Ofori, Isaac Kwesi

Preprint

Catching The Drivers of Inclusive Growth in Sub-Saharan Africa: An Application of Machine Learning

EXCAS Working Paper, No. 21/044

European **X**tramile **C**entre of **A**frican **S**tudies
(EXCAS)

EXCAS Working Paper

WP/21/044

Catching the Drivers of Inclusive Growth in Sub-Saharan Africa: An Application of Machine Learning

Isaac K. Ofori
School of Economics,
University of Insubria,
Via Monte Generoso, 71, 21100, VA, Varese, Italy.
Email: isaac.ofori1@ucc.edu.gh
ikofori@uninsubria.it;

2021 European Xtramile Centre of African Studies WP/21/044

Research Department

Catching the Drivers of Inclusive Growth in Sub-Saharan Africa: An Application of Machine Learning

Isaac K. Ofori

January 2021

Abstract

A conspicuous lacuna in the literature on Sub-Saharan Africa (SSA) is the lack of clarity on variables key for driving and predicting inclusive growth. To address this, I train the machine learning algorithms for *the Standard lasso, the Minimum Schwarz Bayesian Information Criterion (Minimum BIC) lasso,* and *the Adaptive lasso* to study patterns in a dataset comprising 97 covariates of inclusive growth for 43 SSA countries. First, the regularization results show that only 13 variables are key for driving inclusive growth in SSA. Further, the results show that out of the 13, the poverty headcount (US$1.90) matters most. Second, the findings reveal that '*Minimum BIC lasso'* is best for predicting inclusive growth in SSA. Policy recommendations are provided in line with the region's green agenda and the coming into force of the African Continental Free Trade Area.

Keywords: *Clean Fuel, Economic Growth, Machine Learning, Lasso, Sub-Saharan Africa, Regularization, Poverty.*

JEL Codes: *C01; C14; C51; C52; C55; F43; O4; O55*

1.0 Introduction

The question as to what really drives inclusive growth, particularly in sub-Saharan Africa (SSA) has been debated upon for decades. Despite the coming into force of the 17-point sustainable development goals, in the SSA, the focus has largely been on economic growth and poverty instead of shared prosperity (World Bank 2020a). In fact, empirical works on what matters for inclusive in the world's most disadvantaged region remain scanty (Ravallion and Chen 2019; Greenwald and Stiglitz 2013). Conspicuously, a review of the literature shows that there has been a lack of clarity on variables policymakers should target to foster shared prosperity in the developing world[1]. For instance, empirical studies such as Van Niekerk (2020), Mutiiria (2020), Gyamfi (2020), Tella and Alimi (2016), and Zhuang and Ali (2010) provide conflicting results[2] on the effects of covariates such as education, trade openness, government expenditure, institutions, financial development, and foreign direct investment on inclusive growth in SSA.

The relevance of cutting-edge research on inclusive growth is made clear by the disruptive effects of the coronavirus pandemic[3] (Brown *et al.* 2020). Indeed, since the pandemic struck, opinions on how policymakers can build a resilient and all-inclusive SSA continue to dominate the media, political and academic landscapes. However, rigorous empirical content backing such claims in terms of focus and modelling of inclusive growth is hard to find. For instance, albeit bereft of empirical backing, the Bretton Woods institutions— the IMF and World Bank (2020) identify resource allocation[4] as a possible channel through which the welfare setbacks due to covid-19 can be mitigated. However, with development finance from the tax systems and donor agencies expected to fall due to the slowdown in economic activity (OECD 2020), and the fact that recent growth in SSA is driven largely by external developments (UNCTAD 2020; UNECA 2019), achieving sustainable and equitable growth in the medium to long-term would be challenging. This underscores the need for researchers to inform policy in terms of areas crucial for driving/predicting inclusive growth. This fundamentally forms the lacuna in the literature, particularly on SSA, where inclusive growth predictors are selected based researchers' preference/discretion. The challenge with subjective/preferential selection of covariates and

[1] For instance, the choice of structural/macroeconomic, policy and institutional variables have been subjective, clearly raising concerns as to which area need prioritization due to resource constraints.
[2] This is plausibly due to differences in model assumptions and specifications the authors use.
[3] In 2020, the SSA contracted by at least 3 per cent from a mild 0.8 per cent in 2019 (World Bank 2020b; IMF 2020a).
[4] Expenditure on health, and direct transfers to the poor.

the application of traditional techniques such as the ordinary least squares is that: (1) traditional techniques cannot yield sparse results in large datasets, and (2) even weak covariates may be deemed influential under some model assumptions and specifications. More germane, in large datasets like one underpinning this study, using traditional techniques for inference may yield inefficient results due to overfitting/misspecification of the model. However, in the advent of machine learning regularization techniques, the selection of covariates for inference/prediction need not be preferential as algorithms can be trained to study patterns in datasets and catch the salient drivers of the outcome variable (James *et al.* 2013, Zou 2006; Zou and Hastie 2005; Tibshirani 1996). This is where this study contributes to the literature. First, I employ recent advances in machine learning for catching key drivers of inclusive growth in SSA. Second, the study uses machine learning techniques to identify a model best for predicting inclusive growth in SSA.

The rest of the paper is organized as follows: the next section is dedicated to a brief review of the literature on inclusive growth as we as studies employing machine learning techniques. Section 3 also presents the methodological foundation of the paper. The results and discussions are presented in section 4 while chapter 5 concludes with some policy implications.

2.0 Literature survey on measures and drivers of inclusive growth

Achieving economic growth is one thing while achieving shared prosperity is another. If there is any region in the world in need of attention in terms of policy recommendations in fostering inclusive growth, then it is the SSA. In growth sense, the region has felt the brunt of the covid-19 pandemic and of more concern is the bleak outlook on unemployment, precarious employment, poverty and inequality (World Bank 2020a, ILO 2020). However, opinions on what inclusive growth is and what really drives it are varied. For instance, Ravallion (2004) define inclusive growth as one that is largely beneficial to the poor and marginalized (i.e., sustained growth in GDP per capita). Conversely, the IMF (2011) define inclusive growth as growth in incomes of the poor relative to that of the overall population. Taking cues from the absolute and relative perspectives of inclusive growth, Ali and Son (2007) define inclusive growth as one that increases social opportunities in terms of incomes, employment, human capital development, and social safety nets.

Inclusive growth thus encompasses several facets of national development particularly with regards to the creation of equitable opportunities aimed at increasing the incomes, welfare and participation of especially the poor in economic development (Berg and

Ostry 2011; Commission on Growth and Development 2008). A survey of the literature shows that inclusive growth is driven by factors such as globalisation, foreign direct investment, trade openness, and inflation (Anand *et al.* 2013); economic growth, employment (Paramasivan et al. 2014); human capital development, gender equality, and social safety nets (World Bank 2013, 2009; Acemoglu and Robinson 2012; Lustig *et al.* 2012; Zhuang and Ali 2010); resource allocation, infrastructural development, education, and healthcare (Calderón and Servén 2014; Asian Development Bank 2013; Gajigo and Lukoma 2011). While these authors make contributions to the literature, the challenge is that such proliferation of drivers makes it difficult for policymakers to plan and target inclusive growth, signifying the relevance of this paper.

2.1 Brief survey of empirical works applying machine learning techniques in economics

The literature on inclusive growth is vast and an attempt to present them will be a daunting one[5]. The study therefore presents some economics-related empirical works employing machine learning. First is the study by Schneider and Wagner (2009) who focus solely on the machine learning algorithm of lasso (least absolute shrinkage and selection operator) in catching key drivers of growth in the NUTS2 region[6] of the European Union over the period 1995 – 2005. The authors find that variables such as initial GDP per capita, human capital, and initial unemployment rate matter for economic growth. A similar work is that of Dutt and Tsetlin (2016) who employed the elasticnet and lasso regularization techniques to identify which income distribution measure(s)[7] matter(s) for development outcomes— per capita income, schooling, and institutional quality. The authors find that the poverty headcount indicator matters most in predicting the three development outcomes compared to all other 36 distributional measures.

Similarly, Tkacz (2001) applied neural network algorithms in forecasting GDP growth in Canadian. He finds that, relative to traditional methods such as linear and univariate forecasting methods, neural network techniques yield lower forecast errors on annual growth rate. They however indicate that neural techniques perform better in forecasting long-term growth rather than short-term growth. Richardson *et al.* (2021) also explore the machine learning techniques of support-vector machine, neural network, lasso, boosted tree, model and ridge, relative to classical methods in forecasting real GDP growth in

[5] While Mitra and Das (2018) identify 24 variables, the Asian Development Bank (2013) identifies 35 variables as influential for inclusive growth
[6]Nomenclature of Territorial Units for Statistics
[7] In all, a total of 37 income distribution measures were used

New Zealand. They find that machine learning algorithms outperform classic statistical methods in prediction. A similar work in terms of the empirical focus of this study is that of Jung *et al.* (2018) who employ machine learning algorithms of lasso, ridge, elasticnet, neural networks, and super learner in forecasting the GDP growth of the G7 countries. The authors provide strong evidence to conclude that machine learning algorithms outperformed standard prediction techniques.

In the case of SSA, however, attention on how relevant these techniques can be in aiding policymakers plan and target growth is weak. This study shows how machine learning techniques can prove momentous in helping policymakers shape policies to foster inclusive growth in the SSA.

3.0 Data and methodology

3.1 Data

The dataset underpinning the analysis is sourced from a number of micro and macro databases. For the purposes of cross-country analysis, the microdata, which are sourced from recognised surveys[8] are aggregated. The macrodata covers structural, institutional, globalisation, and income distribution indicators drawn from (1) the global consumption and poverty project (Lahoti *et al.* 2016); the Konjunkturforschungsstelle (KOF) index (Dreher 2006; Gygli *et al.* 2019); the World Bank's poverty and equity database, and the world development indicators (World Bank 20201). The dataset spans 1980 – 2019 with 97 covariates. Data on the outcome variable, inclusive growth is however generated following the approach of Anand *et al.* (2013) (*see Appendix A*). The description of the variables is provided in Table A1 (see Appendix B)

3.2 Estimation strategy

Despite the BLUE[9] property of the classical least square estimator, in considerably large datasets like one underpinning this analysis, the least square model is not only less sparse but also, more susceptible to problems like multicollinearity and outliers. That is, as the dataset become large, least square assumptions of no multicollinearity, homoscedasticity, and exogeneity typically break down and therefore overfitting the training sample, causing the out of sample error to increase. Mitigating this challenge is through the use of machine

[8] The surveys on clean fuel, sanitation, under-5 mortality, social equity, HIV prevalence, prenatal care, school enrolment, mobile phone usage, are source from the World Bank's world development indicators.
[9] Best linear and unbiased estimator.

learning regularization techniques[10] for controlling the regression coefficients. The power of the regularization algorithms lies in its ability to reduce model variance and out of sample error by selecting relevant variables that drive the outcome variable (inclusive growth). Per the focus of the study, I relax the estimations on *ridge*[11]. The focus, therefore, is to run the lasso family models— *the Standard lasso, the Minimum BIC*[12] *lasso,* and *adaptive lasso,* and compare their selection and prediction powers with the ordinary least squares technique. To this end, the STATA and R software are employed. The latter is employed primarily for data engineering, partitioning, and descriptive purposes while the regularization and prediction are carried out using the former.

3.2.1 Specification of Standard lasso and Minimum BIC lasso models

The Standard lasso penalizes the model coefficients, using a tuning (regularization) parameter (λ) to reduce model variance, thereby enhancing a better fit (Tibshirani 1996). To identify the salient determinants of inclusive growth, the Standard lasso thus exploits the variance-bias tradeoff. The Standard lasso introduces the penalty $\lambda \sum_{j=1}^{\rho} |\beta_j|$, also referred to ℓ_1-norm penalty to obtain $\hat{\beta}_{lasso}$ defined as:

$$\hat{\beta}_{lasso} = min\left\{SSE + \lambda \sum_{j=1}^{\rho} |\beta_j|\right\} \tag{1}$$

by minimizing the objective function:

$$Q_L = \frac{1}{N} \sum_{i=1}^{N} \omega_i \, f(y_{it}, \beta_0 + X_{it}\beta') + \curlywedge \sum_{j=1}^{p} k_j \, |\beta_j| \tag{2}$$

Where *SSE* is the model sum of square errors; y_{it} is inclusive growth in country *i* at time *t*. If $\lambda = 0$ the standard plunges into the ordinary least square (no variable selection is made); while λ → ∞ gives an intercept-only model. For brevity, it is indicated that the specification of the *Minimum BIC* lasso follows a rerun of the Standard lasso but regularization decisions is based on the model with the least Schwarz Bayesian Information Criterion (Schwarz 1978).

3.2.2 Specification of Adaptive lasso model

To address possible inconsistency associated with the Standard lasso, Zou (2006) introduced the adaptive lasso. A key feature differentiating the Adaptive lasso from the Standard lasso is

[10] The recognized regularization algorithms are the ridge, lasso and elasticnet.
[11] This is because both the ridge algorithms put the coefficients close but not equal to zero (James et 2013).
[12] BIC means Schwarz Bayesian Information Criterion.

the '*oracle*' property (z_j), which adds to the penalty $\lambda \sum_{j=1}^{\rho} |\beta_j|$, therefore ensuring greater consistency and parsimonious regularization even when data attributes grow faster than the number of observations. To select the key determinants of inclusive growth, the Adaptive lasso estimator, $\hat{\beta}_{AdaptiveLasso}$, as specified in (3) minimizes the objective function in (4)

$$\hat{\beta}_{AdaptiveLasso} = min\left\{SSE + \lambda \sum_{j=1}^{\rho} z_j |\beta_j|\right\} \tag{3}$$

$$Q_L = \frac{1}{N} \sum_{i=1}^{N} \omega_i \, f(y_{it}, \beta_0 + X_{it}\beta') + \curlywedge \sum_{j=1}^{p} k_j \, |\beta_j| \tag{4}$$

Where y_{it} is inclusive growth in country in *i* at time *t*, X_{it} is a vector of all 97 possible inclusive growth drivers.

3.2.3 Specification of ordinary least square model

The select the significant drivers of inclusive growth via the standard least squares (pooled), we run model (5), which minimizes only the sum of squared errors without any penalization (i.e., tunning parameter)

$$y_{it} = X_{it}\beta' + \varepsilon_{it} \tag{5}$$

where y_{it} is the inclusive growth in country in *i* at time *t*, and X_{it} denotes all 97 possible inclusive growth drivers. It is imperative to note that selection of variables via the least is strictly based on the statistical significance of the regressors.

3.3 Data engineering and partitioning

To ensure a strongly balanced dataset for training machine learning algorithms, missing values as I show in Figure A1 are computed by applying the K-nearest neighbour data engineering technique, particularly for policy and institutional assessment variables[13] (see results in Figures A2 in Appendix C). Finally, I split the dataset into two equal parts based on the stratified method. The choice of the stratification method is two folds. First, because the response variable, inclusive growth, deviates substantially from normality (positively skewed), and (2) to ensure a balanced representation of the response variable in both the training and testing datasets.

[13] These are data on net migration, and country policy and institutional scores for fiscal policy, macroeconomic management, resource equity, social protection, social inclusion, gender equality, pubic administration, trade, debt management, human resource, and financial sector management.

4.0 Presentation and discussion of results

4.1 Exploratory data analysis

I begin the exploratory data analysis by first presenting the distribution of the training and testing sets, followed by the distribution of the outcome variable (inclusive growth), and the summary statistics (see, Table A2 in Appendix D). Information gleaned from Table A2 indicates that the average inclusive growth figure for the SSA over the study period is US$355.42 in the training set compared to US$354.89 in the testing set. This is significantly lower than the GDP per capita value of US$3756.78 (training set) compared to US$4054.92 in the testing sample, clearly signifying a case of non-inclusive growth trajectories in the SSA. We also observe a mean fiscal policy management score of 3.39 in the training set and 3.44 in the testing sample. The data also shows a mean unemployment figure of 7.77 in the training set as compared to 8.37 in the testing set.

4.1.1 Data partitioning results

Figure 1 shows the 50-50 split of the dataset. It is clear that inclusive growth follows similar in both the training and testing samples.

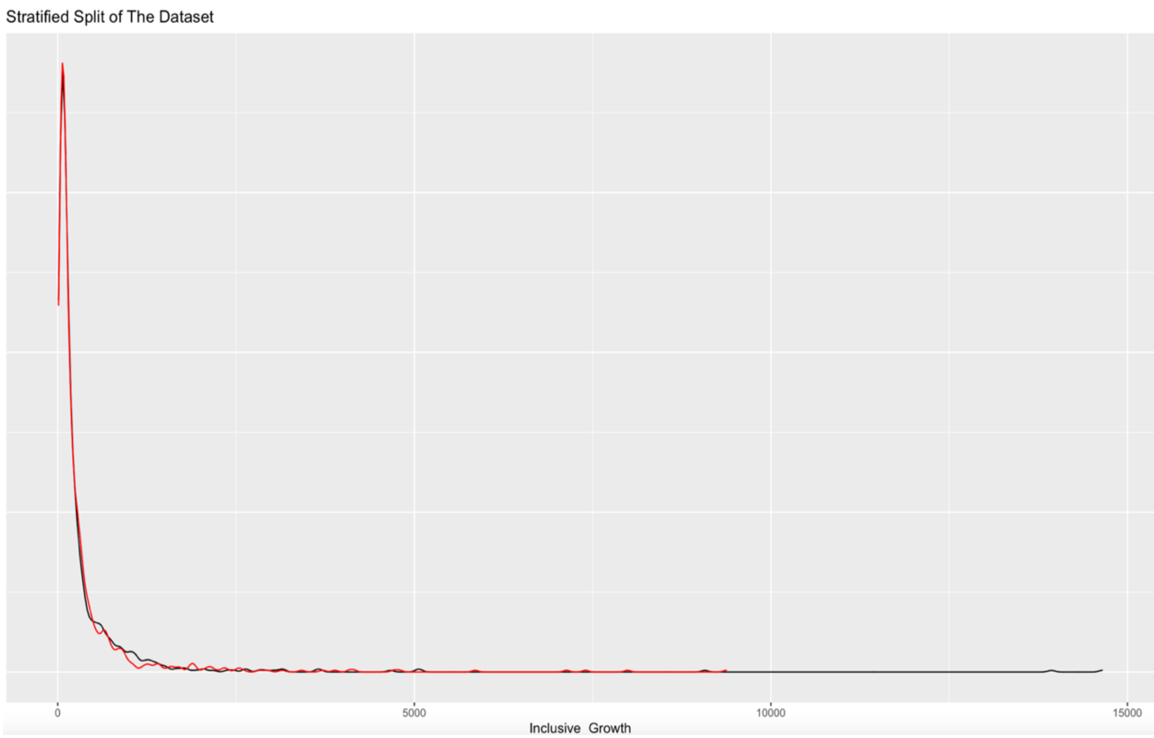


Figure 1: Stratified Split of Inclusive Growth – Training set (Black) vs. Testing set (Red)

4.1.2 The distribution of Inclusive growth

Figure 2 shows the distribution of inclusive growth. It is evident from Figure 2 (left) that inclusive growth is right-skewed. Mindful of the implications of skewed distributions for linear relationships and predictions, inclusive growth is normalized by taking a logarithmic transformation of the series[14]. The result is Figure 2 (right), which shows that inclusive growth is now symmetric (or less heavy-tailed).

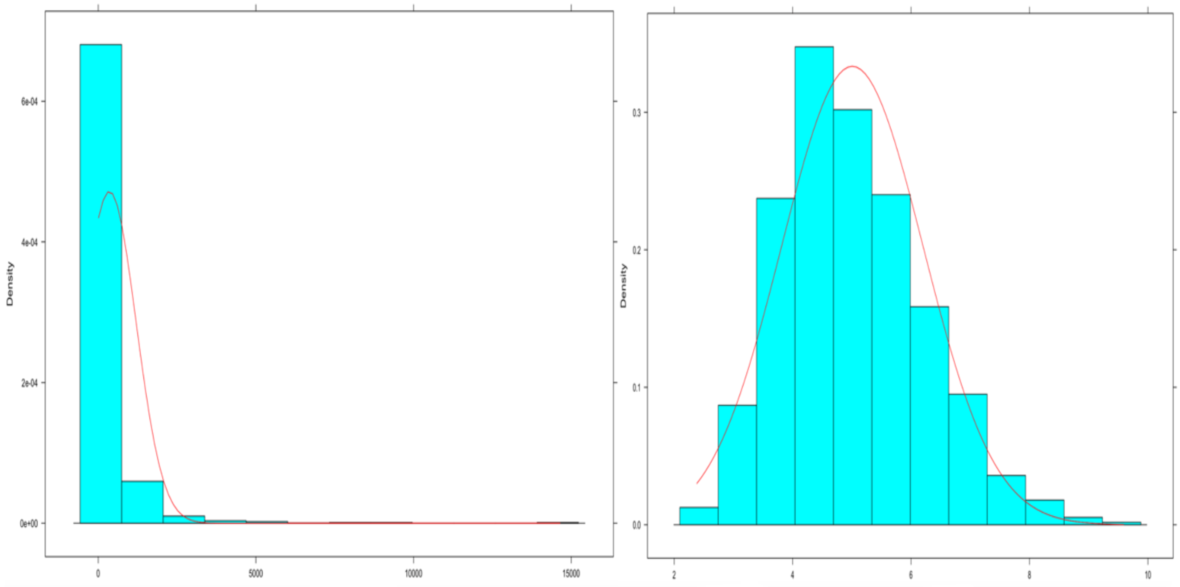


Figure 2: Distribution of Inclusive Growth at level (left), and log-transformed (right)

4.2 Results on drivers of inclusive growth in Sub-Saharan Africa

This section presents results on the first objective of the study— that is variable selection (non-zero coefficients) via the least squares, Standard lasso, Minimum BIC lasso, and Adaptive lasso techniques[15]. For brevity, I present the cross-validation and coefficient path plots for the three models as well as the how the covariates enter and leave the respective models (see, Appendix E). Figures 3 – 5 show that the lasso algorithms select different number of covariates as drivers of inclusive growth.

[14] All variables are standardized to aid appropriate regularization

[15] Note: regularization is done based on 10-fold cross-validation.

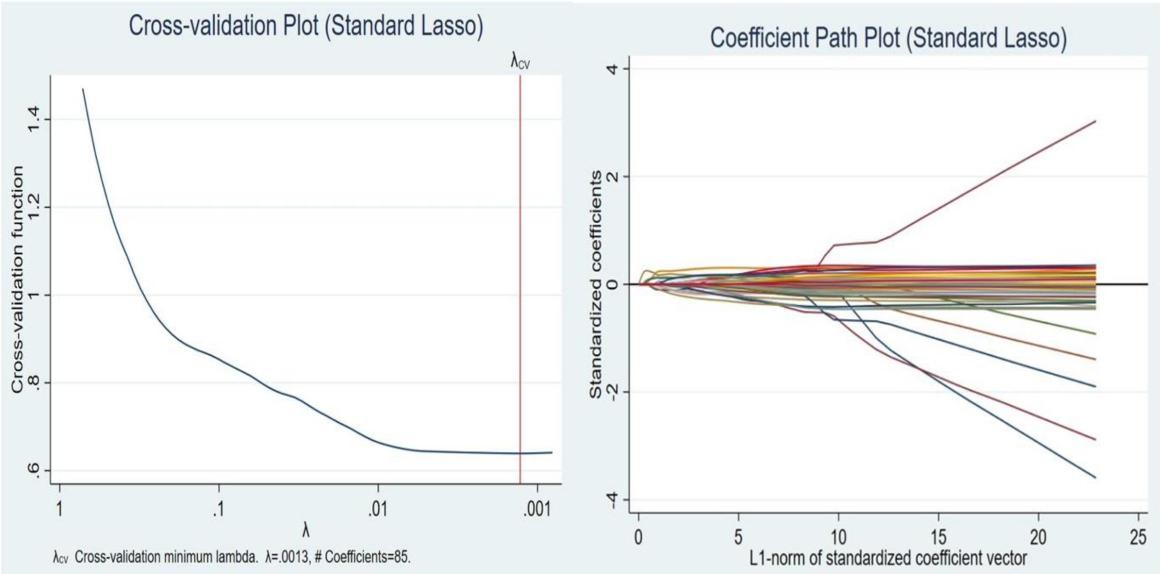


Figure 3: Cross-validation plot (left) and coefficient path plot (right) for Standard lasso

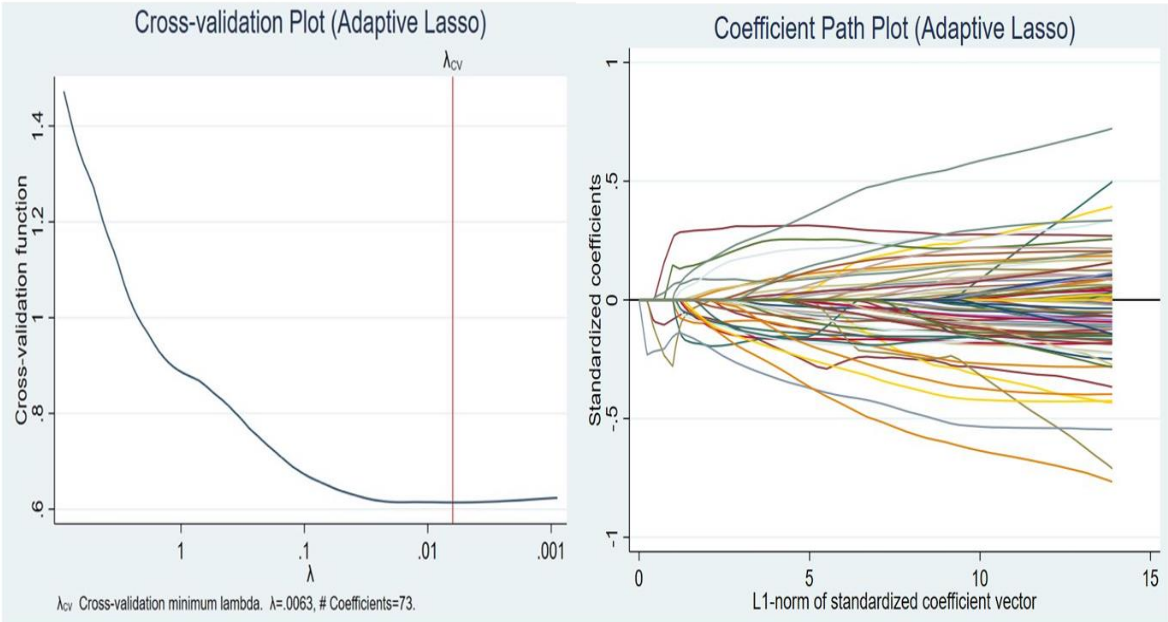


Figure 4: Cross-validation plot (left), and coefficient path plot (right) for Adaptive lasso

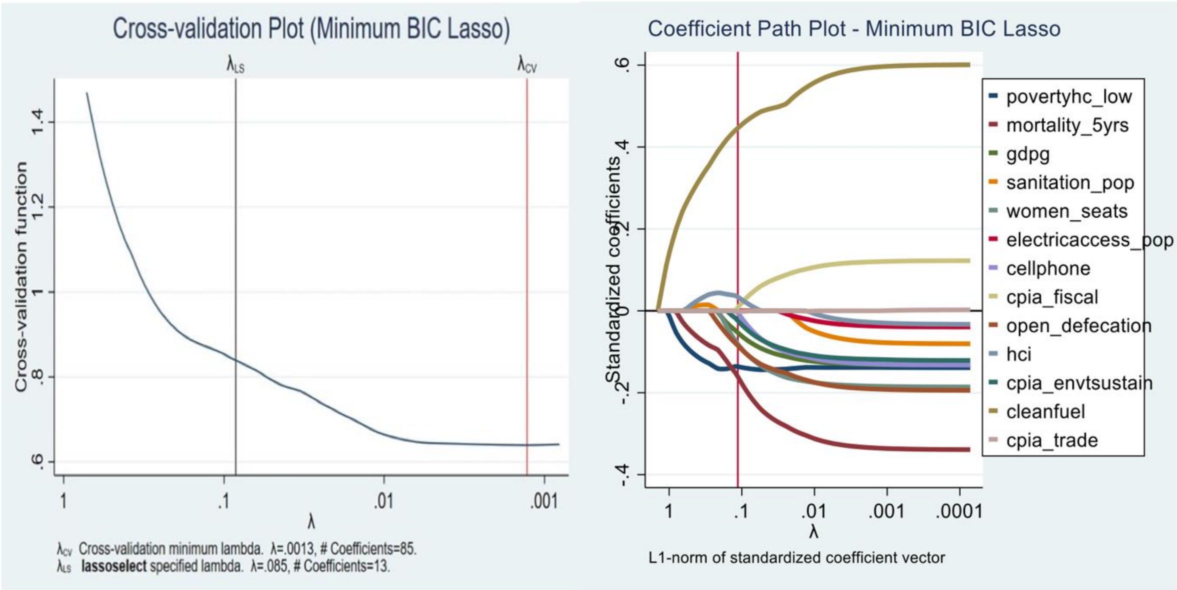


Figure 5: Cross-validation plot (left), and coefficient path plot (right) for Minimum BIC lasso

With a tuning parameter of 0.001, the '*Standard lasso'* model selects a total of 85 variables (see, Figure 3) out of the possible 97. Similarly, the '*adaptive lasso'* model selects 73 variables with a tuning parameter of 0.006 (see, Figure 4). Per the statistical significance of the results in Tables A3 and A4 (supplementary results), it is evident that the least squares technique selects 60 covariates as drivers of inclusive growth. Finally, the study finds a more parsimonious regularization in the '*Minimum BIC'* lasso as it selects only 13 variables out of the possible 97 as key drivers of inclusive growth in SSA (see, Figure 5). It is imperative to point out that these selected drivers of inclusive growth are standardized, implying therefore that the order is important. The results show that, the salient among the 13 selected covariates is the poverty headcount (US$1.90), followed by others such as healthcare (proxied by under-5 mortality), economic growth, clean fuel, sanitation, access to electricity, cellphone, economic growth, sanitation (proxied by personal sanitation facilities), toilet facilities (open defecation), human capital, women voice (proxied by women seats in parliament), electricity access, clean fuel, and country policy and institutional effectiveness for trade (globalisation), fiscal policy, and environmental sustainability.

4.2.1 Pillars of inclusive growth in Sub-Saharan Africa

Based on the '*Minimum BIC*' lasso selection in (Figure 5), I classify the drivers of inclusive growth under three pillars— economic growth, social protection policies, and social inclusion policies (see, Figure 6).

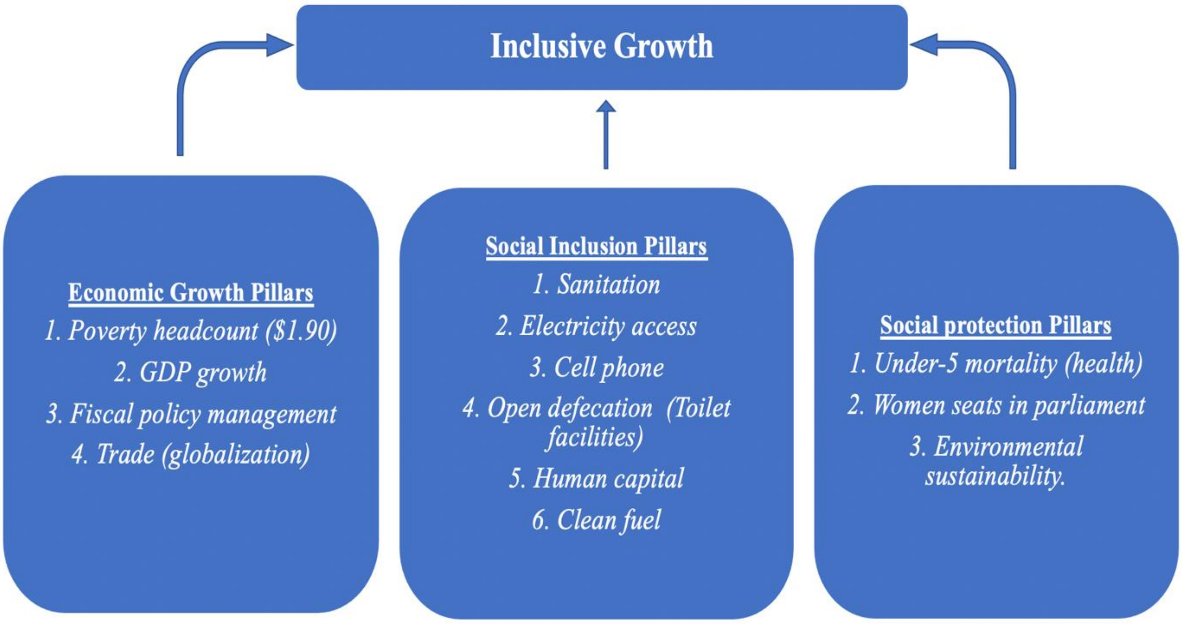


Figure 6: Pillars of inclusive growth based on Minimum BIC lasso selection

For the economic growth pillar, the results show that sustained economic growth, trade, macroeconomic management and poverty reduction are crucial. Also revealing is the selection of clean fuel, sanitation, environmental sustainability, and cellphones under the social equity dimensions (social protection and inclusion), which are in line with the SSA's green growth agenda.

4.3 Predicting inclusive growth in Sub-Saharan Africa

The last objective of the paper is the determination of the model best for predicting inclusive growth in SSA. Which model is the best? The answer is in the goodness-of-fit statistics presented in Table 1.

Table 1: Evaluation of OLS and Lasso models

Prediction Techniques	*Sample*	*RMSE*	*R-Squared*	*Obs*
OLS	*Training set*	0.552	0.999	860
	Testing set	0.692	0.999	860
Standard lasso				
	Training set	0.555	0.624	860
	Testing set	0.680	0.508	860
Minimum BIC lasso				
	Training set	0.560	0.620	860
	Testing set	**0.678**	**0.509**	**860**
Adaptive lasso				
	Training set	0.711	0.518	860
	Testing set	0.763	0.448	860

RMSE: Root Mean Squared Error; Observation

Source: Author's construct, 2021

Table 1 shows that, of all the models, the *'Minimum BIC'* lasso is appropriate for predicting inclusive growth in SSA. Its appropriateness lies in its relative lower out of sample error (0.67) and realistic R-squared of 50.9 per cent. The model for predicting inclusive growth is thus:

$$Inclusivegrowth_{it} = \beta_0 + \beta_1 poverty_{it} + \beta_2 gdpgrowth_{it} + \beta_3 fiscalpolicy_{it} + \beta_4 globalisation_{it} + \beta_5 sanitation_{it} + \beta_6 cellphone_{it} + \beta_7 toiletfacilities_{it} + \beta_8 humancapital_{it} + \beta_9 electricity_{it} + \beta_{10} mortality_{it} + \beta_{11} womenvoice_{it} + \beta_{12} environmentalsustainability_{it} + \beta_{13} cleanfuel_{it} + \varepsilon_{it}$$

where $\varepsilon_{it} = \epsilon_i + \vartheta_t + \mu_{it}$, where ϵ_i is unobserved country-specific fixed effects; ϑ_t is the time effects, and μ_{it} is the idiosyncratic error term.

5.0 Conclusion and policy implications

The study contributes to the policy discourse on how sustainable and pro-poor growth can be achieved in SSA[16]. The contribution I make is non-subjective as recent advances in machine learning are employed to achieve two objectives— first for identifying covariates best for driving inclusive growth in SSA, and second, identifying a model best for predicting inclusive growth in SSA. To this end, I run` three machine learning models— the Standard lasso, Minimum BIC lasso, and the adaptive lasso, in addition to the standard least square technique based on a dataset comprising 97 covariates. On the first objective, the regularization results show that only 13 variables matter for driving inclusive growth in SSA. On the second objective, the results show that the Minimum BIC lasso model is best for predicting inclusive growth. The study shows that, for policymakers to foster inclusive growth in SSA, policy formulations are to target three thematic areas— *the economic growth pillar; the social inclusion pillar;* and *the social protection pillar*. The aforementioned pillars are in themselves strategies for building shared opportunity in SSA through the creation of sustainable employment opportunities, access to environmentally friendly basic amenities, and the strengthening of institutions to provide a level playing field for all. In specifics, the

[16] The SSA countries (43) in the study are: Angola, Benin, Botswana, Burkina Faso, Burundi, Cabo Verde, Cameroon, Central African Republic, Chad, Comoros, Congo, DR., Congo, Cote d'Ivoire, Eswatini, Ethiopia, Gabon, The Gambia, Guinea, Ghana, Guinea-Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Nigeria, Rwanda, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Sudan, South Africa, Tanzania, Togo, Uganda, Zambia.

results show that, to spur inclusive growth from the growth side, policymakers should strive to reduce poverty (poverty headcount US$1.90), ensure prudent fiscal policy management, sustained economic growth (GDP growth), and positioning the region to take advantage of globalisation. Policymakers can foster inclusive growth through the latter by boosting the productive capacity to take advantage of the fertile grounds provided by the Africa Continental Free Trade Area (AfCFTA). The essence of prudent fiscal management in this sense is a conducive environment for private sector to thrive plausibly through infrastructural development while supporting the vulnerable groups through efficient redistribution. The results on social inclusion indicates that enhancing the quality of sanitation, electricity access, ICT diffusion, human capital and clean fuel is key to fostering inclusive growth in SSA. Reliable energy is needed to reduce the cost of production especially small and medium scale enterprises. Also, the power of cellphones in fostering shared growth is seen in its adoption and applicability in various facets of life— for accessing information and opportunities (e.g., wider markets), educational services through e-learning, reduction in transaction costs (e.g., through mobile money/internet banking), and knowledge diffusion. Also germane are the social protection indicators which call for improvement in healthcare system especially for reducing under-5 mortality, while deepening women's voice in decision making (proxied by women seats in parliament), and environmental sustainability. The later calls for policies addressing climate change and its food security concerns by addressing deforestation, illegal mining, and poaching.

The author declares that no funding was received for this research

References

Acemoglu & Robinson (2012) “Why Nations Fail: The Origins of Power, Prosperity and Poverty”. *Profile Books*, London.

Ali, I. & Son, H. (2007). “Measuring Inclusive Growth,” *Asian Development Review*, 24(1), 11–31.

Anand, R., Mishra, S. & Peiris, S. (2013). Inclusive growth: measurement and determinants. *IMF Working Papers No. 13/135.*

ADB. (2013). Framework of inclusive growth indicators: Key indicators for Asia and the Pacific. Mandaluyong City, Philippines: *Asian Development Bank.*

AWID (Association for Women’s Right in Development) (2011) Microcredit Pitfalls: The Experience of Dalit Women In India. http://bit.ly/1lUosvk

Berg A., & Ostry J. D. (2011). Inequality and Unsustainable Growth: Two sides of the same coin? *IMF Staff Discussion Note,* Washington DC, International Monetary Fund,

Brown, C. S., Ravallion, M., & Van De Walle, D. (2020). Can the World’s Poor Protect Themselves from the New Coronavirus? (No. w27200). *National Bureau of Economic Research*.

Calderón, C. & Servén, L. (2014). Infrastructure, Growth and Inequality (Policy Research Working Paper). World Bank.

Commission on Growth and Development (2008) “The Growth Report: Strategies for Sustained Growth and Inclusive Development”. *Washington DC,* World Bank.

Dreher, A. (2006). “Does Globalization Affect Growth? Empirical Evidence from a new index.” *Applied Economics,* 38: 1091-1110.

Dutt, P., & Tsetlin, I. (2016). Income distribution and economic development: Insights from machine learning. *Economics & Politics*.

ECLAC (2011) “Poverty, Inequality and Perceptions of Work in Latin America”. CEPAL, Santiago de Chile

Estache, A., Ianchovichina, E., Bacon, R., & Salamon, I. (2013). Infrastructure and employment creation in the Middle East and North Africa. The World Bank.

Gajigo, O. & Lukoma, A. (2011). Infrastructure and agricultural productivity in Africa. *Market Brief.* African Development Bank.

Greenwald, B., & Stiglitz, J. E. (2013). Industrial policies, the creation of a learning society, and economic development. In The Industrial Policy Revolution (43-71). *Palgrave Macmillan*, London.

Gyamfi, N. M., Bokpin, G. A., Aboagye, A. Q., & Ackah, C. G. (2019). Financial development, institutional quality and inclusive growth in Africa. *Global Business Review*, Gygli, S., Haelg, F., Potrafke, N., & Sturm, J. E. (2019). The KOF Globalisation Index – Revisited. *Review of International Organizations*, *14*(3), 543–574.

Ianchovichina, E., & Lundstrom S. G. (2012). What Is Inclusive Growth? In Commodity Prices and Inclusive Growth in Low-Income Countries, ed. by Rabah Arezki, Catherine Pat- tillo, Marc Quintyn, and Min Zhu. *International Monetary Fund*.

ILO (International Labour Organization) (2020a). World Economic and Social Outlook Trends – 2020. *International Labour Office – Geneva.*

IMF (International Monetary Fund) (2020a). World Economic Outlook: A Long and Difficult Ascent. October. *Washington, DC,* International Monetary Fund.

———. (2011). "Regional Economic Outlook: Asia and Pacific." October, Washington, DC. International Monetary Fund.

———. (2007). "Globalization and Inequality." In World Economic Outlook, chapter 4. Washington, DC. International Monetary Fund.

International Monetary Fund and World Bank. (2020). Enhancing Access to Opportunities. International Monetary Fund and World Bank, *Washington, DC*: World Bank.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112, p. 18). New York: springer.

Jung, J.-K., Patnam, M., & Ter-Martirosyan, A. (2018). *An algorithmic crystal ball: Forecasts-based on machine learning.* IMF Working Papers. Washington, D.C.: International Monetary Fund.

Lahoti, R., Jayadev, A., & Reddy, S. (2016). The global consumption and income project (GCIP): An overview. *Journal of Globalization and Development*, *7*(1), 61-108.

Lustig, N., Lopez-Calva, L. F., & Ortiz-Juarez, E. (2012). *Declining inequality in Latin America in the 2000s: the cases of Argentina, Brazil, and Mexico*. The World Bank.

Mutiiria, O. M., Ju, Q., & Dumor, K. (2020). Infrastructure and inclusive growth in sub-Saharan Africa: An empirical analysis. *Progress in Development Studies*, *20*(3), 187-207.

OECD (Organization for Economic Co-operation and Development). (2020). The Impact of The Coronavirus (COVID-19) Crisis on Development Finance, June, 2020.

Paramasivan, S. Mani, K. & Utpal, C. (2014). A Theoretical Model for Inclusive Economic Growth in Indian Context. *International Journal of Humanities and Social Science* 4(13), 228-234.

Ravallion, M., & Chen, S. (2019). Global poverty measurement when relative income matters. *Journal of public economics, 177*, 104046.

Richardson, A., van Florenstein Mulder, T., & Vehbi, T. (2021). Nowcasting GDP using machine-learning algorithms: A real-time assessment. *International Journal of Forecasting*, *37*(2), 941-948.

———. (2004). Measuring Pro-Poor Growth. *Economics Letters*, 78: 93−99.

Schneider, U., & Wagner, M. (2012). Catching growth determinants with the adaptive lasso. *German Economic Review*, *13*(1), 71-85.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of statistics*, *6*(2), 461-464.

Tella, S. A., & Alimi, O. Y. (2016). Determinants of inclusive growth in Africa: Role of health and demographic changes. *African Journal of Economic Review*, 4(2), 138-146.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267-288.

Tkacz, G. (2001). Neural network forecasting of Canadian GDP growth. *International Journal of Forecasting*, *17*(1), 57-69.

UNECA. 2019. "Fiscal Policy for Sustainable Development in Africa." https://repository .uneca.org/handle/10855/41804.

Van Niekerk, A. J. (2020). Towards inclusive growth in Africa. *Development Southern Africa*, *37*(3), 519-533.

World Bank. (2021). World Development Indicators. February 17 Editiom. *Washington, DC:* World Bank.

———. (2020a). Poverty and Shared Prosperity 2020: Reversals of Fortunes. October. *Washington, DC*: World Bank.

———. (2020b). Global Economic Prospects, June 2020. *Washington, DC*: World Bank.

———. (2019). Global Economic Prospects. Darkening Skies. January. *Washington, DC*: World Bank.

———. (2013). "Inclusion Matters: The Foundation for Shared Prosperity". Washington DC, World Bank.

———. (2009) "What is Inclusive Growth?" Washington DC, World Bank.

Zhuang, J., & Ali, I. (2010). Poverty, inequality, and inclusive growth in Asia. *Poverty, Inequality, and Inclusive Growth: Measurement, Policy Issues, and Country Studies*, 1-32.

Zou, H. (2006). The Adaptive lasso and its oracle properties. *Journal Of The American Statistical Association*, 101(476), 1418-1429.

Appendix A

Measurement of Inclusive Growth by Anand et al. (2013)

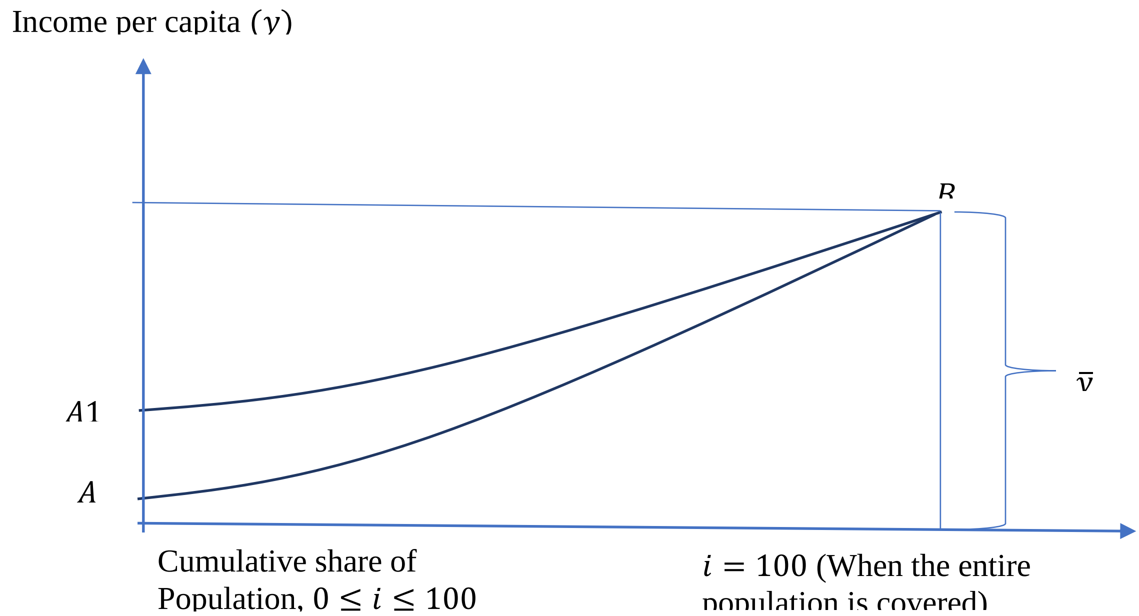
This writeup is reproduced from the original article with permission of the IMF as per the IMF copyright and usage effective January 02, 2020.

To integrate equity and growth in a unified measure, Anand, Mishra and Peiris (2013) proposed a measure of inclusive growth based on a utilitarian social welfare function drawn from consumer choice literature, where inclusive growth depends on two factors: (i) income growth; and (ii) income distribution. Similar to the consumer theory where the indifference curves represent the changes over time in aggregate demand, Anand, Mishra and Peiris (2013) decomposed the income and substitution effect into growth and distributional components. The underlying social welfare function must satisfy two properties to capture these features: (i) it is increasing in its argument (to capture growth dimension) and (ii) it satisfies the transfer property – any transfer of income from a poor person to a richer person reduces the value of the function (to capture distributional dimension).

A measure of inclusiveness is based on the concept of a concentration curve. Following Ali and Son (2007), Anand, Mishra and Peiris (2013) defined a generalized concentration curve, which they called social mobility curve, S^c, such that:

$$S^c \approx \left(y_1, \frac{y_1 + y_2}{2}, \dots \dots \dots, \frac{y_1 + y_2 + \cdots + y_n}{n}\right)$$

Where n is the number of persons in the population with incomes $y_1, y_2, \dots \dots, y_n$, where $y_1$is the poorest person and y_n is the richest person. This generalized concentration curve is basically a cumulative distribution of a social mobility vector $S \approx (y_1, y_2, \dots \dots \dots, y_n)$ with an underlying function $W = W(y_1, y_2, \dots \dots \dots, y_n)$ satisfying the two properties mentioned above to capture growth and distribution dimensions. Since S^c satisfies the transfer property, a superior income distribution will always have a higher generalized concentration curve. Similarly, since it is increasing in its argument, higher-income will also have a higher generalized concentration curve. As in Ali and Son (2007), the generalized concentration curves can be presented in continuous time to be more amendable to econometric analysis. The population is arranged in the ascending order of their income. Let $\bar{y_i}$ is the average income of the bottom i per cent of the population, where i varies from 0 to 100 and y_i is the mean income. Anand, Mishra and Peiris (2013) plotted $\bar{y_i}$ for different values of i (curve AB in Appendix A below). Curve AB represents a social mobility curve discussed above. Since a higher curve implies greater social mobility, growth is inclusive if the social mobility curve moves upward at all points. However, there may be degrees of inclusive growth depending on: (i) how much the curve moves up (growth); and (ii) how the distribution of income changes (equity). This feature of the social mobility curve is the basis of our integrated measure of inclusive growth. Thus, if two generalized concentration curves do not intersect, they could be ranked on social mobility (i.e. inclusiveness of growth). To illustrate the point made above, Appendix A depicts two social mobility curves with the same average income ($\bar{y}$) but different degrees of inclusiveness (i.e. different income distribution). Social mobility curve (A1B) is more inclusive than the social mobility curve AB, as the average income of the bottom segment of the society is higher.



Source: Anand, Mishra and Peiris (2013)

To capture the magnitude of the change in income distribution, Anand, Mishra and Peiris (2013) used a simple form of the social mobility function by calculating an index (or social mobility index) from the area under the social mobility curve:

$$\bar{y}^* = \int_0^{100} \bar{y}_i \, di$$

The greater the $\bar{y}^*$, the greater is the income. If the income of everyone in the population is the same (i.e. if income distribution is completely equitable) then $\bar{y}^*$ will be equal to $\bar{y}$. If $\bar{y}^*$ is lower than $\bar{y}$, it implies that the distribution of income is inequitable. So, the deviation of $\bar{y}^*$ from $\bar{y}$ is an indication of inequality in income distribution. Ali and Son (2007 use this feature of $\bar{y}^*$ and propose an income equity index (IEI) as:

$$\omega = \frac{\bar{y}^*}{\bar{y}}$$

For a completely equitable society, $\omega = 1$. Thus, a higher value of ω (closer to one) represents higher income equality. Rearranging,

$$\bar{y}^* = \omega * \bar{y}$$

Inclusive growth requires increasing$\bar{y}^*$, which could be achieved by: (i) increasing $\bar{y}$, that is increasing average income through growth; (ii) increasing the equity index of income, ω, through increasing equity; or (iii) a combination of (i) and (ii). Differentiating the above equation:

$$d\bar{y}^* = \omega * d\bar{y} + d\omega * \bar{y}$$

Where $d\bar{y}^*$ is the change in the degree of inclusive growth. Growth is more inclusive if $d\bar{y}^* > 0$. It also allows us to decompose inclusive growth into income growth and change in equity. The first term is the contribution of an increase in average income (keeping income distribution constant) while the second term is the contribution of changes in the income distribution (keeping the average income unchanged). Inclusive growth depends on the sign and the magnitude of the two terms.

Appendix B

Table A.1: Variable description and sources

Variable	Description	Source
wagessalary	Wages and salaried workers (percentage of total employment)	WDI
Vulnerable_employment	Vulnerable/precarious employment	WDI
unempl	Annual unemployment rate	WDI
sec_teachers	Trained teachers in secondary education (% of total teachers)	WDI
debt_service	Overall national debt (%GDP)	WDI
registry_time	Time required to register property is the number of calendar days needed for businesses to secure rights to property.	WDI
business_time	Time required to start a business is the number of calendar days needed to complete the procedures to legally operate a business.	WDI
tertiary	Academic staff (% female) is the share of female academic staff in tertiary education.	WDI
taxrev	Tax revenue (%GDP)	WDI
Tariff	Trade policy (tariff weighted mean)	WDI
soc_contri	Contributions by employees, employers, and self-employed individuals, and other contributions whose source cannot be determined.	WDI
service_VA	Service sector, value added	WDI
self_employ	Self-employment, total (% total employment)	WDI
self_emplFE	Self-employment, total (% female)	WDI
sch_enrolTER	Gross enrolment ratio is the ratio of total enrolment in tertiary institutions, regardless of age, to the population of the age group	WDI
infrast_qual	The Quality of Port Infrastructure measures business executives' perception of their country's port facilities.	WDI
electricaccess_pop	Electricity access (overall population)	WDI
rd	Expenditure on research and development	WDI
labforce_MAFE	Labour force participation rate	WDI
pupiltea_ratio	Pupil teacher ratio	WDI
women_seats	Women in parliaments are the percentage of parliamentary seats in a single or lower chamber held by women.	WDI
progressto_sec	Number of junior high student progressing to high school	WDI
HIV_preva	Prevalence of HIV, total (% of population ages 15-49)	WDI
prenatal	Pregnant women receiving antennal care	WDI
poverty_hc	International poverty headcount	PED
povertyhc_mid	International poverty headcount (US$3.20)	PED
povertyhc_low	International poverty headcount (US$1.90)	PED
povgap_mid	Lower-middle income poverty gap (Poverty gap US$3.20)	PED
povgap_low	Poverty intensity (poverty gap US$1.90)	PED
urbanization	Annual population growth rate (urban)	WDI
popgrof	Annual population growth rate (overall)	WDI
remit	Remittance inflows to GDP (%)	WDI
sanitation	People using at least basic sanitation services	WDI
opendefeca_pop	People practicing open defecation	WDI
unfpa_aid	Net official development assistance from UNFPA	WDI
unicef_aid	Net official development assistance from UNICEF	WDI
undp_aid	Net official development assistance from UNDP	WDI
noda	Net official development assistance	WDI
netmigration	Net migration (immigrants less emigrants)	WDI
mortality_5yrs	Prevalence of infant (under-5) mortality	WDI
cellphone	Active mobile cellular phones per 100 people	WDI
manuf_VA	Manufacturing sector, value added	WDI
logisticquality_TT	Logistic quality score (road transportation)	WDI
logisticquality_overal	Overall logistic quality score	WDI
logisticquality_ship	Logistic quality score (shipping)	WDI
electricity	Electricity access (rural population)	WDI

industry_VA	Industrial sector, value added	WDI
povgapmid_increase		
povgaplow_increase		
hci	Human Capital Index (HCI) (scale 0-1)	WDI
gfcf	Gross fixed capital formation	WDI
gov_educ	Government expenditure on education	WDI
Government_Expenditure	Overall government expenditure (%GDP)	WDI
gpc_GDP_Per_Capita_Growth	GDP per capita growth	WDI
GDP_Per_Capita	GDP per capita (US$' 2017 PPP)	WDI
gdpg	Annual GDP growth rate	WDI
Foreign_Direct_Investment	Foreign direct investment, net inflows (%GDP)	WDI
telefon	Fixed telephone subscription per a million population	WDI
emp_ind	Number of people employed (industrial sector)	WDI
emp_agric	Number of people employed (agricultural sector)	WDI
health_exp	Government health expenditure (%GDP)	WDI
Financial_Deepening	Financial institutions credit to private sector	WDI
health_expcurrent	Level of current health expenditure expressed as a percentage of GDP.	WDI
cpia_transparency	Transparency, corruption, and accountability rating (1=low to 6=high)	CPIA
cpia_trade	Trade rating (1=low to 6=high)	CPIA
cpia_socprotection	Institutions for Social protection rating (1=low to 6=high)	CPIA
cpia_publicadmi	Assesses the extent to which civilian central government staff is structured to design and implement government policy and deliver services effectively.	CPIA
cpia_publicmgt	Public sector management and institutions cluster average (1=low to 6=high)	CPIA
cpia_envtsustain	Assesses the extent to which environmental policies foster the protection and sustainable use of natural resources and the management of pollution.	CPIA
cpia_socinclusion	Institutions for Social inclusion rating (1=low to 6=high)	CPIA
cpia_macro	Macroeconomic management rating (1=low to 6=high)	CPIA
cpia_gender	Gender equality rating (1=low to 6=high)	CPIA
cpia_fiscal	Fiscal policy assesses the short- and medium-term sustainability of fiscal policy (taking into account monetary and exchange rate policy and the sustainability of the public debt) and its impact on growth.	CPIA
cpia_finsector	Financial sector management rating (1=low to 6=high)	CPIA
cpia_resourceeqity	The extent to which the pattern of public expenditures and revenue collection affects the poor and is consistent with national poverty reduction priorities.	CPIA
cpia_humanresouce	Effectiveness of national policies and public and private sector service delivery that affect the access to and quality of health and education services, including prevention and treatment of HIV/AIDS, tuberculosis, and malaria.	CPIA
Social_Protection_Score	Percentage of population participating in social insurance, social safety net, and unemployment benefits and active labour market programs.	WDI
socialinsurance	Percentage of population participating in programs that provide old age contributory pensions and social security and health insurance benefits	WDI
importcost	Cost to import, documentary compliance (US$)	DBP
exportcost	Cost to export, documentary compliance (US$)	DBP
Inflation	End-of-period inflation	WDI
banks	Retail locations of resident commercial banks and other resident banks that function as commercial banks that provide financial services to customers	WDI
atm	Computerized telecommunications devices that provide clients of a financial institution with access to financial transactions in a public place.	WDI
agric_VA	Agricultural sector (value added)	WDI
bankacc	Individuals having an account at a bank or another type of financial institution or using a mobile money service in the past 12 months (female, % age 15+).	WDI
cleanfuel	Access to clean fuels and technologies for cooking is the proportion of total population primarily using clean cooking fuels and technologies for cooking.	WDI
salary	Payments in cash, as well as in kind to employees in return for services rendered, and government contributions to social insurance schemes	WDI
salt	Percentage of households which have salt they used for cooking that tested positive (>0ppm) for presence of iodine.	WDI

natresourcerent	Natural resource rent %GDP)	WDI
Inclusive_Growth	Inclusive growth value calculated as presented in Appendix A	Generated
kofgi	Kof. eoverall globalisation index (de jure)	KOF
Economic_Globalisation_Index	Kof. Economic globalisation index (de jure)	KOF
koffin	Kof. financial globalisation index (de jure)	KOF
kofso	Kof. social globalisation index (de jure)	KOF
palma	The Palma ratio of income inequality	GCIP
theil	Theil index of income inequality	GCIP
gini	Gini index inequality indicators	GCIP

Data sources: World Bank Group, CPIA (Country Policy and Institutional Assessment) database: http://www.worldbank.org/ida); WDI (World Development Indicators: http://www.worldbank.org/lpi); World Bank, WDB (Doing Business Project): http://www.doingbusiness.org/).; World Bank (Poverty and Equity Database: https://databank.worldbank.org/reports.aspx?source=poverty-and-equity-database ; KOF Index: https://kof.ethz.ch/en/forecasts-and-indicators/indicators/kof-globalisation-index.html ; GCIP: http://gcip.info/graphs/download

Source: Author’s Construct, 2021

Appendix C

Figure A1: Overview of dataset before data engineering

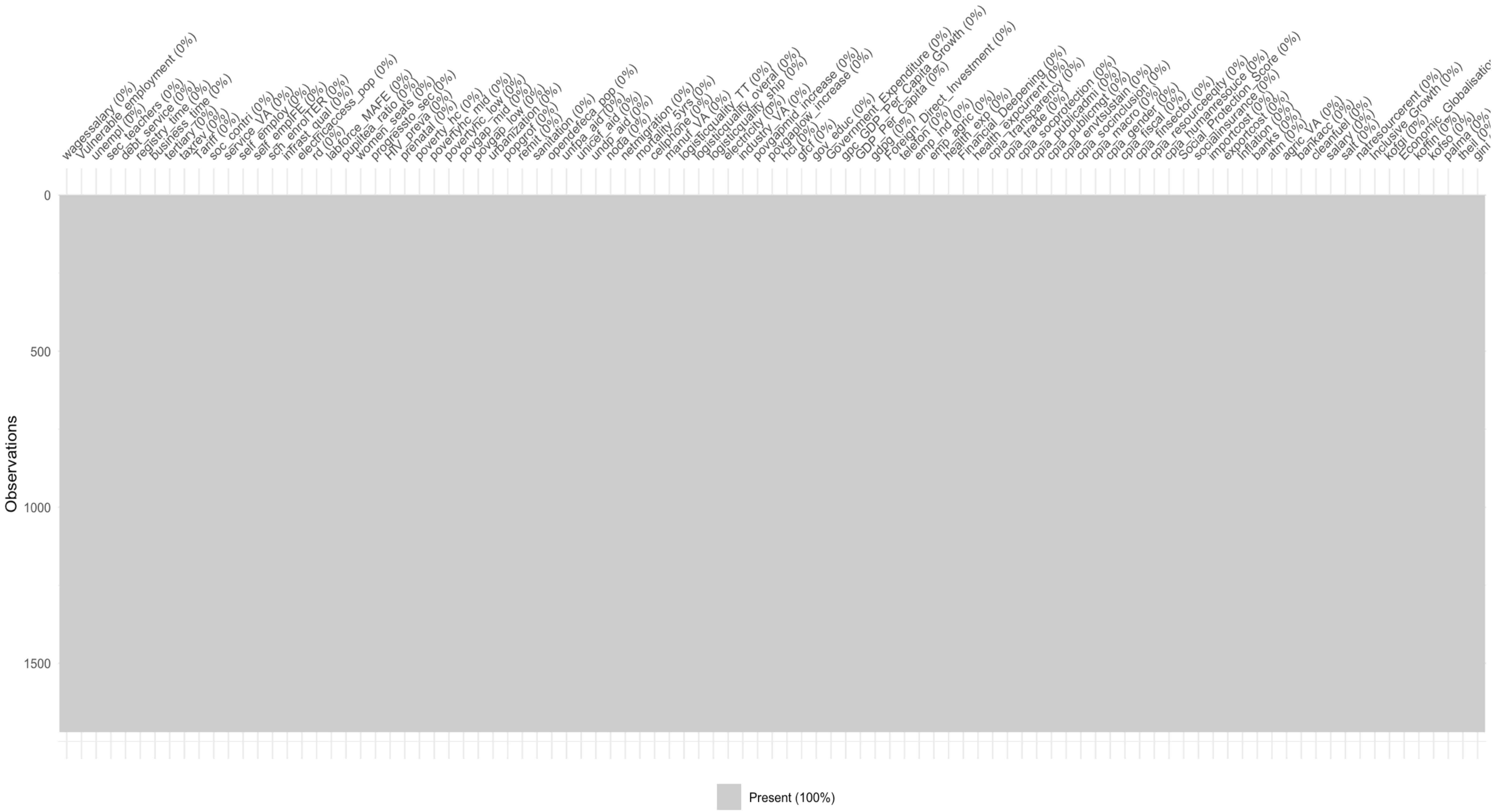


Figure A2: Data engineering plot of variables in the dataset

APPENDIX D

Table A2: Descriptive statistics

Variables	Obs	Mean (TR)	Std. Dev. (TR)	Min (TR)	Max (TR)	Mean (TS)	Std. Dev. (TS)	Min (TS)	Max (TS)
wagessalary	***860(860)***	27.004	22.272	5.049	85.135	28.909	23.663	5.106	85.871
Vulnerable_employment	***860(860)***	70.913	22.908	9.529	94.75	68.983	24.341	8.826	94.759
unempl	***860(860)***	7.778	7.494	.3	37.932	8.376	7.987	.3	37.976
sec_teachers	***860(860)***	70.8	24.701	12.903	100	70.998	24.483	12.038	100
debt_service	***860(860)***	3.651	5.001	.053	73.283	3.44	3.586	.027	30.36
registry_time	***860(860)***	92.179	83.85	7	389	93.762	86.916	7	389
business_time	***860(860)***	53.764	47.245	2.5	259.5	53.447	46.033	4	260.5
tertiary	***860(860)***	19.091	12.527	1.225	54.314	19.35	13.039	1.225	57.143
taxrev	***860(860)***	15.075	7.338	4.099	38.532	15.229	7.285	4.204	39.258
Tariff	***860(860)***	12.346	5.821	.85	91.27	12.365	5.349	.84	32.6
soc_contri	***860(860)***	.645	1.446	0	6.391	.719	1.458	0	6.391
service_VA	***860(860)***	4.819	6.101	-57.122	37.156	4.948	7.057	-33.233	97.466
self_employ	***860(860)***	72.418	22.143	14.865	94.951	70.534	23.428	14.129	94.894
self_emplFE	***860(860)***	79.159	22.436	12.337	99.23	77.161	24.062	11.816	99.225
sch_enrolTER	***860(860)***	4.703	5.292	.075	38.904	4.931	5.772	.09	40.596
infrast_qual	***860(860)***	3.794	5.573	.043	44.43	4.139	6.289	.043	47.457
electricaccess_pop	***860(860)***	64.856	22.807	13.654	99.479	65.817	23.298	10.805	98.662
rd	***860(860)***	.257	.191	.005	.898	.272	.192	.007	.888
labforce_MAFE	***860(860)***	79.577	17.035	30.61	108	78.111	17.219	30.61	107.883
pupiltea_ratio	***860(860)***	25.188	9.466	8.442	80.052	24.92	8.932	5.314	80.052
women_seats	***860(860)***	12.847	10.117	0	63.75	13.508	10.6	0	63.75
progressto_sec	***860(860)***	70.478	21.454	7.45	100	71.295	20.975	8.678	100
HIV_preva	***860(860)***	2.882	4.095	.1	24.2	2.912	4.317	.1	24.2
prenatal	***860(860)***	78.705	17.611	23.4	99.4	78.064	19.025	23.4	99.4
poverty_hc	***860(860)***	49.384	13.654	7.9	73.2	47.86	14.328	7.9	73.2
povertyhc_mid	***860(860)***	69.8	23.367	3.1	98.5	67.587	25.041	2.2	98.5
povertyhc_low	***860(860)***	49.823	24.368	.4	94.3	48.307	25.738	-6.3	94.3
povgap_mid	***860(860)***	38.505	18.585	.7	86.7	37.554	19.874	-4.9	86.7
povgap_low	***860(860)***	23.072	16.218	.1	86.7	22.811	17.421	-13.3	86.7
urbanization	***860(860)***	39.767	14.896	10.838	100	39.365	13.667	10.954	90.707
popgrof	***860(860)***	2.562	1.062	-6.766	8.118	2.545	.956	-5.539	6.989
remit	***860(860)***	4.275	17.579	0	235.924	4.473	17.521	0	232.217
sanitation	***860(860)***	28.268	20.998	3.404	100	30.142	21.957	3.404	100
opendefeca_pop	***860(860)***	10.41	12.082	0	64.501	9.924	11.94	0	64.501
unfpa_aid	***860(860)***	1610000	1610000	10000	9270000	1480000	1510000	-10000	9040000
unicef_aid	***860(860)***	6290000	9120000	-980000	66857658	5890000	9050000	-5490000	68126160
undp_aid	***860(860)***	5020000	4280000	-1610000	31520000	4780000	4270000	10000	36919998
noda	***860(860)***	10.983	10.896	-.251	94.946	10.062	9.82	0	77.868
netmigration	***860(860)***	-19600	256000	-1374270	1287106	-19400	300000	-1374270	1457943
mortality_5yrs	***860(860)***	125.892	63.626	13.7	336.2	123.854	65.85	13.7	337.4
cellphone	***860(860)***	23.975	38.478	0	198.152	24.108	38.971	0	173.497
manuf_VA	***860(860)***	2.974	16.054	-43.84	375.158	2.597	10.165	-37.933	97.709
logisticquality_TT	***860(860)***	2.174	.337	1.4	3.79	2.174	.363	1.27	3.776
logisticquality_overal	***860(860)***	2.397	.309	1.61	3.775	2.397	.329	1.77	3.775
logisticquality_ship	***860(860)***	2.828	.443	1.67	4.03	2.831	.436	2	4.018
electricity	***860(860)***	58.381	21.67	0	95.868	60.301	21.483	0	95.868
industry_VA	***860(860)***	23.049	12.188	.96	72.123	24.029	12.229	1.305	72.717
povgapmid_increase	***860(860)***	3.198	2.979	.005	27.729	2.981	2.238	.166	27.729
povgaplow_increase	***860(860)***	1.635	1.65	0	15.555	1.517	1.22	.004	15.555
hci	***860(860)***	.394	.068	.293	.678	.397	.072	.293	.678

gfcf	***860(860)***	21.542	10.515	0	93.547	21.332	10.363	-2.424	79.158
gov_educ	***860(860)***	15.479	5.538	4.997	34.309	15.706	5.477	4.673	37.521
Government_Expenditure	***860(860)***	14.706	6.521	0	51.975	14.767	6.573	0	45.959
gpc_GDP_Per_Capita_Growth	***860(860)***	1.027	4.93	-47.503	21.028	1.071	5.154	-31.333	37.536
GDP_Per_Capita	***860(860)***	3756.78	4325.921	436.72	29223.465	4054.927	4484.312	471.325	27242.656
gdpg	***860(860)***	3.69	5.164	-50.248	26.417	3.636	5.168	-30.145	35.224
Foreign_Direct_Investment	***860(860)***	2.976	5.905	-8.703	86.989	2.927	6.862	-28.624	103.337
telefon	***860(860)***	182000	622000	0	5492840	178000	608000	0	5075420
emp_ind	***860(860)***	12.876	8.185	1.505	42.903	13.44	8.812	1.465	43.114
emp_agric	***860(860)***	53.784	21.893	4.6	92.298	52.37	23.075	4.65	92.303
health_exp	***860(860)***	1.667	1.093	.062	5.826	1.694	1.13	.062	6.049
Financial_Deepening	***860(860)***	17.417	20.497	.491	160.125	19.159	21.548	0	150.974
health_expcurrent	***860(860)***	5.37	2.39	1.453	20.413	5.186	2.214	1.453	16.62
cpia_transparency	***860(860)***	2.817	.584	1.5	4.5	2.791	.59	1.5	4.5
cpia_trade	***860(860)***	3.717	.524	2	4.5	3.728	.498	2	4.5
cpia_socprotection	***860(860)***	2.999	.51	2	4.5	3	.513	2	4.5
cpia_publicadmi	***860(860)***	2.913	.457	2	4	2.922	.455	2	4
cpia_publicmgt	***860(860)***	3.015	.453	2	4.1	3.018	.459	2	4
cpia_envtsustain	***860(860)***	3.084	.542	2	4	3.095	.501	2	4.5
cpia_socinclusion	***860(860)***	3.17	.47	2.2	4.3	3.18	.462	2.2	4.3
cpia_macro	***860(860)***	3.64	.641	1.5	5	3.669	.636	1.5	5
cpia_gender	***860(860)***	3.19	.525	2	4.5	3.195	.543	2	4.5
cpia_fiscal	***860(860)***	3.395	.653	1.5	4.5	3.445	.635	1.5	4.5
cpia_finsector	***860(860)***	2.955	.422	2	4	2.952	.422	2	4
cpia_resourceeqity	***860(860)***	3.301	.651	1.5	4.5	3.313	.655	1.5	4.5
cpia_humanresouce	***860(860)***	3.275	.533	2	4.5	3.294	.52	2	4.5
Social_Protection_Score	***860(860)***	18.688	19.584	-17.878	81.201	18.94	19.889	.452	81.201
socialinsurance	***860(860)***	4.199	4.802	.496	59.52	4.475	5.535	.496	59.52
importcost	***860(860)***	664.334	492.041	98.1	3039	681.679	537.369	98.1	3039
exportcost	***860(860)***	591.627	456.34	108.9	2222.7	610.202	484.685	108.9	2222.7
Inflation	***860(860)***	70.852	67.604	0	1344.193	70.536	59.407	0	890.229
banks	***860(860)***	5.007	8.135	.137	54.362	5.342	8.106	.137	54.043
atm	***860(860)***	13.288	19.069	0	83.906	14.657	19.826	0	83.906
agric_VA	***860(860)***	24.667	14.564	1.828	66.033	23.485	14.668	1.881	79.042
bankacc	***860(860)***	22.522	16.281	1.452	79.998	23.604	18.159	1.452	87.113
cleanfuel	***860(860)***	16.488	23.489	.15	93.34	19.306	24.787	.15	93.34
salary	***860(860)***	37.187	10.446	9.339	69.497	37.001	10.18	10.264	60.741
salt	***860(860)***	67.241	26.568	2	98	65.084	27.571	2	98
natresourcerent	***860(860)***	10.946	10.147	0	59.604	10.626	10.069	0	56.939
Inclusive_Growth	***860(860)***	**355.424**	**843.313**	**10.834**	**13934.83**	**354.89**	**848.076**	**14.852**	**14647.05**
kofgi	***860(860)***	40.81	9.949	16.922	72.262	40.93	10.238	17.578	72.354
Economic_Globalisation_Index	***860(860)***	40.869	11.02	15.039	84.48	40.642	11.265	13.188	85.299
koffin	***860(860)***	43.956	11.882	14.067	84.754	43.197	12.155	12.224	86.737
kofso	***860(860)***	32.642	14.399	5.461	78.558	33.071	15.208	4.642	78.383
palma	***860(860)***	7.142	3.191	2.484	30.065	7.09	3.309	2.484	30.065
theil	***860(860)***	.68	.114	.35	1.164	.676	.117	.35	1.165
gini	***860(860)***	60.012	5.539	44.082	86.276	59.827	5.678	44.082	86.832

Note: TR is Training set; TS is Testing set; Std. Dev is Standard deviation; Min is Minimum; Obs is Observation, 860(860 denotes observations in Training and Testing sets, and Max is Maximum

Table A3: Ordinary least square regression on training sample

Variable	Coef.	St.Err.	t-value	[95% Conf	Interval]	Sig
Vulnerable_emp	-.65	.015	-43.21	-.679	-.62	***
unempl	.021	.01	2.03	.001	.042	**
sec_teachers	0	.002	0.00	-.003	.003	
debt_service	.012	.005	2.32	.002	.023	**
registry_time	.001	0	2.43	0	.002	**
business_time	.003	.001	3.83	.001	.005	***
tertiary	.001	.004	0.15	-.008	.009	
taxrev	-.033	.006	-5.29	-.046	-.021	***
Tariff	-.018	.005	-3.66	-.028	-.009	***
soc_contri	.1	.024	4.19	.053	.147	***
service_VA	-.001	.004	-0.23	-.009	.007	
self_employ	-.247	.021	-11.88	-.287	-.206	***
self_emplFE	-.11	.012	-9.34	-.134	-.087	***
sch_enrolTER	.044	.019	2.29	.006	.081	**
infrast_qual	-.031	.018	-1.69	-.066	.005	*
electricaccess_pop	.006	.002	2.89	.002	.011	***
rd	.787	.191	4.13	.413	1.162	***
labforce_MAFE	-.006	.003	-1.85	-.013	0	*
pupiltea_ratio	.015	.003	4.37	.008	.022	***
women_seats	-.025	.004	-6.25	-.033	-.017	***
progressto_sec	-.002	.002	-0.89	-.006	.002	
HIV_preva	.007	.01	0.67	-.013	.026	
prenatal	.003	.003	1.10	-.002	.009	
poverty_hc	-.014	.003	-4.06	-.021	-.007	***
povertyhc_mid	.022	.039	0.56	-.055	.099	
povertyhc_low	-.065	.036	-1.79	-.136	.006	*
povgap_mid	.129	.17	0.76	-.205	.464	
povgap_low	-.054	.098	-0.55	-.247	.138	
urbanization	.029	.003	9.08	.023	.035	***
popgrof	.051	.041	1.25	-.03	.132	
remit	0	.002	-0.23	-.005	.004	
sanitation	-.012	.004	-3.18	-.019	-.005	***
opendefeca_pop	-.004	.004	-0.96	-.013	.004	
unfpa_aid	0	0	-0.86	0	0	
unicef_aid	0	0	-1.08	0	0	
undp_aid	0	0	0.97	0	0	
noda	-.005	.003	-1.59	-.011	.001	
netmigration	0	0	-0.22	0	0	
mortality_5yrs	.003	.001	3.41	.001	.005	***
cellphone	-.002	.001	-1.47	-.005	.001	
manuf_VA	-.001	.001	-0.94	-.004	.001	
logisticquality_TT	.189	.183	1.04	-.17	.548	
logisticqualityl	-.717	.268	-2.67	-1.244	-.19	***
logisticquality_ship	.339	.115	2.94	.113	.565	***
electricity	-.001	.002	-0.78	-.004	.002	
industry_VA	-.004	.004	-1.07	-.012	.004	
povgapmid_increa	.071	.043	1.66	-.013	.154	*
povgaplow_increas	-.227	.076	-3.00	-.375	-.078	***
hci	-12.715	1.013	-12.55	-14.704	-10.726	***
gfcf	0	.003	0.06	-.005	.006	
gov_educ	.007	.006	1.23	-.004	.019	
Government_Exp	-.019	.005	-3.71	-.03	-.009	***
gpc_GDP_Per_Ca	.027	.046	0.58	-.063	.116	
GDP_Per_Capita	0	0	-0.50	0	0	
gdpg	-.023	.044	-0.52	-.11	.064	
Foreign_Direct_In	-.008	.004	-1.90	-.016	0	*
telefon	0	0	-5.91	0	0	***
emp_ind	.035	.01	3.69	.017	.054	***

emp_agric	.015	.004	3.32	.006	.023	***
health_exp	-.106	.053	-2.02	-.21	-.003	**
Financial_Deepen	.002	.003	0.55	-.004	.008	
health_expcurrent	.112	.021	5.31	.071	.154	***
cpia_transparency	-.292	.122	-2.40	-.53	-.053	**
cpia_trade	.039	.096	0.41	-.149	.228	
cpia_socprotection	-.222	.379	-0.59	-.965	.521	
cpia_publicadmi	-.428	.162	-2.64	-.746	-.109	***
cpia_publicmgt	.504	.284	1.77	-.053	1.061	*
cpia_envtsustain	.798	.401	1.99	.011	1.585	**
cpia_socinclusion	-1.791	1.947	-0.92	-5.612	2.031	
cpia_macro	-.004	.087	-0.04	-.175	.168	
cpia_gender	.566	.417	1.36	-.252	1.383	
cpia_fiscal	-.232	.083	-2.79	-.396	-.069	***
cpia_finsector	.061	.113	0.54	-.16	.282	
cpia_resourceeqity	.696	.402	1.73	-.094	1.486	*
cpia_humanresouc	1.112	.411	2.71	.306	1.918	***
Social_Protection_	.002	.002	0.92	-.002	.007	
socialinsurance	-.002	.007	-0.22	-.016	.013	
importcost	.001	0	4.60	0	.001	***
exportcost	-.001	0	-4.57	-.001	0	***
Inflation	-.001	0	-2.90	-.002	0	***
atm	-.001	.002	-0.66	-.005	.003	
agric_VA	0	.003	-0.07	-.006	.006	
bankacc	.004	.003	1.39	-.002	.011	
cleanfuel	.019	.005	3.80	.009	.028	***
salary	-.005	.004	-1.17	-.012	.003	
salt	.004	.002	2.39	.001	.007	**
natresourcerent	.012	.004	3.32	.005	.019	***
kofgi	.057	.011	5.42	.036	.078	***
Economic_Globali	-.017	.008	-1.96	-.033	0	*
koffin	.004	.006	0.68	-.007	.015	
kofso	-.002	.008	-0.25	-.017	.014	
palma	-.08	.022	-3.63	-.124	-.037	***
theil	6.809	1.112	6.12	4.626	8.992	***
gini	-.062	.017	-3.77	-.095	-.03	***
Constant	95.766	1.329	72.08	93.158	98.375	***
Mean dependent var		355.42		843.31		
R-squared		0.999		860.000		
F-test		14837.639		0.000		
Akaike crit. (AIC)		1509.661		1961.569		
Root MSE		**0.5525**				

Table A4: Ordinary least square regression on testing sample

Variable	Coef.	St.Err.	t-value	[95% Conf	Interval]	Sig
Vulnerable_empl	-.772	.015	-52.04	-.801	-.743	***
unempl	-.023	.012	-1.96	-.046	0	*
sec_teachers	.004	.002	2.41	.001	.008	**
debt_service	.016	.01	1.63	-.003	.035	
registry_time	.001	.001	2.17	0	.002	**
business_time	.003	.001	2.88	.001	.005	***
tertiary	.001	.005	0.16	-.009	.011	
taxrev	-.02	.008	-2.60	-.036	-.005	***
Tariff	-.009	.008	-1.21	-.025	.006	
soc_contri	.087	.029	3.01	.03	.144	***
service_VA	-.006	.004	-1.36	-.013	.002	
self_employ	-.106	.023	-4.56	-.151	-.06	***
self_emplFE	-.118	.014	-8.32	-.145	-.09	***
sch_enrolTER	.041	.019	2.12	.003	.079	**
infrast_qual	-.042	.018	-2.29	-.077	-.006	**
electricaccess_pop	.008	.003	2.73	.002	.014	***
rd	.534	.225	2.37	.092	.975	**
labforce_MAFE	-.01	.004	-2.39	-.019	-.002	**
pupiltea_ratio	.018	.004	4.33	.01	.026	***
women_seats	-.006	.005	-1.20	-.015	.004	
progressto_sec	.001	.002	0.57	-.003	.006	
HIV_preva	.031	.012	2.59	.008	.054	***
prenatal	-.001	.003	-0.15	-.007	.006	
poverty_hc	-.005	.004	-1.41	-.013	.002	
povertyhc_mid	.009	.054	0.17	-.097	.116	
povertyhc_low	-.004	.048	-0.08	-.099	.091	
povgap_mid	.022	.235	0.09	-.44	.483	
povgap_low	-.012	.136	-0.09	-.28	.255	
urbanization	.037	.004	8.83	.029	.045	***
popgrof	-.029	.049	-0.60	-.125	.067	
remit	.006	.003	2.39	.001	.011	**
sanitation	-.016	.004	-3.75	-.025	-.008	***
opendefeca_pop	-.001	.005	-0.28	-.012	.009	
unfpa_aid	0	0	-0.54	0	0	
unicef_aid	0	0	0.10	0	0	
undp_aid	0	0	0.80	0	0	
noda	-.002	.004	-0.43	-.009	.006	
netmigration	0	0	1.12	0	0	
mortality_5yrs	.004	.001	3.18	.001	.006	***
cellphone	0	.002	-0.02	-.003	.003	
manuf_VA	0	.003	-0.16	-.005	.005	
logisticquality_TT	-.791	.219	-3.61	-1.222	-.361	***
logisticquality_ov~	.698	.328	2.13	.054	1.342	**
logisticquality_ship	.058	.141	0.41	-.22	.335	
electricity	-.003	.002	-1.81	-.007	0	*
industry_VA	-.002	.005	-0.38	-.011	.007	
povgapmid_increa	-.035	.059	-0.59	-.15	.081	
povgaplow_increas	-.018	.107	-0.17	-.227	.191	
hci	-14.423	1.217	-11.85	-16.811	-12.034	***
gfcf	.001	.004	0.14	-.007	.008	
gov_educ	.019	.007	2.60	.005	.033	***
Government_Exp	-.016	.006	-2.65	-.028	-.004	***
gpc_GDP_Per_Ca pita~h	-.077	.034	-2.27	-.144	-.01	**
GDP_Per_Capita	0	0	4.55	0	0	***
gdpg	.067	.034	2.00	.001	.133	**
Foreign_Direct_In	-.001	.004	-0.32	-.01	.007	
telefon	0	0	-6.42	0	0	***

emp_ind	.056	.01	5.56	.036	.075	***
emp_agric	.018	.005	3.57	.008	.028	***
health_exp	-.009	.063	-0.15	-.132	.114	
Financial_Deepeni	-.006	.004	-1.75	-.013	.001	*
health_expcurrent	.076	.027	2.85	.024	.129	***
cpia_transparency	-.716	.148	-4.83	-1.008	-.425	***
cpia_trade	-.172	.111	-1.54	-.39	.047	
cpia_socprotection	.582	.685	0.85	-.763	1.928	
cpia_publicadmi	-.526	.202	-2.60	-.924	-.129	***
cpia_publicmgt	1.358	.357	3.80	.657	2.059	***
cpia_envtsustain	1.903	.712	2.67	.506	3.3	***
cpia_socinclusion	-5.99	3.508	-1.71	-12.875	.896	*
cpia_macro	.024	.108	0.22	-.188	.237	
cpia_gender	1.444	.716	2.02	.038	2.85	**
cpia_fiscal	-.233	.118	-1.98	-.463	-.002	**
cpia_finsector	.199	.141	1.42	-.077	.476	
cpia_resourceeqity	1.362	.712	1.91	-.036	2.759	*
cpia_humanresouc	1.618	.719	2.25	.206	3.03	**
Social_Protection_	-.004	.003	-1.34	-.009	.002	
socialinsurance	-.016	.008	-1.97	-.031	0	**
importcost	.001	0	5.24	.001	.001	***
exportcost	-.001	0	-4.79	-.001	-.001	***
Inflation	-.002	.001	-3.52	-.004	-.001	***
atm	-.006	.002	-2.80	-.011	-.002	***
agric_VA	0	.004	-0.02	-.007	.007	
bankacc	.006	.003	1.78	-.001	.013	*
cleanfuel	.023	.006	4.14	.012	.035	***
salary	-.022	.005	-4.98	-.031	-.014	***
salt	.003	.002	1.72	0	.007	*
natresourcerent	.007	.004	1.67	-.001	.016	*
kofgi	.042	.012	3.63	.019	.064	***
Economic_Globali	-.03	.01	-3.12	-.05	-.011	***
koffin	.016	.007	2.42	.003	.029	**
kofso	.013	.008	1.57	-.003	.028	
palma	-.03	.025	-1.19	-.079	.019	
theil	3.025	1.229	2.46	.612	5.438	**
gini	-.004	.017	-0.22	-.038	.03	
Constant	93.972	1.494	62.89	91.038	96.905	***
Mean dependent var		354.89		848.07		
R-squared		0.999		860.000		
F-test		12005.269		0.000		
Akaike crit. (AIC)		1795.912		2247.820		
Root MSE		**0.69263**				

Appendix E

	cv	minBIC	adaptive
wagessalary	x		x
povgap_mid	x		x
Vulnerable_Employment	x		x
povertyhc_low	x	x	x
self_employ	x		x
povgap_low	x		x
importcost	x		x
mortality_5yrs	x	x	x
gdpg	x	x	x
emp_ind	x		x
sanitation_pop	x	x	x
kofgi	x		x
women_seats	x	x	x
exportcost	x		x
industry_VA	x		x
povgaplow_increase	x		x
palma	x		x
theil	x		x
povgapmid_increase	x		x
cpia_resourceeqity	x		x
agric_VA	x		x
progressto_sec	x		x
electricaccess_pop	x	x	x
logisticqua_TT	x		x
povert_hc	x		x
gpc_GDP_Per_Capita_Growth	x		x
resallocation	x		x
cellphone	x	x	x
bankacc	x		x
Financial_Deepening	x		x
logisticqua_ship	x		x
cpia_fiscal	x	x	x
cpia_humanresouce	x		x
telefon	x		x
banks	x		x
salary	x		x
cpia_publicadmi	x		x
opendefeca_pop	x	x	x
netmigration	x		x
noda	x		x
infrast_qua	x		x
remit	x		x
gfcf	x		x
hci	x	x	x
logisticqua_overal	x		
Tariff	x		x
exportburden	x		
cpia_envtsustain	x	x	x
undp_aid	x		x
socprotectpop	x		x
cleanfuel	x	x	
GDP_Per_Capita	x		
Economic_Globalisation_Index	x		
sch_enrolTER	x		x
manuf_VA	x		x
Human_Capital	x		x

Figure A2 continued

Human_Capital	x		x
taxrev	x		x
cpia_trade	x	x	
Social_Protection_Score	x		
emp_agric	x		
cpia_finsector	x		
unfpa_aid	x		
Government_Expenditure	x		
electricity	x		
cpia_gender	x		
cpia_publicmgt	x		
gov_educ	x		
Inflation	x		
HIV_preva	x		
sec_teachers	x		
popgrof	x		
business_time	x		
unempl	x		
gini	x		
service_VA	x		
rd	x		
Foreign_Direct_Investment	x		
cpia_transparency	x		
health_exp	x		
_cons	x	x	x

Figure A2: Selected variables via Standard lasso (cv), Minimum BIC lasso (MinBIC), and Adaptive lasso (adaptive)