_2["dem"]
lag1_2 <- coef_2["lag(y, 1)"]
lag2_2 <- coef_2["lag(y, 2)"]
lre2 <- dem_coef_2 / (1 - (lag1_2 + lag2_2))
pers2 <- lag1_2 + lag2_2
eff_25_2 <- compute_dynamic_effect(
  dem_coef_2, c(lag1_2, lag2_2), 25
)

coef_3 <- coef(model_3_gmm)
dem_coef_3 <- coef_3["dem"]
lag1_3 <- coef_3["lag(y, 1)"]
lag2_3 <- coef_3["lag(y, 2)"]
lag3_3 <- coef_3["lag(y, 3)"]
lag4_3 <- coef_3["lag(y, 4)"]
lre3 <- dem_coef_3 / (1 - (lag1_3 +
  lag2_3 + lag3_3 + lag4_3))
pers3 <- lag1_3 + lag2_3 + lag3_3 + lag4_3
eff_25_3 <- compute_dynamic_effect(
  dem_coef_3, c(lag1_3, lag2_3, lag3_3, lag4_3), 25
)

coef_4 <- coef(model_4_gmm)
dem_coef_4 <- coef_4["dem"]
lag1_4 <- coef_4["lag(y, 1)"]
lag2_4 <- coef_4["lag(y, 2)"]
lag3_4 <- coef_4["lag(y, 3)"]
lag4_4 <- coef_4["lag(y, 4)"]
lag5_4 <- coef_4["lag(y, 5)"]
lag6_4 <- coef_4["lag(y, 6)"]
lag7_4 <- coef_4["lag(y, 7)"]
lag8_4 <- coef_4["lag(y, 8)"]
lre4 <- dem_coef_4 / (1 - (lag1_4 +
  lag2_4 + lag3_4 + lag4_4 + lag5_4 +
  lag6_4 + lag7_4 + lag8_4))
pers4 <- lag1_4 + lag2_4 + lag3_4 +
```

```r
  lag4_4 + lag5_4 + lag6_4 + lag7_4 + lag8_4
eff_25_4 <- compute_dynamic_effect(
  dem_coef_4,
  c(
    lag1_4, lag2_4, lag3_4, lag4_4,
    lag5_4, lag6_4, lag7_4, lag8_4
  ),
  25
)

lre <- round(c(lre1, lre2, lre3, lre4), 3)
pers <- round(c(pers1, pers2, pers3, pers4), 3)
eff_25 <- round(
  c(eff_25_1, eff_25_2, eff_25_3, eff_25_4),
  3
)

se1 <- sqrt(diag(vcov(model_1_gmm)))
se2 <- sqrt(diag(vcov(model_2_gmm)))
se3 <- sqrt(diag(vcov(model_3_gmm)))
se4 <- sqrt(diag(vcov(model_4_gmm)))

override.coef.1 <- c(
  coef_1["dem"],
  coef_1["lag(y, 1)"],
  rep(NA, 7)
)
override.se.1 <- c(
  se1["dem"],
  se1["lag(y, 1)"],
  rep(NA, 7)
)
override.coef.2 <- c(
  coef_2["dem"],
  coef_2["lag(y, 1)"],
  coef_2["lag(y, 2)"],
  rep(NA, 6)
)
override.se.2 <- c(
  se2["dem"],
  se2["lag(y, 1)"],
  se2["lag(y, 2)"],
  rep(NA, 6)
)
override.coef.3 <- c(
  coef_3["dem"],
  coef_3["lag(y, 1)"],
  coef_3["lag(y, 2)"],
  coef_3["lag(y, 3)"],
  coef_3["lag(y, 4)"],
  rep(NA, 4)
)
override.se.3 <- c(
```

```
  se3["dem"],
  se3["lag(y, 1)"],
  se3["lag(y, 2)"],
  se3["lag(y, 3)"],
  se3["lag(y, 4)"],
  rep(NA, 4)
)
override.coef.4 <- c(
  coef_4["dem"],
  coef_4["lag(y, 1)"],
  coef_4["lag(y, 2)"],
  coef_4["lag(y, 3)"],
  coef_4["lag(y, 4)"],
  coef_4["lag(y, 5)"],
  coef_4["lag(y, 6)"],
  coef_4["lag(y, 7)"],
  coef_4["lag(y, 8)"]
)
override.se.4 <- c(
  se4["dem"],
  se4["lag(y, 1)"],
  se4["lag(y, 2)"],
  se4["lag(y, 3)"],
  se4["lag(y, 4)"],
  se4["lag(y, 5)"],
  se4["lag(y, 6)"],
  se4["lag(y, 7)"],
  se4["lag(y, 8)"]
)
```

5.2.4 Tabulation

```
models <- list(model_1_gmm, model_2_gmm, model_3_gmm, model_4_gmm)

#Generating LaTex Table
texreg(
  models,
  override.coef = list(
    override.coef.1,
    override.coef.2,
    override.coef.3,
    override.coef.4
  ),
  override.se = list(
    override.se.1,
    override.se.2,
    override.se.3,
    override.se.4
  ),
  custom.model.names = c("(1)", "(2)", "(3)", "(4)"),
  custom.coef.names = c(
    "Democracy", "Lag 1", "Lag 2",
```

```
    "Lag 3", "Lag 4", "Lag 5",
    "Lag 6", "Lag 7", "Lag 8"
  ),
  custom.gof.rows = list(
    "Persistence" = pers,
    "Long run effect" = lre,
    "Effect after 25 years" = eff_25
  ),
  file = "output/table_2_GMM.tex",
  caption = "Effect of Democracy on (Log) GDP per Capita: Arellano-Bond GMM Estimation"
)
```

There are two possible reasons for the discrepancy between the results of the original paper and our estimates. The first concerns the setting regarding the number of lags used. We employ lags up to a maximum of 49 periods as instruments, whereas the original paper's Stata replication code appears to use an automatic selection procedure provided by a package. Consequently, the moment conditions being estimated may differ, leading to different estimation results. The second reason involves data preprocessing. In the Stata code, procedures such as bootstrap sample extraction are performed, which may result in a different sample composition and, consequently, different estimation outcomes.

	(1)	(2)	(3)	(4)
Democracy	2.79	2.29	0.05	1.51
	(2.12)	(1.63)	(1.42)	(0.51)
Lag 1	0.96^{***}	0.99^{***}	0.94^{***}	0.93^{***}
	(0.03)	(0.03)	(0.03)	(0.01)
Lag 2		-0.02	-0.00	-0.01
		(0.01)	(0.01)	(0.00)
Lag 3			0.00	0.00
			(0.01)	(0.00)
Lag 4			-0.02^{*}	-0.01
			(0.01)	(0.00)
Lag 5				-0.00
				(0.00)
Lag 6				0.00
				(0.00)
Lag 7				-0.00
				(0.00)
Lag 8				-0.00
				(0.00)
Persistence	0.96	0.97	0.92	0.91
Long run effect	63.18	74.26	0.65	16.40
Effect after 25 years	42.76	40.77	0.59	15.27
n	175	175	175	175
T	50	49	47	43
Num. obs.	6790	6642	6336	5688
Num. obs. used	6542	6311	5824	4779
Sargan Test: chisq	145.66	147.27	140.10	146.09
Sargan Test: df	2398.00	2297.00	2095.00	1691.00
Sargan Test: p-value	1.00	1.00	1.00	1.00
Wald Test Coefficients: chisq	808.19	984.51	1143.95	2227.71
Wald Test Coefficients: df	2	3	5	9
Wald Test Coefficients: p-value	0.00	0.00	0.00	0.00
Wald Test Time Dummies: chisq	533.24	491.67	497.42	453.37
Wald Test Time Dummies: df	48	46	42	34
Wald Test Time Dummies: p-value	0.00	0.00	0.00	0.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 6: Effect of Democracy on (Log) GDP per Capita: Arellano–Bond GMM Estimation