Financial Development and Educational Attainment: A Cross Country Comparison

Does education covary with financial development? An Instrumental variables panel regression approach

John Karuitha

Monday December 04, 2023

Contents

1 Background

In this analysis, I examine the relationship between financial development and education attainment. My premise is that people who attain higher education are better placed to enter the formal labor market and hence demand financial services such as bank accounts. Education also raises awareness even among people in the informal and semi-formal sectors to better manage and access finance from formal financial intermediaries (Allen et al., 2014). I examine financial development as the number of people (aged 15 years and above) that have an account with a financial intermediary. I capture education attainment using secondary school enrollment ratio to primary school enrollment ratio. I also include variables such as institutional quality, region, trade openness as controls as they are well known drivers of financial development (Klapper & Singer, 2015). I use years of compulsory education as the instrumental variable for education.

We estimate an equation of the form;

$findev_{ij} = \alpha + \beta_1 Education_{ij} + \beta_2 Governance_{ij} + \beta_4 TradeOpenness_{ij} + \epsilon_{ij}$

- Financial Development refers to the extent that people can access and use affordable financial services. In our case, we proxy financial development as the percentage of people with bank accounts in a country.

- We proxy general level of education as high school (secondary school) turnover to primary school turnover.

  Prior: We expect a positive relationship between education and financial development.

- Governance captures the quality of institutions in a country and how well these institutions improve the quality of life of the citizens. We use the first principle component of the Worldwide Governance Indicators to proxy governance or institutional quality.

  Prior: We expect a positive relationship between governance and financial development.

- Trade Openness: The extent to which a country trades with the outside world. We proxy this using the ratio of imports and exports to GDP.

  Prior: We expect a positive relationship between trade openness and financial development.

Instrumental variables: I use years of compulsory education in a country as the instrumental variable for education. Rationale: Years of compulsory education can be a good instrument for the education variable, because it is correlated with education levels but is not directly related to financial development. Similarly, I use the air freight of passengers and press freedom index as instruments for trade openness and governance, respectively.

2 Data

We source the data from the World Bank, World Development Indicators (WDI) and the Worldwide Governance Indicators (WGI) for the years 1998-2002. The data consists of 8875 observations of 11 variables. The data consists of the following variables (see Table 1).

Instrumental variables: I use years of compulsory education in a country as the instrumental variable for education. Rationale: Years of compulsory education can be a good instrument for the education variable, because it is correlated with education levels but is not directly related to financial development. Similarly, I use the air freight of passengers and press freedom index as instruments for trade openness and governance, respectively. All this data is available in the World Bank Website.

3 Data Exploration

I explore the data by creating summary tables and visualizations. First, I examine the distribution of the variables (see Figure 1).

Table 1: Variables Definition

Variable	Definition
time	Year the data collected, 1998-2022.
country_name	Name of the country.
country_code	ISO3c country code
Continent	Continent of associated country.
Accounts	% of people aged 15+ with an account
Governance	KKM Governance Index, defined as the first principal component of the KKM indicators
Educ	Educational attainment- ratio of high school turnover tp primary school turnover
Openess	Trade openess, defined as the ratio of imports and exports to GDP
compulsory_education	Number of years of compulsory education in country.
Press freedom index	The data is collected through an online questionnaire sent to journalists, media lawyers, researchers and other media specialists selected by Reporters without Borders (RSF) in the 180 countries covered by the Index.
Air Transport, passengers carried	Passengers carried by air into and out of a country.

Note:

Source: The World Bank, https://databank.worldbank.org/source/world-development-indicators

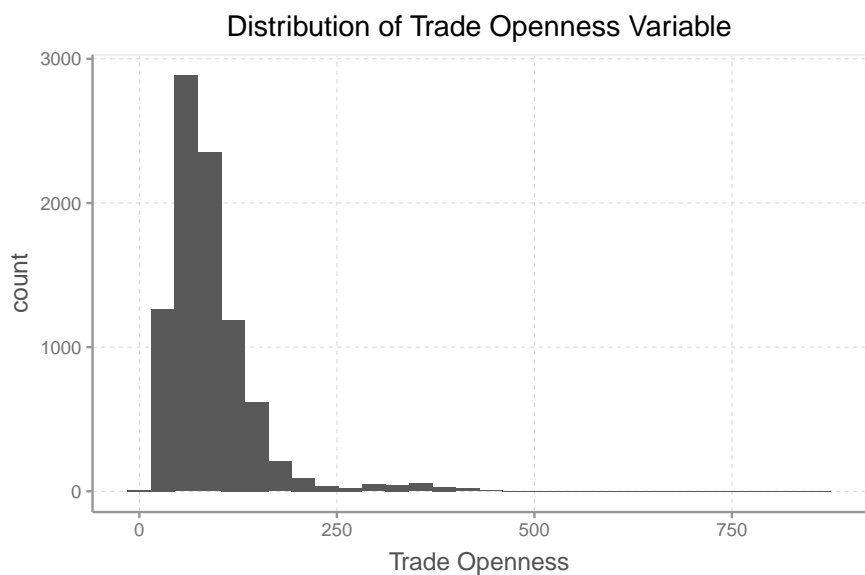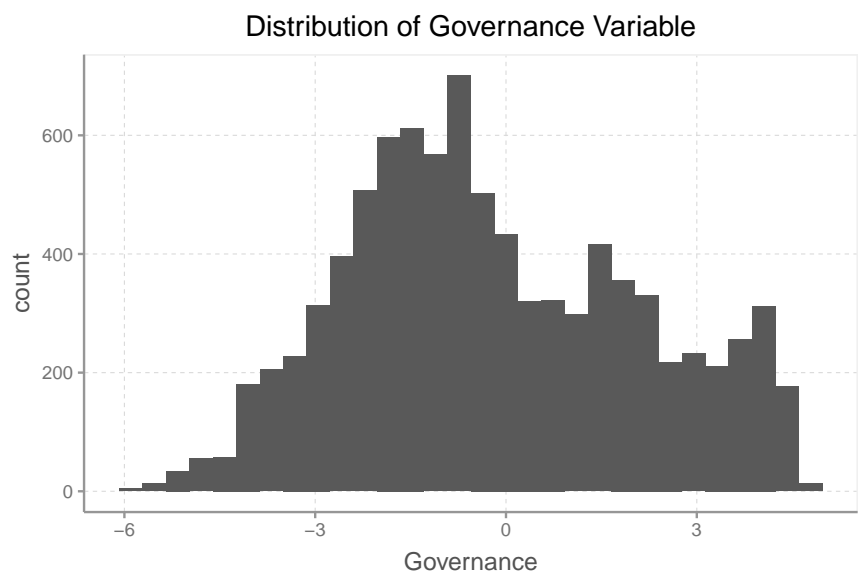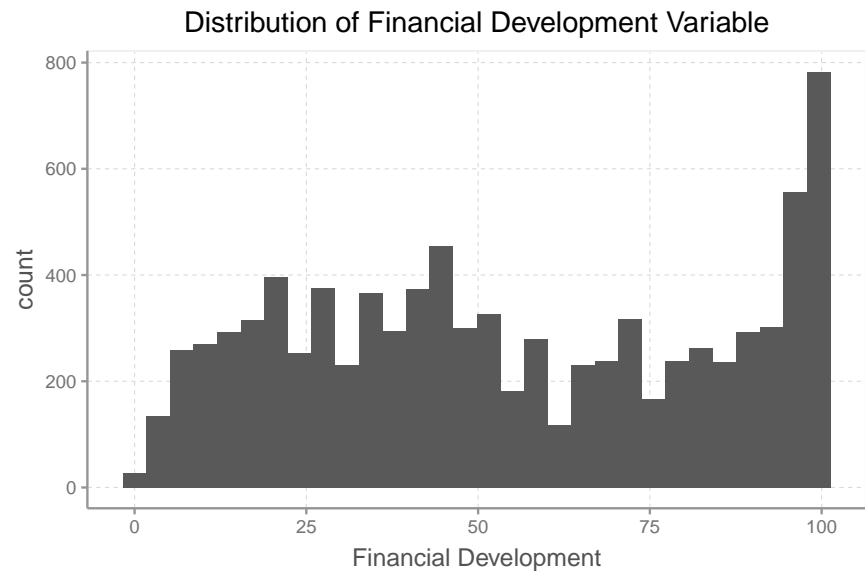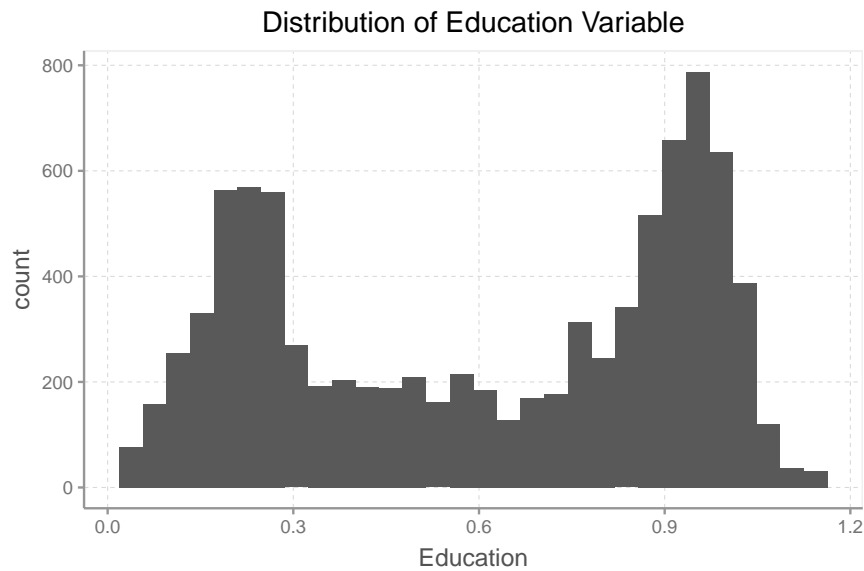[1] The last three variables serve as instruments

Figure 1: Distribution of the variables

Next, I look at the relationship between education and financial development by continent. We see a consistently strong positive relationship between the two variables across all continents.

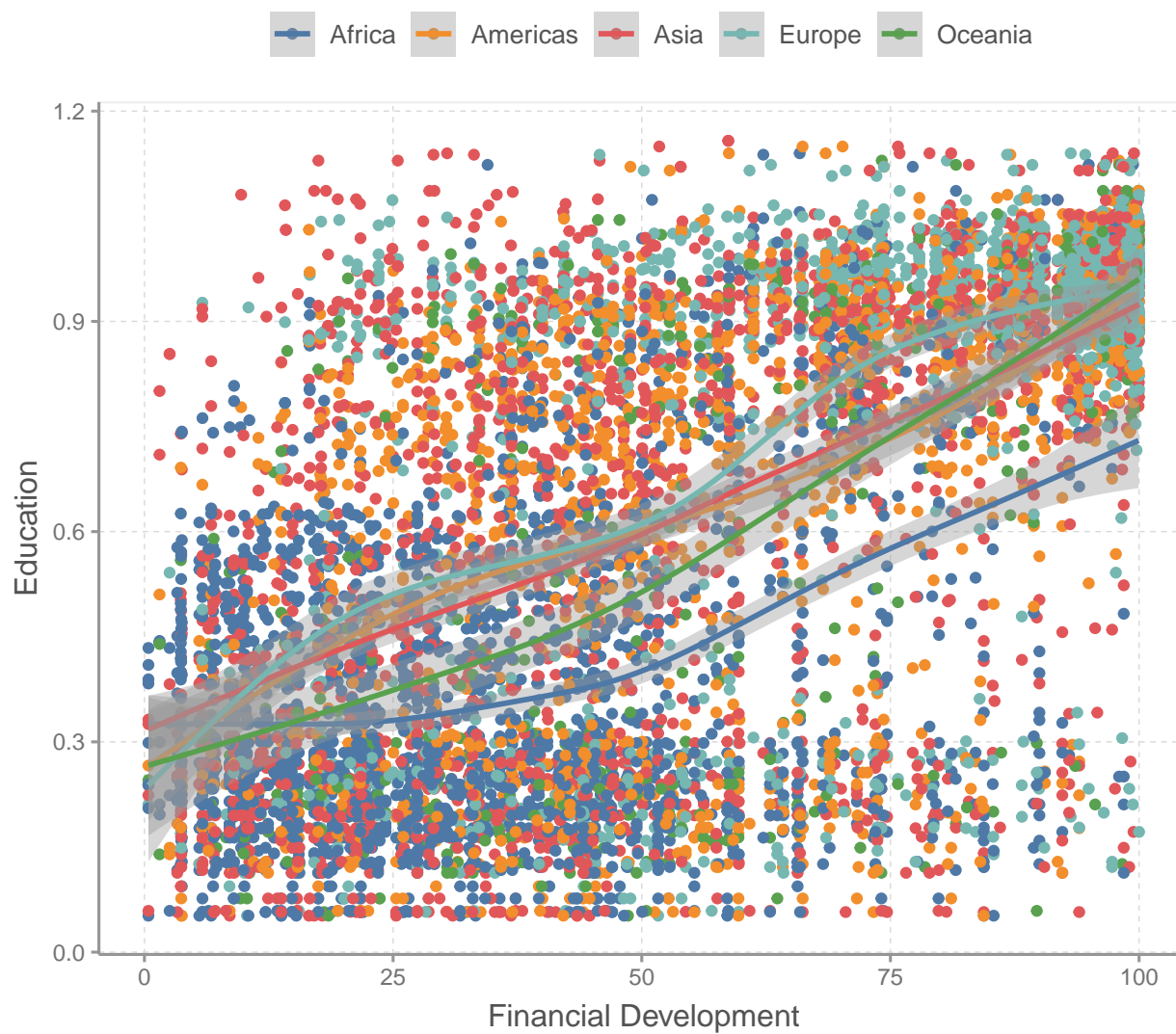## Scatterplot of Financial Development against Education

Figure 2: Scatterplot of Financial Development against Education
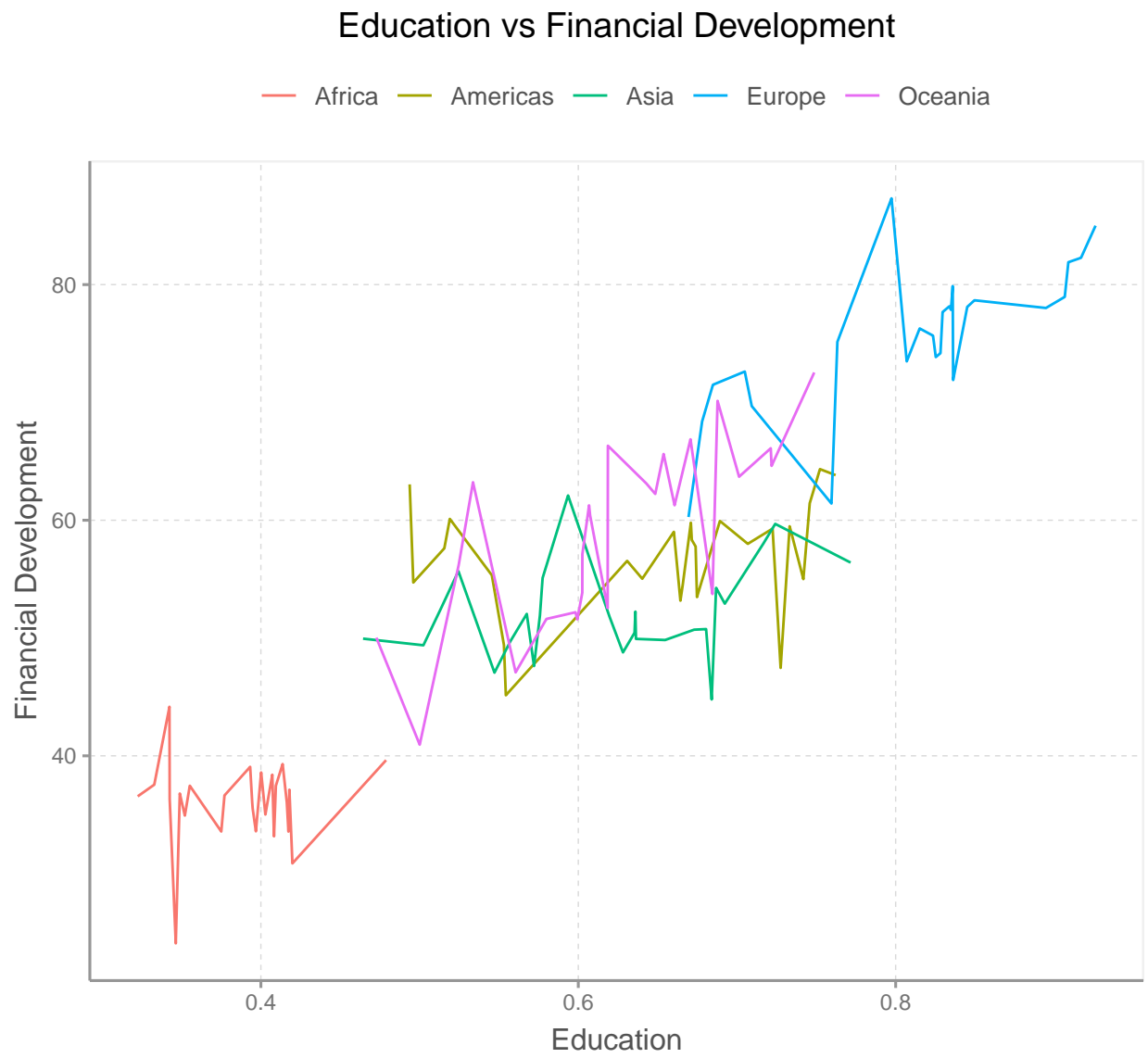
Education vs Financial Development

Figure 3: Education vs Financial Development

I then examine the average trends in financial development by continent over time. We see that Africa lies far below the other regions.

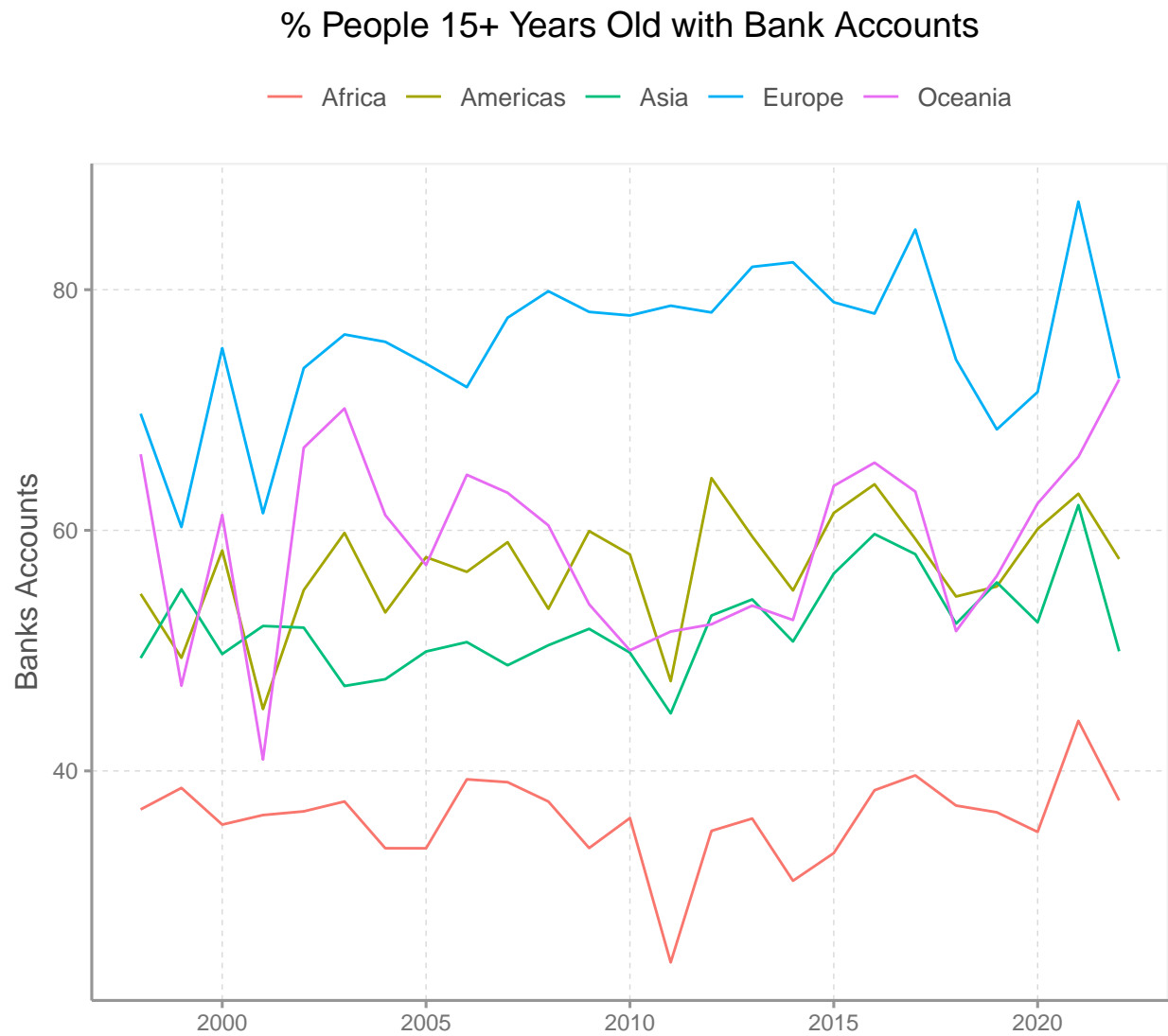## % People 15+ Years Old with Bank Accounts

Figure 4: % People 15+ Years Old with Bank Accounts

The figure below reinforces the low level of financial development in Africa, with Europe exhibiting the highest median financial development.
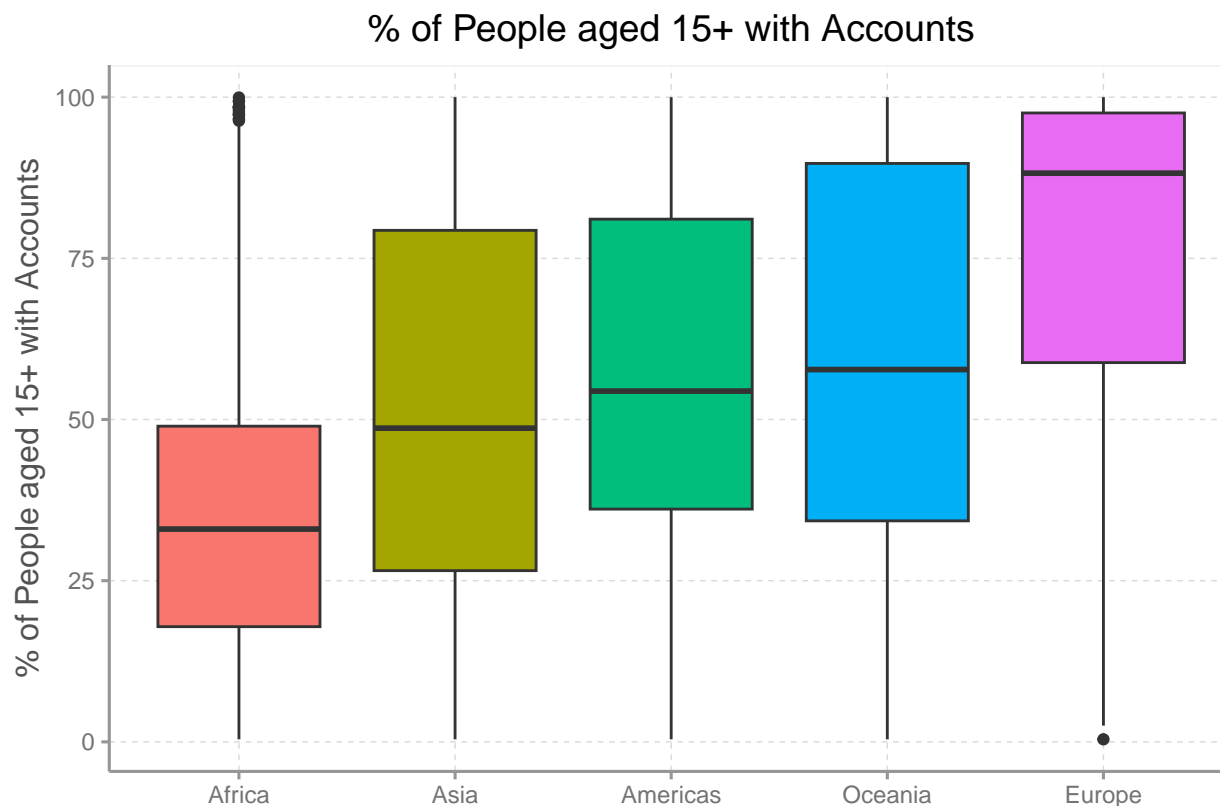
% of People aged 15+ with Accounts

Figure 5: % of People aged 15+ with Accounts

I also look at the correlation between these financial development and the other variables. We see a very high correlation between financial development (accounts) and governance and education. Financial development has moderate relationship with trade openness. The high correlation between education and governance could bring about multicollinearity and cross-sectional dependence problems.

I also summarise the data below

Table 2: Summary Statistics

| skim_type	Variable	Mean	SD	Min	Q1	Median	Q3	Max
numeric	accounts	54.529	30.195	0.400	28.570	50.760	83.560	100.00
numeric	governance	-0.198	2.288	-5.910	-1.896	-0.576	1.574	4.78
numeric	educ	0.608	0.326	0.052	0.268	0.665	0.921	1.16
numeric	openess	89.432	57.769	2.699	54.632	77.116	106.332	863.20
numeric	compulsory_education	8.947	2.291	0.000	8.000	9.000	10.000	17.00
numeric	air_freight	13.937	2.432	0.000	12.515	14.044	15.454	20.65
numeric	press_freedom	66.228	45.364	-10.000	28.000	63.440	90.000	180.00
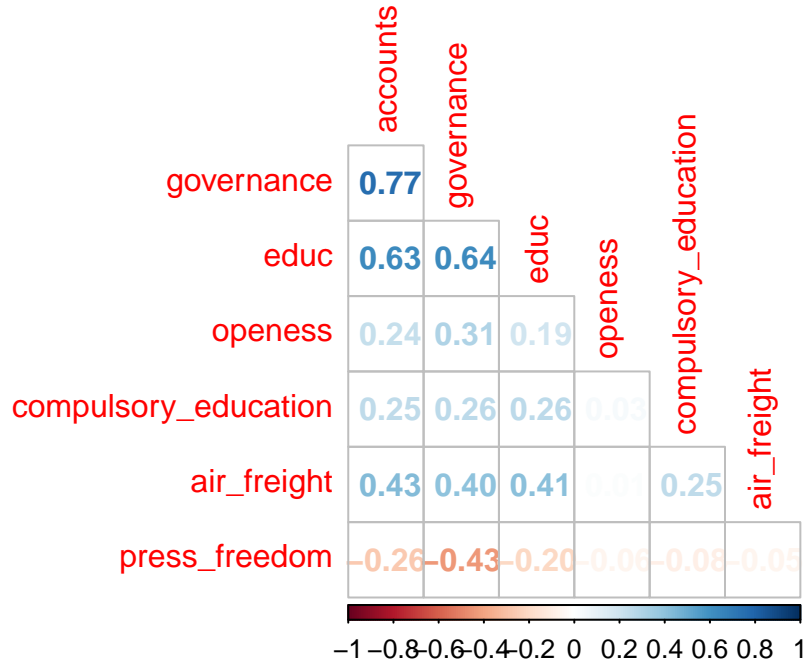
Figure 6: Correlation Analysis

4 Regression Analysis

I run panel regressions as follows;

- Fixed effects model.
- Random effects model.
- Pooled OLS

We estimate an equation of the form;

$$findev_{ij} = \alpha + \beta_1 Education_{ij} + \beta_2 Governance_{ij} + \beta_4 Trade_o penness_{ij}$$

The individual regression output follows.

4.1 The fixed effects model

4.1.1 Fixed effects without Instruments

Below is a summary of the fixed effects model without instruments.

```
## Twoways effects Within Model
##
## Call:
## plm(formula = accounts ~ educ + openness + governance, data = my_data,
##     effect = "twoway", model = "within", index = c("continent",
##         "time"))
##
## Balanced Panel: n = 5, T = 25, N = 8875
##
## Residuals:
##    Min. 1st Qu.  Median 3rd Qu.    Max.
## -58.267 -12.367   0.404  11.173  82.002
##
```

```
## Coefficients:
##            Estimate Std. Error t-value          Pr(>|t|)
## educ       21.76081    0.81517   26.69 <0.0000000000000002 ***
## openess    -0.00196    0.00357   -0.55                0.58
## governance  8.04936    0.12505   64.37 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    6210000
## Residual Sum of Squares: 2930000
## R-Squared:      0.528
## Adj. R-Squared: 0.526
## F-statistic: 3295.38 on 3 and 8843 DF, p-value: <0.0000000000000002
```

4.1.2 Fixed Effects with Instruments

Then the fixed effects model WITH instruments.

```
## Twoways effects Within Model
## Instrumental variable estimation
##
## Call:
## plm(formula = accounts ~ educ + openess + governance | compulsory_education +
##     air_freight + press_freedom, data = my_data, effect = "twoway",
##     model = "within", index = c("continent", "time"))
##
## Balanced Panel: n = 5, T = 25, N = 8875
##
## Residuals:
##    Min. 1st Qu.  Median 3rd Qu.    Max.
##  -76.94  -19.43   -4.49   14.62  296.60
##
## Coefficients:
##            Estimate Std. Error z-value Pr(>|z|)
## educ         28.683     24.176    1.19  0.23545
## openess      -0.446      0.168   -2.66  0.00786 **
## governance    7.167      1.883    3.81  0.00014 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    6210000
## Residual Sum of Squares: 8540000
## R-Squared:      0.122
## Adj. R-Squared: 0.118
## Chisq: 989.338 on 3 DF, p-value: <0.0000000000000002
```

4.2 Pooled OLS Model

4.2.1 Pooled OLS without Instruments

Here, I present the pooled OLS model without instruments

```
## Pooling Model
##
## Call:
## plm(formula = accounts ~ educ + openess + governance, data = my_data,
```

```
##      effect = "twoway", model = "pooling", index = c("continent",
##         "time"))
##
## Balanced Panel: n = 5, T = 25, N = 8875
##
## Residuals:
##     Min. 1st Qu.  Median 3rd Qu.     Max.
## -59.218 -12.368   0.606  11.467   81.153
##
## Coefficients:
##               Estimate Std. Error t-value         Pr(>|t|)
## (Intercept) 42.12694    0.61939   68.01 <0.0000000000000002 ***
## educ        22.77823    0.78150   29.15 <0.0000000000000002 ***
## openess      0.00174    0.00357    0.49              0.63
## governance   8.04644    0.11469   70.16 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    8090000
## Residual Sum of Squares: 3040000
## R-Squared:       0.624
## Adj. R-Squared: 0.624
## F-statistic: 4916.52 on 3 and 8871 DF, p-value: <0.0000000000000002
```

4.2.2 Pooled OLS with Instruments

Then look at the pooled OLS with instruments

```
## Pooling Model
## Instrumental variable estimation
##    (Balestra-Varadharajan-Krishnakumar's transformation)
##
## Call:
## plm(formula = accounts ~ educ + openess + governance | compulsory_education +
##     air_freight + press_freedom, data = my_data, effect = "twoway",
##     model = "pooling", index = c("continent", "time"))
##
## Balanced Panel: n = 5, T = 25, N = 8875
##
## Residuals:
##    Min. 1st Qu.  Median 3rd Qu.    Max.
##  -91.44  -25.70   -6.43   17.65  434.32
##
## Coefficients:
##             Estimate Std. Error z-value Pr(>|z|)
## (Intercept)   76.334     39.832    1.92   0.0553 .
## educ          57.623     12.494    4.61 0.000004 ***
## openess       -0.621      0.370   -1.68   0.0934 .
## governance     6.293      2.099    3.00   0.0027 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    8090000
## Residual Sum of Squares: 14700000
## R-Squared:       0.129
```

```
## Adj. R-Squared: 0.129
## Chisq: 1281.95 on 3 DF, p-value: <0.0000000000000002
```

4.3   Combined results

I combine the regression output into one table below.

Table 3: Panel Regression Analysis

	Dependent variable:			
	Financial Development			
	Fixed Effects	Fixed Effects (Instruments)	Pooled OLS	Pooled OLS (Instruments)
	(1)	(2)	(3)	(4)
Education	21.800***	28.700	22.800***	57.600***
	(0.815)	(24.200)	(0.781)	(12.500)
Trade Openness	−0.002	−0.446***	0.002	−0.621*
	(0.004)	(0.168)	(0.004)	(0.370)
Governance	8.050***	7.170***	8.050***	6.290***
	(0.125)	(1.880)	(0.115)	(2.100)
Constant			42.100***	76.300*
			(0.619)	(39.800)
Observations	8,875	8,875	8,875	8,875
R^2	0.528	0.122	0.624	0.129
Adjusted R^2	0.526	0.118	0.624	0.129
F Statistic	3,295.000*** (df = 3; 8843)	989.000***	4,917.000*** (df = 3; 8871)	1,282.000***

Note: *p<0.1; **p<0.05; ***p<0.01

The output shows that the three variables have a statistically significant relationship with financial development. Holding governance and trade openness constant, a better education is positively associated with higher average financial development. Governance has a positive relationship with financial development, ceteris paribus. On the contrary, trade openness has a marginal negative relationship with financial development. The relationship is uniform across the models. The models without instruments have a better explanatory power than models with instruments. For instance, the fixed effects model with instruments have an R^2 of 52.6% compared to 11.8% for the model with instruments. Similarly, the pooled OLS without instruments has an adjusted R^2 of 62.4% compared to the pooled OLS with instruments that has an adjusted R^2 of 12.9%. But which of these models is better? Let us examine the AIC.

Table 4: AIC for Panel Models

Model	AIC
Pooled OLS with instruments	-45482
Fixed effects with instruments	-43076
Pooled OLS	-38490
Fixed effects	-38331

This analysis shows that the pooled OLS with instruments is the better model, followed by the fixed effects with instruments.

The models could be subject to omitted variable bias though it does fairly well with a R^2. Hence increased investments in raising institutional quality and education attainment should be encouraged.

Policy recommendation: Given that our analysis is between education and financial development, we recomend substantial investments in education to raise the levels of financial development in the long term. This recomendation is especially pertinent for regions with low levels of financial development like Africa.

5 Conclusion

In this analysis, we have examined the relationship between financial development proxied by the number of bank accounts and three variables that are postulated to drive financial development; governance, education, and trade openness. We find that Africa lags in financial development. All the three measures have a statistically significant relationship with financial development, and the models are all significant going by the F-statistic. Hence increased investments in raising education attainment should be encouraged holding institutional quality trade openness constant in order to raise financial development in the long term.

6 Acknowledgements

References

Allaire, J., Xie, Y., Dervieux, C., McPherson, J., Luraschi, J., Ushey, K., . . . Iannone, R. (2023). *rmarkdown: Dynamic documents for r.* Retrieved from https://github.com/rstudio/rmarkdown

Allen, F., Carletti, E., Cull, R., Qian, J. 'QJ'., Senbet, L., & Valenzuela, P. (2014). The african financial development and financial inclusion gaps. *Journal of African Economies*, *23*(5), 614–642.

Arel-Bundock, V. (2022). *WDI: World development indicators and other world bank data.* Retrieved from https://CRAN.R-project.org/package=WDI

Arel-Bundock, V., Enevoldsen, N., & Yetman, C. (2018). countrycode: An r package to convert country names and country codes. *Journal of Open Source Software*, *3*(28), 848. Retrieved from https://doi.org/10.21105/joss.00848

Arnold, J. B. (2023). *ggthemes: Extra themes, scales and geoms for "ggplot2".* Retrieved from https://CRAN.R-project.org/package=ggthemes

Corporation, M., & Weston, S. (2022). *doParallel: Foreach parallel adaptor for the "parallel" package.* Retrieved from https://CRAN.R-project.org/package=doParallel

Croissant, Y., & Millo, G. (2008). Panel data econometrics in R: The plm package. *Journal of Statistical Software*, *27*(2), 1–43. https://doi.org/10.18637/jss.v027.i02

Croissant, Y., & Millo, G. (2018). *Panel data econometrics with R.* Wiley.

Csárdi, G., Hester, J., Wickham, H., Chang, W., Morgan, M., & Tenenbaum, D. (2023). *remotes: R package installation from remote repositories, including "GitHub".* Retrieved from https://CRAN.R-project.org/package=remotes

Dancho, M. (2020). *correlationfunnel: Speed up exploratory data analysis (EDA) with the correlation funnel.* Retrieved from https://CRAN.R-project.org/package=correlationfunnel

Firke, S. (2023). *janitor: Simple tools for examining and cleaning dirty data.* Retrieved from https://CRAN.R-project.org/package=janitor

Hlavac, M. (2022). *stargazer: Well-formatted regression and summary statistics tables.* Bratislava, Slovakia: Social Policy Institute. Retrieved from https://CRAN.R-project.org/package=stargazer

Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A program for missing data. *Journal of Statistical Software*, *45*(7), 1–47. https://doi.org/10.18637/jss.v045.i07

Klapper, L., & Singer, D. (2015). The role of informal financial services in africa. *Journal of African Economies*, *24*(suppl_1), i12–i31.

Millo, G. (2017). Robust standard error estimators for panel models: A unifying approach. *Journal of Statistical Software*, *82*(3), 1–27. https://doi.org/10.18637/jss.v082.i03

Pedersen, T. L. (2023). *patchwork: The composer of plots.* Retrieved from https://CRAN.R-project.org/package=patchwork

Piburn, J. (2020). *wbstats: Programmatic access to the world bank API.* Oak Ridge, Tennessee: Oak Ridge National Laboratory. Retrieved from https://doi.org/10.11578/dc.20171025.1827

R Core Team. (2023). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rinker, T. W., & Kurkiewicz, D. (2018). *pacman: Package management for R.* Buffalo, New York. Retrieved from http://github.com/trinker/pacman

Smeets, B. (2023). *artyfarty: Themes for ggplot2.*

Tierney, N., & Cook, D. (2023). Expanding tidy data principles to facilitate missing data exploration, visualization and assessment of imputations. *Journal of Statistical Software*, *105*(7), 1–31. https://doi.org/10.18637/jss.v105.i07

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in r. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Waring, E., Quinn, M., McNamara, A., Arino de la Rubia, E., Zhu, H., & Ellis, S. (2022). *skimr: Compact and flexible summaries of data.* Retrieved from https://CRAN.R-project.org/package=skimr

Wei, T., & Simko, V. (2021). *R package "corrplot": Visualization of a correlation matrix.* Retrieved from https://github.com/taiyun/corrplot

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., . . . Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Xie, Y. (2014). knitr: A comprehensive tool for reproducible research in R. In V. Stodden, F. Leisch, & R. D. Peng (Eds.), *Implementing reproducible computational research.* Chapman; Hall/CRC.

Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Boca Raton, Florida: Chapman; Hall/CRC. Retrieved from https://yihui.org/knitr/

Xie, Y. (2023). *knitr: A general-purpose package for dynamic report generation in r.* Retrieved from https://yihui.org/knitr/

Xie, Y., Allaire, J. J., & Grolemund, G. (2018). *R markdown: The definitive guide.* Boca Raton, Florida: Chapman; Hall/CRC. Retrieved from https://bookdown.org/yihui/rmarkdown

Xie, Y., Dervieux, C., & Riederer, E. (2020). *R markdown cookbook.* Boca Raton, Florida: Chapman; Hall/CRC. Retrieved from https://bookdown.org/yihui/rmarkdown-cookbook

Zhu, H. (2023). *kableExtra: Construct complex table with "kable" and pipe syntax.*