

Prosperity and Democracy: Examining the Correlation through Two-Way Fixed-Effects Modelling

Nikita Sidorov MA, Central European University

6/5/2023

Contents

1	Abstract	2
2	Introduction	3
3	Theory: Economic Development and Democracy	3
3.1	Hypothesis & RQ	4
4	Research Design & Data Description	4
4.1	Research Design	4
4.2	Data Source and Variables	4
5	Exploratory Data Analysis	7
5.1	Summary Statistics	7
5.2	NAs Handling	8
5.3	Univariate Statistics	9
5.4	Correlations	9
6	Model Building	11
6.1	Basic Model	11
6.2	Two-Way Fixed-Effects Model	12
7	Discussion & Future Improvements	15
	References	15

1 Abstract

The causal link between democracy and economic development has been established half a century ago by Lipset (1959), however, it remains the subject of still ongoing debate among scholars. While modernisation theory suggests that economic growth leads to increased demand for democracy, there are exceptions to this rule, such as China's experience. This paper tests the hypothesis of modernisation theory using Varieties of Democracy (V-Dem) data from 1991 to 2019 with a help of a Two-Way Fixed-Effects OLS regression model. We find a statistically significant positive relationship between GDP per capita and the Polyarchy index of democracy. However, the relationship is relatively weak, and the issue of the direction of causality remains open for future research. The study fills an important gap in the literature by providing a more contemporary analysis of the relationship between wealth and democracy.

2 Introduction

This paper contributes to the ongoing debate regarding the relationship between wealth and democracy. While many scholars have studied this topic, there is currently no consensus on the existence and the direction of the relationship. Lipset (1959) has developed a modernisation theory that suggests that *economic growth leads to increased demand for democracy, making democratisation more likely*. However, there are important exceptions to this rule, such as the People’s Republic of China, which challenges the universality of this theory (Liu, Su, and Wu 2022). Additionally, modernisation theory was developed just after the second - post WWII - wave of democratisation. Thus, it does not take into account a) the third wave, (Huntington 1993) when democratisation was experienced by relatively less developed countries as Hungary and Poland that were hailed as “post-communist success stories” (Herman 2015, 1058), b) the third reverse-wave when de-democratisation happened in less-established democracies and countries in so-called transitional state (Carothers 2002).

Therefore, in this paper, we address this gap by testing the hypothesis of modernisation theory using the V-Dem data from 1991 to 2019. We employ OLS regression coupled with a two-way fixed-effects (TWFE) model to test the validity of the hypothesis. Our findings indicate a statistically significant positive relationship between the polyarchy index of democracy and GDP per capita. However, two caveats must be considered: the relationship is relatively weak with 1% increase in Log GDP PPC increasing polyarchy index by approximately ~ 0.0001 . Secondly, the study leaves the issue of determination of causality to the future scholarship, as, despite using TWFE models, we cannot be sure we isolate the effect of economic development on democracy (for this purposes, instrumental variable approach may be preferred). Nevertheless, our study provides a more contemporary and comprehensive analysis of the relationship between wealth and democracy, filling an important gap in the literature.

3 Theory: Economic Development and Democracy

In his seminal study “Some Social Requisites of Democracy: Economic Development and Political Legitimacy” Lipset (1959) claimed that more democratic states would have higher levels of wealth, industrialisation, education, and urbanisation, hence it can be assumed that democracy is the final stage of the process of modernisation: *the more developed the state becomes the higher the chance it will be democratic*. Wealth is considered as a factor that can contribute to democratisation, and several potential causal links exist between them. Firstly, individuals with sufficient wealth are more likely to have the time to participate in political processes. People without jobs usually exhibit lower participation levels in elections, which are one of the easiest options for participation, as they require less time compared to other types of involvement. Secondly, wealth is associated with increased education levels, and educated people typically demand democratic changes (Lipset 1959). Thus, as a country’s wealth increases, it becomes more likely to fulfill the necessary conditions for democracy to take root.

At the same time, Treisman (2014) provides further insights into the relationship between economic development, democracy, and political leadership. He argues that leader turnover matters because leaders themselves have a significant impact on the political landscape, and the distribution of incoming leaders starts to include some who permit reform due to education or modern values. Over time, reformers are selected out as they introduce and lose elections or are ousted by coups. Treisman’s analysis shows that economic development has different consequences in the short and medium run. Higher income promotes political liberalisation in the medium run, but faster growth entrenches incumbent leaders, helping autocrats resist pressures for reform in the short run.

However, China’s unique experience challenges this theory (Liu, Su, and Wu 2022). In the years leading up to the mid-2010s, the country had experienced the rapid growth of GDP per capita, and a demand for liberal democracy had emerged among certain socio-economic groups, such as the better educated, media exposed, urban, and young. Nevertheless, this effect had almost disappeared in a decade (Liu, Su, and Wu 2022). Hence, authors argue that China’s exceptionalism challenges modernisation theory’s universality, as the state control of ideology and political indoctrination can counteract the modernisation effect. The authors

suggest that empirical analysis of the Chinese case supports the theory that state intervention, particularly through political indoctrination, can decrease the likelihood of individuals holding liberal democratic values. This evidence calls for a more context-specific analysis of political and cultural change. Furthermore, this theory does not account for affluent non-democracies as Singapore, Qatar, and United Arab Emirates - why they have not democratised despite economic prosperity? Additionally, at the time of writing Lipset could have assumed that many countries, that today we deem as second-wave democratised, have done so due to exogenous circumstances - i.e., influence and military occupation by the Allies after the WWII with obvious example of Axis countries (Germany, Japan, and Italy). It is worth noting that the “external democratising pressure” would later become a crucial component in Levitsky and Way (2006) theory.

Having examined the theoretical underpinnings of this work, we now state our hypothesis and research question.

3.1 Hypothesis & RQ

Research question: What is the impact of modernisation on the level of democracy?

Our main Hypothesis states that:

H_1 : in line with the modernisation theory, the economic prosperity leads to democratisation.

4 Research Design & Data Description

4.1 Research Design

To answer our research question (RQ), we employ an OLS model to investigate a potential causal effect between economic development (GDP per capita) and the level of democracy, measured by the Polyarchy index from a V-Dem dataset. We use a *two-way fixed-effects* (TWFE) model to control for both cross-sectional and time-series factors. The idea behind this design is to account for unobserved heterogeneity across different countries and years in the data, thus producing more robust coefficients. Additionally, it helps to address *omitted variable bias* by controlling for the fixed effects, which account for the influence of unobserved or excluded variables, thus theoretically increasing the internal validity of our study (ensuring causality works in the way we propose).

However, the TWFE model may not be capable of controlling for all possible confounding variables, such as those that may vary over time. Therefore, we also include a control variable, the Index of Corruption. Indeed, high corruption may negatively affect the GDP by promoting solutions that are less than efficient (Mustapha 2014) and democracy (Kubbe and Engelbert 2017) by influencing political culture and discourse in the society, therefore, therefore preventing (or subverting) democratisation efforts. We also contend that other factors, such as level of education, colonial history, or high share of natural resources in country’s export (Ross 2001) may as well be potential confounding variables. Unfortunately, the testing of their potential influence is beyond the scope of this paper. The proposed model can be described via the following general equation:

$$Y_{it} = \alpha + x_{1it}\beta_1 + x_{2it}\beta_2 + u_i + v_t + e_{it}$$

where Y_{it} is a Polyarchy index for country i at the time t , $x_{1it}\beta_1$ and $x_{2it}\beta_2$ are independent variables (GDP and Control), u_i - cross-sectional fixed-effects (FEs) for the i -th country, v_t - time-series FEs for the t -th year, and e_{it} is an error term.

4.2 Data Source and Variables

To test our hypothesis, we employ one of the largest datasets on political indicators - V-Dem (Varieties of Democracy) - developed and maintained by the University of Gothenburg. The full database consists of more

than 27000 observations for 4600 variables. This is the example of a panel data (i.e., both cross-sectional and cross-temporal). The operations on such data are usually harder, as there is a need to control for the country and/or yearly specific variation.

However, for the purposes of this project, we use only a small fraction of data, given that:

1. We use data only from 1991 to 2019 to capture the mentioned third-wave of democratisation (c.f.: (Huntington 1993)) and not the earlier variation. The upper bound corresponds to the last year before the beginning of the COVID-19 outbreak that tighten the regulations worldwide (Greitens 2020) and impacted many democracy indicators. Hence, we exclude the concurrent years to keep this exogenous factor out.
2. We use only several out of 4600 variables: the list of variables and their description can be found in Table 1.

Table 1: Description of the Variables

Variable	Class	Scale	Description
Polyarchy Index	Dependent Variable	continuous (0-1)	Polyarchy index based on Dahl and Shapiro (2015) concept of polyarchy and includes measures of freedom of elections, access to alternative information, and extension of suffrage.
GDP Per Capita	Independent Variable	Numeric	Point estimate of yearly GDP per capita provided by Fariss et al. (2021).
Corruption Index	Control	continuous (0-1)	The degree of corruption of the regime (averaging the indices for public service, legislative, executive, and judicial corruption).
Year	Fixed-Effect I	Character	The year for observation (from 1991 to 2019)
Country ID	Fixed-Effect II	Character	Country Code

Our main Dependent Variable is the polyarchy index. It is an aggregate estimate of the country’s level of democracy, which is a high-level democracy indicator produced by taking the average and weighted average of indicators measuring the freedom of associations, fairness and freedom of elections, access to alternative information, and the extension of suffrage (which are, in turn, either aggregated themselves or based on the expert responses). We use this variable, as it captures the multifaceted nature of democracy. Indeed, it goes beyond the procedural definition (Schumpeter 2013) and encompasses variety of aspects of a vibrant liberal democracy.

Our main independent variable is the country’s Gross Domestic Product Per Capita. Interestingly, it is an external variable in the dataset provided by Fariss et al. (2021). Originally, it was created to better estimate the data in the past (i.e., the original index goes back half a millennium), however, we use it (and not other estimation) mainly for the purpose of convenience, as we need not parse the data from the other sources and thus avoid a potential pitfall of mismatched identifiers (when in different datasets a country is coded differently, e.g., Portugal being coded as *PRT* and *POR* in different classifiers (ISO vs IOC)). We use GDP per capita as, by dividing the GDP by the population size, we are able to inadvertently control for the difference in population sizes.

At the same time, we control for the country’s level of corruption, since it may be a confounding variable in the regression system. The variable is also an aggregate of aggregates that are based on the expert evaluations. Theoretically, it may influence both the GDP and the polyarchy index. We do not employ further controls, since our idea is to use a TWFE model to control for the country- and time-specific factors that may influence the democratic index, which has become an ‘industry standard’ (but its effectiveness is not ubiquitous (Imai and Kim 2020)).

As for the data type, most of the variables are not experimental, however, a caveat must be considered. While some variables can be considered fully observational - for example, percentage of population with suffrage rights (a part of the polyarchy index) - as they can be deduced from the government policies/reports/statistics, it is not a universal case. Beyond the high-level aggregation, low-level variables, as government censorship effort (a part of the freedom of expression index, which is used to calculate the additive polyarchy index, which is used to measure the polyarchy index itself) - relies on the expert grading. At the same time, the GDP values are constructed by modelling the ‘true’ GDP per capita values based on the 9 existing datasets (Fariss et al. 2021). Having canvassed the existing literature and described the data, we turn to the Exploratory Data Analysis.

5 Exploratory Data Analysis

Now, allow us to examine the variables. Firstly, we will present the technical statistics (NA cases and their handling, descriptive univariate statistics), then, we will explore some of the noteworthy relationships graphically and after that turn to the model building process.

5.1 Summary Statistics

We operate with the following number of unique observations.

Table 2: Unique Observations

Countries	Years
179	29

Overall, we have the data on 179 countries and 29 years - meaning the original dataset included most (if not all) UN recognised states and all the years are, at least partially, presented in the data.

Additionally, we are interested in the year-country pairs to ensure that for yearly data is available for every state.

Table 3: Country-Year Pairs

Country Name	Count
Bosnia and Herzegovina	28
Slovakia	27
Montenegro	22
Kosovo	21
Palestine/Gaza	13
South Sudan	9

As expected, for some of the countries the data cannot be obtained as they probably gained independence more recently. To avoid potential biases we remove countries that have less than half (14) years available in the data (i.e., South Sudan and Palestine/Gaza). While it is not strictly necessary, given the control for the country-specific effects, we, nevertheless, handle the NAs ourselves, since the `fixest` package that we use for TWFE works better without them (also, by doing this we explicitly introduce the preferred NA handling method)

Secondly, the basic summary statistics of the data (apart from the ID variables):

Table 4: Summary Statistics

Year	Polyarchy	GDP PPC	Corruption Index
Min. :1991	Min. :0.01	Min. : 0.3	Min. :0.002
1st Qu.:1998	1st Qu.:0.27	1st Qu.: 2.8	1st Qu.:0.236
Median :2005	Median :0.50	Median : 7.8	Median :0.565
Mean :2005	Mean :0.51	Mean : 13.9	Mean :0.512
3rd Qu.:2012	3rd Qu.:0.76	3rd Qu.: 20.1	3rd Qu.:0.780
Max. :2019	Max. :0.92	Max. :100.9	Max. :0.967
NA	NA	NA's :116	NA's :24

The summary shows that that:

- No years are missing in the data;
- The variation of the Polyarchy index is high with values ranging from the most undemocratic (0.01) to strongly democratic (0.92). The decent range of values would allow to avoid over-fitting of the model and thus lead to better applicability of the results;
- The variation of GDP is also high, however, there are 116 missing cases which we will handle later;
- The Corruption Index also displays significant variation with only 24 cases missing.

5.2 NAs Handling

Now, we deal with the missing cases problem. Usually, when the amount of missing data is low (5% of *randomly* missing data is considered a threshold (Emmanuel et al. 2021)) it is possible to omit it. In our case, we have just over 2% of missing data. While it is theoretically possible to just disregard it, upon further investigation it is revealed that the data is **not** randomly missing. For Hong Kong, Somaliland, Palestine/West Bank, and Zanzibar all the data on GDP PPC is missing. Hence, these countries will also have to be removed from the dataset.

Table 5: Missing GDP PPC Cases Count

Country Name	Count
Hong Kong	29
Palestine/West Bank	29
Somaliland	29
Zanzibar	29

As for the Corruption index, there are 24 missing cases total for two countries.

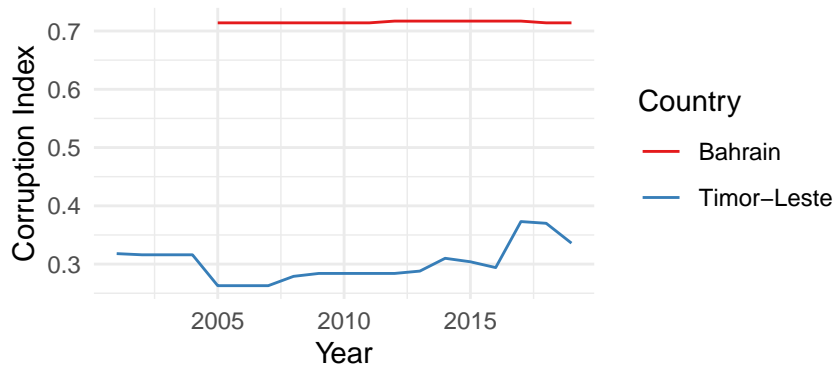
Table 6: Missing Corruption Index Cases Count

Country Name	Count
Bahrain	14
Timor-Leste	10

Interestingly, the values for Corruption index do not vary heavily for Bahrain and exhibit just a slight variation for Timor-Leste. Hence, it is reasonable to impute the missing values with mean values for the respective states.

Figure 1. Corruption Index for Selected Countries

Missing years are included for demonstration



After all manipulations, we are left with 4999 observations for 173 countries over 29 years.

5.3 Univariate Statistics

Now, we shall explore the distributions of the relevant variables and relationships between them.

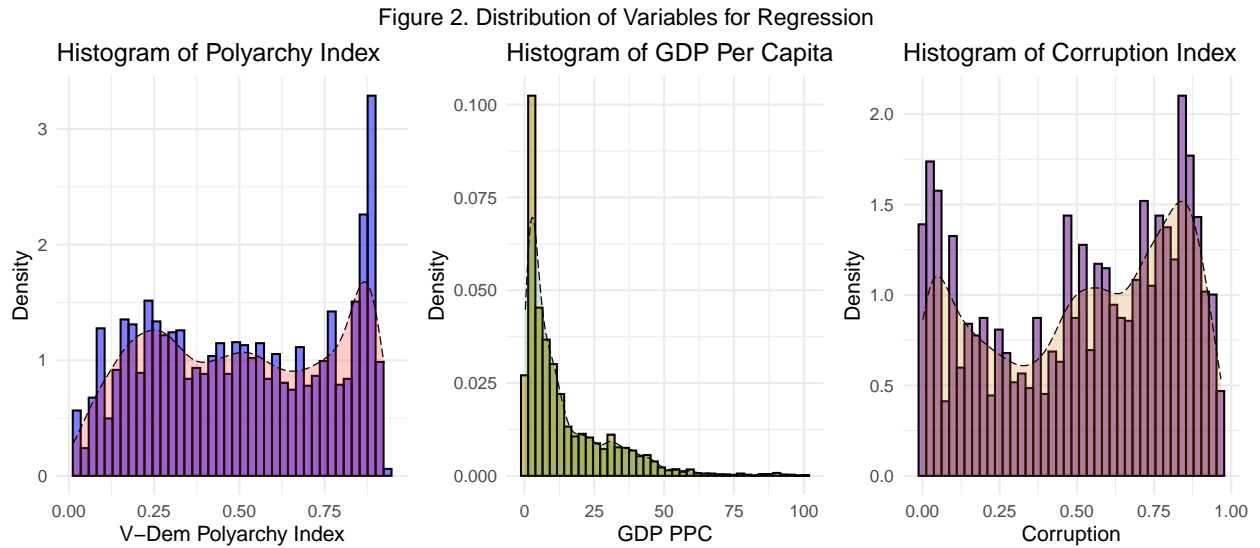


Table 7: Normality Check for the Variables

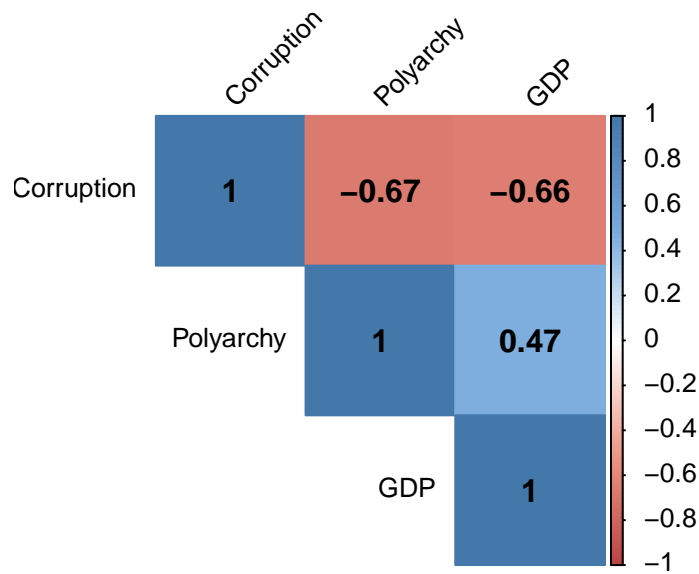
Variable	W	p-value
Polyarchy Index	0.9381	< 0.001
GDP PPC	0.7753	< 0.001
Corruption Index	0.9241	< 0.001

These plots paint an interesting picture. As we can see, all the variables are, in fact, not normally distributed. Additionally, the results of Shapiro-Wilk tests lead us to the same conclusion. P-value for all the tests is less than significance threshold $\alpha = 0.05$ meaning they are not distributed normally. However, we can at least partially improve the situation by applying a logarithmic transformation to the GDP PPC to lessen the skewness. The same procedure would not work with other two variables, given that the logarithmic transformation would change the signs (given that $\log(1) = 0$ and $\log(0) = \text{undefined}$). Additionally, the small scale of observation can lead to distorted interpretation of the transformed data.

5.4 Correlations

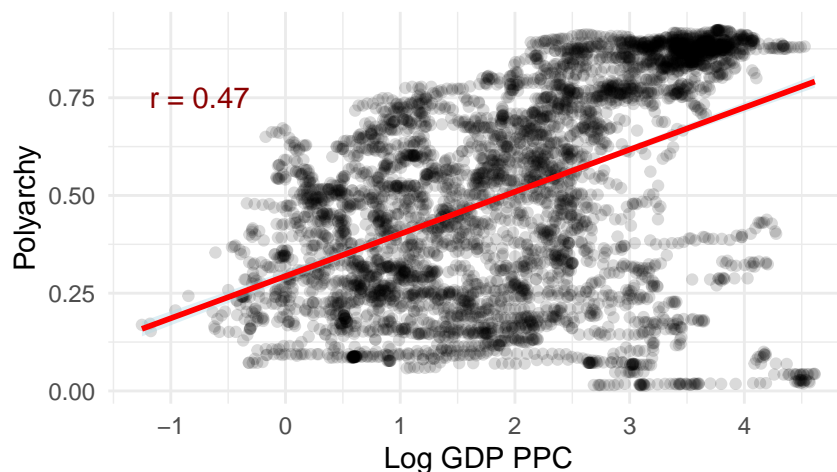
Now, we examine the correlations between the variables of interest. For this purpose, we use a correlation matrix. The matrix shows that there are significant negative correlations between corruption index and GDP PPC; and corruption and polyarchy index. On the other hand, there is a positive relationship between polyarchy and GDP. This hints at the possible presence of multicollinearity in the model. As the correlation matrix indicates, the inclusion of both these variables may be discouraged given their correlation, which can lead to the undermined statistical significance of our key explanatory variable. However, as corruption is still a possible confounding variable, we justify its inclusion. Additionally, the presence of at least some form of linear relationship between our dependent variable and independent variable is encouraging.

Figure 3. Correlation Matrix



As for the correlation between our main variables, we observe that Pearson's $r = .47$ indicating a moderately strong correlation between the GDP and democracy. The superimposed trend-line confirms this relationship. At the same time, a closer examination (Figure 4) shows the presence of a cluster of observation in high GDP and Polyarchy category - that may skew the correlation coefficient upwards. Additionally, there is a small cluster of large GDP and small (close to non-existing) democracy. It is formed by the observations from the Gulf states, Qatar and United Arab Emirates in particular. It is worth remembering, that the plots depicts panel data - thus, data points for countries are plotted for all the years (1991-2019), thereby overemphasising the effect of clusters of outliers.

Figure 4. Relationship Between Polyarchy and GDP



Additionally, when plotting the Polyarchy scores over time (figure 5), we see that the median scores have mostly remained stable, fluctuating near 0.50 mark. However, there is a profound difference between the rest of the graph and the first two years of data - indicating the ongoing democratisation process. Indeed, during the early years, the Polyarchy score was around 0.30 - much lower than later.

Plotting Polyarchy scores by the Log of GDP PPC shows an interesting picture. Firstly, we can observe the

relationship predicted by modernisation theory - the more wealthy the country is - the higher the polyarchy score. Secondly, it also accounts for the caveat outlined in the theoretical section and pictured in Figure 4- there is, indeed, a selection of countries that are extremely wealthy but are extremely undemocratic. (The outlier - wealthy *and* democratic is Luxembourg). Thirdly, even while the median for wealthy (log GDP PPC = 4) is the highest observed - we can identify a trailing tail of outliers. It can be due to the rounding of the Log GDP PPC to the nearest integer number.

Figure 5. Distribution of Polyarchy Scores over Time

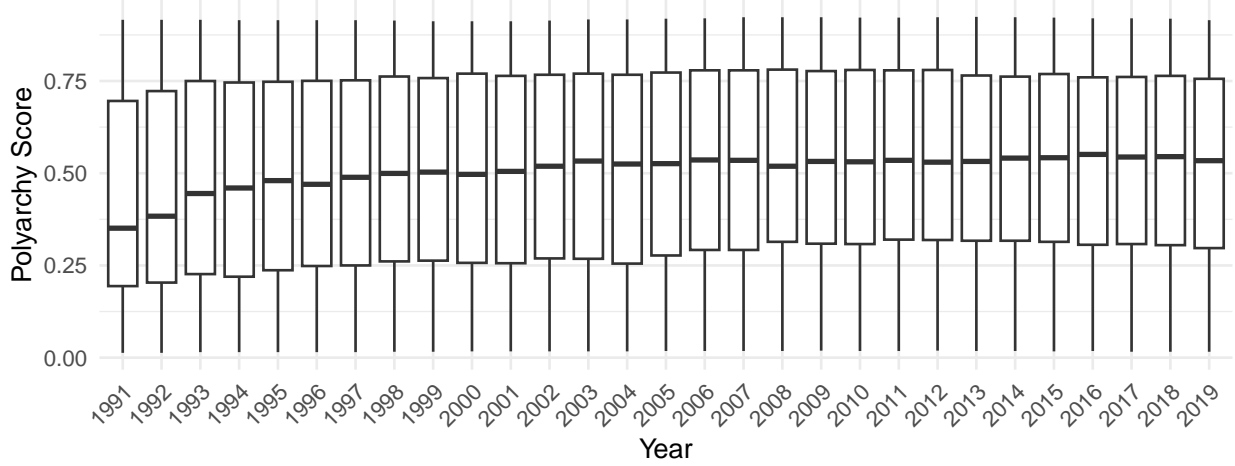
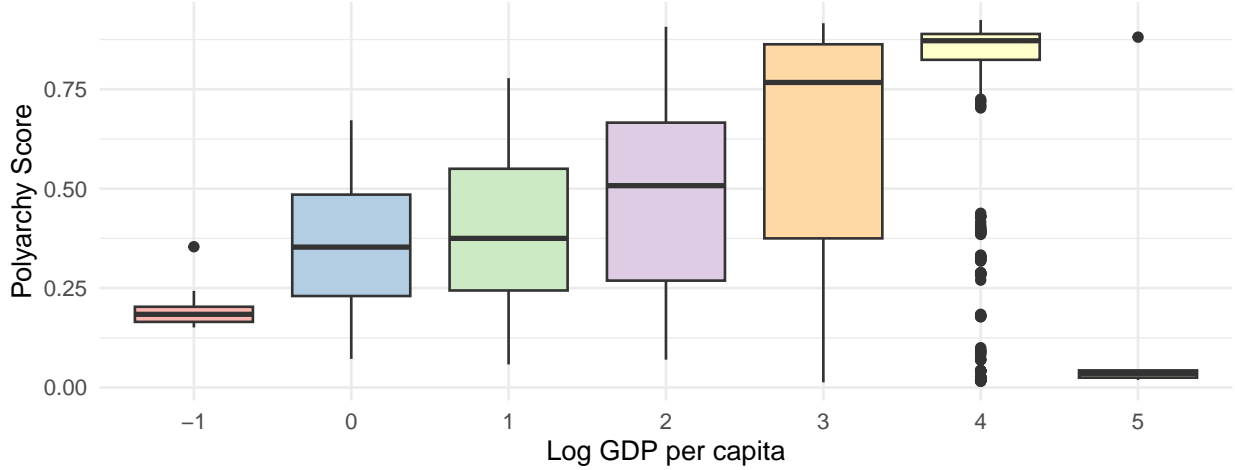


Figure 6. Distribution of Polyarchy Scores by Log GDP per Capita



Having examined our main variables and relations between them, we move to the model building section.

6 Model Building

6.1 Basic Model

To start, we fit a basic model (i.e., it does include a confounding variable but does not control for the FEs) with the following formula:

$$Polyarchy_i = \beta_0 + \beta_1 \text{Log}(\text{GDP}) + \beta_2 \text{CorruptionIndex} + v$$

meaning we predict the polyarchy index using the log of the GDP PPC and the corruption index. For the basic model we use a standard OLS regression. While its usage in such a form is discouraged when working with panel data, we employ it as a base model upon which the final one will be built. The coefficients are the following:

Table 8: Coefficient-Level Estimates for the Basic Model

	<i>Dependent variable:</i>
	Polyarchy Index
Log(GDP PPC)	0.013*** (0.003)
Corruption Index	-0.570*** (0.012)
Constant	0.777*** (0.012)
Observations	4,999
R ²	0.457
Adjusted R ²	0.457
Residual Std. Error	0.198 (df = 4996)
F Statistic	2,100.192*** (df = 2; 4996)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

The initial model gives hope. Let us examine it in more details:

- the F-statistics is high with a p-value lower than a threshold for a 5% significance means that some of the coefficients, at least, do differ from 0.
- Adjusted R^2 is 0.45 meaning the model explains half of the variation in the Dependent Variable, which is rather mediocre.
- Both coefficients are significant at 5% level (p-value < 0.05, furthermore, they are actually significant at 0.1% level (a = 0.001), given that p-values are smaller than 0.001). P-value of regression coefficient tests a hypothesis that its true value is 0 (i.e., it has no effect on the DV). In more technical terms, it means that the probability of observing the estimated coefficient, or a more extreme coefficient, assuming the null hypothesis is true (i.e., it is equal to 0) is less than 5%. Thus, with 95% (even 99.9%) confidence we can conclude that our estimator for the coefficient does not equal to 0 and has some influence over the DV. For our Log GDP PPC key IV it means that we are thus able to reject H_0 that GDP does not influence the level of democracy. Since it is only the basic model, we will not go into the details of interpreting the effect size - something, we will do for the final model.

6.2 Two-Way Fixed-Effects Model

However, this model is far from perfect. The current version does not account for country-specific time-invariant factors that could influence the measure of democracy thus making the estimate of GDP coefficient biased. Probably there are some factors that influence both the predictor and the dependent variable, thus, the assumption of mean independence of error can be violated. Hence, we employ a two-way fixed-effect model controlling for the country and year.

Furthermore, Breusch-Pagan test shows that heteroskedasticity (i.e., non-constant variance of residuals) is present in the model, leading to the possible biased estimation of standard errors (and thus to the incorrect

Table 9: Breusch-Pagan test for the Basic Model

	Statistic	DF	P.value
BP	101.3557	2	<0.001

values of p-values). Hence, the fixed-effect model shall also cluster the errors (Colin Cameron and Miller 2015). Now let us examine the new model.

Table 10: Coefficient-Level Estimates for the TWFE Model

Dependent Variable: Model:	Polyarchy Index (1)
<i>Variables</i>	
Log GDP PPC	0.0120*** (0.0039)
Corruption Index	-0.4354*** (0.0328)
<i>Fixed-effects</i>	
year	Yes
country_name	Yes
<i>Fit statistics</i>	
Observations	4,999
R ²	0.92113
Within R ²	0.13774
<i>Clustered (year) standard-errors in parentheses</i>	
<i>Signif. Codes: ***: 0.01, **: 0.05, *: 0.1</i>	

The final model shows a considerable improvement. The main coefficient remains significant at 5% significance level, thus, allowing us to further reject the H_0 that economic development does not influence the level of democracy, providing an empirical evidence for the modernisation theory. To be more precise, we can reject H_0 with 99% certainty, given that our coefficient is significant at 0.01 level. However, the strength of a relationship is rather small: for every 1% increase in Log GDP PPC, the polyarchy index increases approximately by ~0.0001. At the same time, our control variable - corruption index - shows a more strong effect on the DV. Thus, a unit increase in the corruption index results in a -0.4 change to the democracy *ceteris paribus*. This effect is significant at 0.001 level.

The standard error (SE) for the main independent variable shows that the variability (estimated standard deviation) of the sampling distribution of the coefficient is 0.0039. Its value is instrumental in calculating the p-value for the t-statistic of the coefficient, as is shown in the formula:

$$t = \frac{coef.estimate(\beta)}{StandardError(SE)}$$

Table 11: Additional Statistics for the TWFE Model

Predictor	B	SE	t	p
Log GDP PPC	0.01	0.004	3.09	0.004
Corruption Index	-0.44	0.033	-13.28	<0.001

Table 12: 95% CIs for Final Model

	2.5%	97.5%
Log GDP PPC	0.004	0.020
Corruption Index	-0.503	-0.368

Indeed, for our main coefficient (Log GDP PPC) it translates into $t = 3.09$. Which, coupled with 28 Degrees of Freedom results in a p-value 0.004464 (4.464e-3), just as in our model.

As for the Confidence Intervals (CI), for a β_i coefficient, they can be interpreted as the interval that has a 95% probability to contain the true values of β_i . CIs are calculated as follows:

$$CI_{0.95}^{\beta_i} = [\beta_i \pm 1.96 * SE(\beta_i)]$$

The 95% confidence intervals for our Beta-Coefficients are presented in the Table 12. Since neither of them include 0 we can be sure in the direction of the influence of each of the variables. For example, for our main explanatory variable, we can conclude with 95% certainty that the true value of its coefficient lies between [0.004 and 0.020] (our estimate is 0.01).

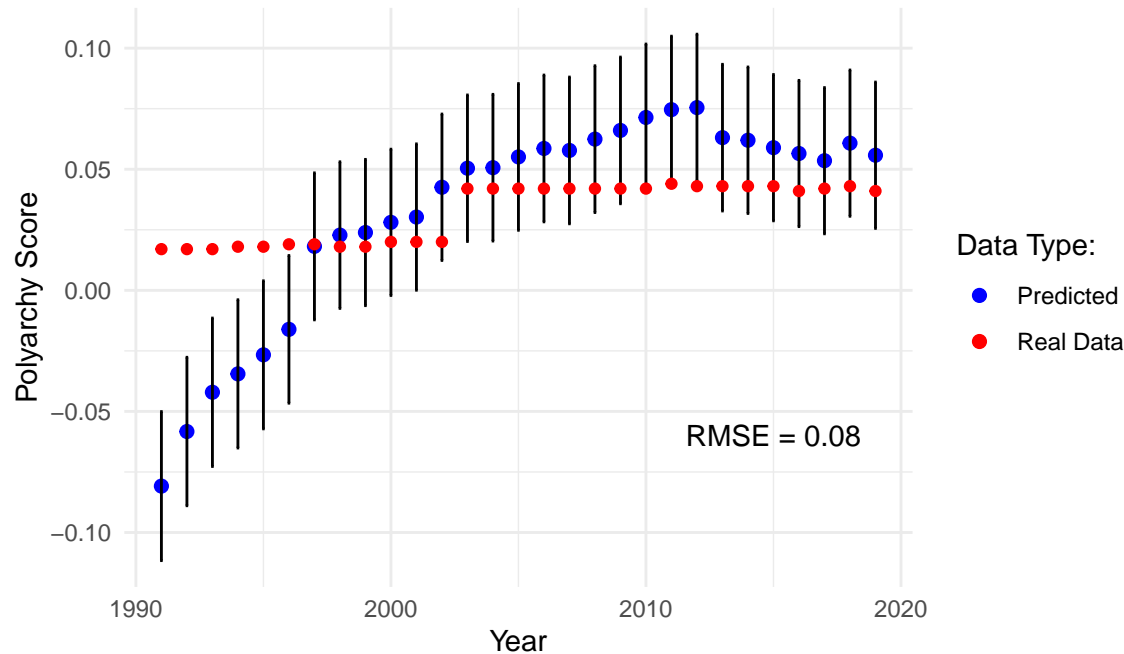
Additionally, the adjusted (for the number of predictors) R^2 is 0.91 - indicating that our model explains more than 90% of variation across the dependent variable, which is a substantial improvement. However, the within R^2 is perniciously low - 0.13 - hinting at the low explanatory power of the model in explaining the variation within the panel units (countries).

6.2.1 Prediction

To substantively test our model¹, we decided to predict the values of Polyarchy index for Qatar - one of the outliers in the data, as we mentioned earlier. We expect a rather poor performance, given that we have already dedicated Qatar as a problem for the modernisation theory. However, as Figure 7 illustrates, from the year 1997 onwards the real data for Qatar falls within the CIs for predicted values - meaning our CIs, indeed, contain a true value of the DV. At the same time, in the beginning of the time period, the model predicted a negative Polyarchy score which cannot be the case, given the scale of measurement. The performance might be impaired by the lack of variation in the real data - the polyarchy score for Qatar is rather stable (the one major deviation corresponding to the less than 0.1 change). The RMSE of the model is 0.08, meaning that the weighted average error between the predictions and the actual data is 0.08, which is relatively mediocre, since our DV is measured from 0 to 1 - meaning we have an 8% error.

¹Full disclosure: we actually test a different model. The issue is that the original model (described above) does not natively work with `predict` function. Our new model is essentially the same (apart from clustering the errors) when compared to the TWFE model with the same parameters (the model used for predictions is using dummy coding to control for FEs). It results in the same coefficients, significance levels, R^2 and other parameters - they are, indeed, almost identical. Therefore, we deem predicting using the typical OLS model with dummy coding valid.

Figure 7. Model Performance
The Case of Qatar



7 Discussion & Future Improvements

In this paper we aimed to contribute to the ongoing debate regarding the relationship between wealth and democracy by testing the hypotheses of modernisation theory using V-Dem data from 1991 to 2019. Given the panel nature of our data, the study employed OLS regression with two-way fixed-effects (TWFE) controls to investigate the potential causal effect between economic development (GDP per capita) and the level of democracy, controlling for both cross-sectional and time-series factors. Our main conclusion appears to favour modernisation hypothesis: *a statistically significant positive relationship between the polyarchy index of democracy and GDP per capita can be observed*, providing empirical evidence for the modernisation theory. However, there are several limitations that must be considered. Firstly, we cannot entirely rule out the possibility of reverse causality, meaning that it is plausible that the causation may also run in the opposite direction - the level of democracy influences the prosperity. Secondly, we acknowledge that there may be additional confounding variables that we have not included in our model, which could affect the relationship between democracy and GDP (thus, undermining the internal validity of our project). For example, colonial legacies or education level can serve as controls in the future research. As we have mentioned above, future scholarship should probably make use of instrumental variable approach to find a variable that is correlated with GDP *but not with democracy level* - to help cement a direction of causality and isolate the effects of GDP to a higher degree that TWFE allows. Lastly, it seems that the overall influence of economic prosperity while significant is rather small overall, as emphasised above. Nevertheless, despite the limitations of the study, it provides a more contemporary and comprehensive analysis of the relationship between wealth and democracy, filling an important gap in the literature.

References

Carothers, Thomas. 2002. "The End of the Transition Paradigm." *Journal of Democracy* 13(1): 5–21. <http://dx.doi.org/10.1353/jod.2002.0003>.

- Colin Cameron, A., and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50(2): 317–72. <http://dx.doi.org/10.3368/jhr.50.2.317>.
- Dahl, Robert A., and Ian Shapiro. 2015. *On Democracy*. 2nd edition. New Haven: Yale University Press.
- Emmanuel, Tlamelo et al. 2021. "A Survey on Missing Data in Machine Learning." *Journal of Big Data* 8(1). <http://dx.doi.org/10.1186/s40537-021-00516-9>.
- Fariss, Christopher, Therese Anders, Jonathan Markowitz, and Miriam Barnum. 2021. "Replication Data for: New Estimates of over 500 Years of Historic GDP and Population Data." <https://dataverse.harvard.edu/citation?persistentId=doi:10.7910/DVN/DC0ING>.
- Greitens, Sheena Chestnut. 2020. "Surveillance, Security, and Liberal Democracy in the Post-COVID World." *International Organization* 74(S1): E169–90. <http://dx.doi.org/10.1017/S0020818320000417>.
- Herman, Lise Esther. 2015. "Re-Evaluating the Post-Communist Success Story: Party Elite Loyalty, Citizen Mobilization and the Erosion of Hungarian Democracy." *European Political Science Review* 8(2): 251–84. <http://dx.doi.org/10.1017/S1755773914000472>.
- Huntington, Samuel P. 1993. 4 *The Third Wave: Democratization in the Late Twentieth Century*. University of Oklahoma press.
- Imai, Kosuke, and In Song Kim. 2020. "On the Use of Two-Way Fixed Effects Regression Models for Causal Inference with Panel Data." *Political Analysis* 29(3): 405–15. <http://dx.doi.org/10.1017/pan.2020.33>.
- Kubbe, Ina, and Annika Engelbert. 2017. "Corruption and the Impact of Democracy." *Crime, Law and Social Change* 70(2): 175–78. <http://dx.doi.org/10.1007/s10611-017-9732-0>.
- Levitsky, Steven, and Lucan A. Way. 2006. "Linkage Versus Leverage. Rethinking the International Dimension of Regime Change." *Comparative Politics* 38(4): 379. <http://dx.doi.org/10.2307/20434008>.
- Lipset, Seymour Martin. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Political Science Review* 53(1): 69–105. <http://dx.doi.org/10.2307/1951731>.
- Liu, Yu, Yu-Sung Su, and Wenquan Wu. 2022. "How Modernization Theory Has Stumbled in China: A Political Interception Perspective." *Democratization* 30(2): 302–24. <http://dx.doi.org/10.1080/13510347.2022.2137793>.
- Mustapha, Nazar. 2014. "The Impact of Corruption on GDP Per Capita." *Journal of Eastern European and Central Asian Research* 1(2). <http://dx.doi.org/10.15549/jeecar.v1i2.76>.
- Ross, Michael L. 2001. "Does Oil Hinder Democracy?" *World Politics* 53(3): 325–61. <http://dx.doi.org/10.1353/wp.2001.0011>.
- Schumpeter, Joseph A. 2013. *Capitalism, Socialism and Democracy*. routledge.
- Treisman, Daniel. 2014. "Income, Democracy, and Leader Turnover." *American Journal of Political Science* 59(4): 927–42. <http://dx.doi.org/10.1111/ajps.12135>.