Estimating and Forecasting Growth Model by Least Absolute Shrinkage and Selection Operator (LASSO): A Cross Country Analysis

Nasir Mahmood[1] and Benish Rashid[2]

Abstract

Least Absolute Shrinkage and Selection Operator (LASSO) is a method of regression analysis that uses regularization to increase prediction accuracy for the selection of variable and model interpretability. The model which is having least root mean square error is the best model. We also estimate the retention frequency for each model to show the repeatedly significant variable in each modeling for all the countries. In this study six growth models have been used for analyzing the main determinants of economic growth in case of cross countries. Time Series Data from 1980 to 2020 were used to analyze the cross country growth factors therefore, the current study looked at 43 countries with modelling these different comparative studies based on growth modelling. So, we can make six individual models and we can estimate the General Unrestricted Model with the use of econometric technique Least Absolute Shrinkage and Selection Operator. Current study found that in case of nested model or full model it is concluded that model with lag value of GDP, trade openness, population, real export, and gross fixed capital formation are the main and potential determinants to boost up the economic growth in most of the countries.

Key words: LASSO, Economic Growth, Cross Country,

Introduction

The traditional method assumes that there must be more observations than parameters in the regression model. What if the number of observations is more than the number of parameters? Then the old ways of doing things don't work. A shrinkage method is one way to solve the problem. Most of the shrinkage method is based on mathematical programming techniques and their tools. The shrinkage methods get rid of high-dimensional data for lower gains, which makes it easy to get rid of irrelevant candidates. Least Absolute Shrinkage and Selection Operator (LASSO), which was made by Tibshirani in 1996, was the first popular method of this type. It can shrink some coefficients to zero.

Second, LASSO is a method for both estimating the model and choosing the variables at the same time. LASSO can work when there are more variables than observations, and it gives the sparse models (Zhao and Yu, 2006, Meinshausen and Yu, 2009). As Efron et al. (2004) and Friedman et al. (2010) have shown, the LASSO's regularization path can be figured out quickly. Tibshirani has a number of generalizations and different types of the LASSO procedure that can be used to solve a number of problems (2011). Particular attention has been paid to Elastic Net (E-Net) and Adaptive LASSO (adaLASSO).

Meinshausen (2006) compared the two shrinkage procedures of model selection LASSO and RELAXED LASSO. A two-stage procedure, referred to as the relaxed Lasso, was found to be effective in overcoming the conflicting requirements of a proficient computational procedure and fast convergence rates of the 2-loss. Relaxed Lasso solutions for orthogonal designs provide a continuum of solutions that include both soft-and hard-thresholding of estimators. It is possible to compute all relaxed Lasso solutions at the same time as computing all regular

[1] *Lecturer in Economics, Higher Education Department, Punjab.*
Email: nasirmahmoodlec2012@gmail.com
[2] *Ph.D Scholar, Pakistan Institute of Development Economics, Islamabad.*
Email: benishrashid83@gmail.com

Lasso solutions, and computing all relaxed Lasso solutions is often as expensive as computing all regular Lasso solutions. It has been demonstrated in the study's numerical results that the relaxed Lasso estimator produces lighter models with the same or lower prediction loss than the regular Lasso estimator when dealing with high-dimensional data.

FU (1998) analyzed the different models such as Bridge, Ridge and Lasso. Researchers have compared different criteria for model selection based on a compression approach. He performed an empirical exercise and used the prostate cancer data. According to results, Bridge regression performs better from ridge and LASSO.

In cross-sectional modeling, Epprecht et al. compare the LASSO and AdaLASSO estimate with classical technique (Autometrics) in forecasting and covariate selection. The result indicates that LASSO and AdaLASSO estimates outperform Autometrics in prediction.

Methodology

Model Selection Procedures based on shrinkage methodology

The model selection procedures that are consisted on Shrinkage approach are based on mathematical programming techniques. These techniques remove high dimensionality of the data and shrink irrelevant variable to zero. The Least Absolute Shrinkage and Selection Operator (LASSO) is a popular estimation method in a linear regression framework, introduced by Tibshirani (1996). LASSO method is like ridge regression; however, it set some coefficients precisely equal to zero with a substantial bias. The resulting model is easy to interpret and possesses the least forecast error. LASSO has ability to estimate the parameters and to select the variable at a time. Before LASSO stepwise selection method is most widely used for choosing the regressors in which only prediction accuracy is improved in certain cases mostly prediction is worse. When there are more variables then observations LASSO has ability to handle this. (H. Zou ,)

Computational Detail of Least Absolute shrinkage and selection operator (LASSO)

It is a method of regression analysis that uses regularization to increase prediction accuracy selection of variable and model interpretability. The most commonly used method before LASSO was the selection of stepwise regression, which only improved the accuracy of the prediction if only a few covariates were strongly related to the result. In some cases, a prediction error may worsen. Ridge regression reduces fitting by reducing large coefficients of regression, but does not perform covariate selection, and hence does not contribute to the model's interpretability.

By forcing the absolute value of the regression coefficients to be less than a fixed value, LASSO can achieve which goal by effectively choosing a simple model without these coefficients. Although this concept is similar to ridge regression, the squares of the coefficients are forced to sum up below a fixed value in ridge regression.

Where The LASSO coefficients, $\hat{B}_j^L$, minimize the quantity

$$\sum_{i=0}^{n}\left(y_i - B_0 - \sum_{j=1}^{p} B_j x_{ij}\right)^2 + \lambda\sum_{j=1}^{p}\left|B_j\right| = \mathrm{RSS} + \lambda\sum_{j=1}^{p}\left|B_j\right|$$

The LASSO, like ridge regression, reduces coefficient estimates to zero. However, Because of the L1 penalty, some coefficient estimates are forced to be exactly zero when the tuning parameter is large enough. Bias increases with. Variance rises as falls.

No parameters are removed when $\lambda = 0$. The estimate is the same as linear regression.So, like best subset selection, the LASSO selects variables. As a result, LASSO models are easier to interpret than ridge regression models. There are several algorithms of LASSO such as:Adaptive LASSO, Elastic Net Relaxed LASSO, etc. are all based on shrinkage methodologies.

Theory Based Models for Economic Growth

Model 1	Akram, et al. (2011). LnGDP =f (FDI(inf) , T Debts, DI , Inf)
Model 2	Mihaela, et al. (2017). LnGDP =f (Inf, LnTLF, TOTP, FDI (inf), GExp)
Model 3	Sami, et al. (2014). LnGDP =f (Edu, RExp, P(remi), FDI)
Model 4	Ajmair, et al. (2015). LnGDP =f (Inf, LnGCF , Rexp, P(remi)
Model 5	Al-Smadi, (2020). LnGDP = f (FDI ,TOP, LG, DI , LnGCF ,)
Model 6	. Udeaja, et al. (2015). LnGDP =f (DI, FDI, Edu, TOP)

The Econometric Model takes following form.

LnGDP =f (FDI, TOP, LG, DI, LnGCF, TDebts, INF , LnTLF, LnTOTP EDU, LnREXP , LnGEXP , REMI)

$$LNGDP_t = \beta_o + +\beta_1 \mathrm{FDI(inf)}_t + \beta_2 \mathrm{TOP}_t\ \beta_3 \mathrm{LG}_t + \beta_4 \mathrm{DI}_t + \beta_5 \mathrm{LnGCF}_t + \beta_6 \mathrm{TDebtS}_t + \beta_7 \mathrm{Inf}_t + \beta_8 \mathrm{LnTLF}_t + \beta_9 \mathrm{LnTOTP}_t + \beta_{10} \mathrm{Edu}_t + \beta_{11} \mathrm{LnRExp}_t + \beta_{12} LnGExp_t + \beta_{13} P(remi)_t + \mu_t$$

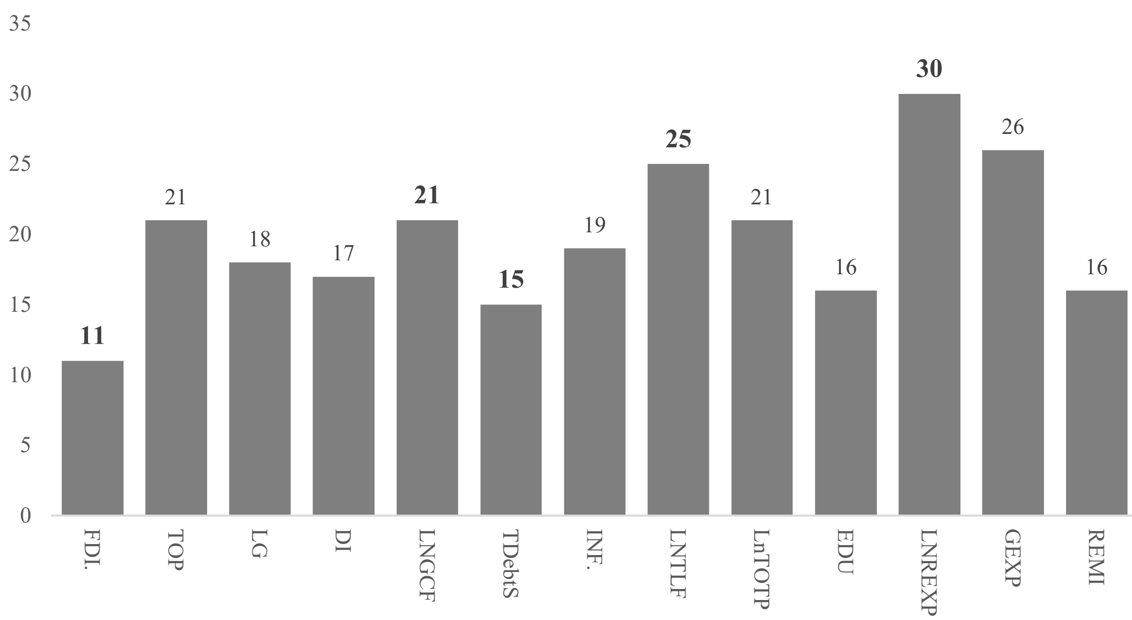
Selecting Models for Economic Growth
In Table 1 the results of Absolute Shrinkage and Selection Operator (LASSO) are presented. The results in Table 1 are based on the modeling of economic growth. In this model the economic growth (LNGDP) is dependent variable and Gross fixed capital formation (LNGCF), total Labor force (LNTLF), foreign direct investment (FDI) and independent variables are trade openness (TOP), labor growth (LG), domestic interest (DI), total debts (TDebts), inflation (INF), total population (LNTOTP), education expenditure (EDU), exports of goods and services (LNREXP), personal remittances (REMI), and government expenditure (LNGEXP). In this modeling, the FDI, LNGCF, and LNTLF are our focus variables while the LNGEXP, REMI, LNREXP, EDU, LNTOTP, INF, TDebts, DI, LG, and TOP are the auxiliary variables.

Results of LASSO Regression
Table below shows results of estimation using LASSO. The table provides regression coefficients of the variables in the model. The last column gives root mean square errors for

each estimated model. The cells marked with (…) indicates the variables that were excluded from the model by LASSO. Row 1 indicates that for Argentina, the focus variable FDI and LNGC, and the next most common auxiliary variables are DI, INF and EDU found to be insignificant were excluded by LASSO. Similarly for Austria Focus Variable FDI and other auxiliary variables LG, TDebts, LNTOTP,EDU and REMI traced out be insignificant and therefore dropped by estimation procedures. In the last row, the frequency of retention of each variable is provided. The table shows that out of 43 countries, the variable LNEXP was significant in 31 cases. The final results of Table 1 shows that the forecast RMSE of LASSO modeling for growth of Argentina is lowest 0.01 while the forecast RMSE for United States growth model is highest 3.57. So, according to the LASSO modeling based on forecast RMSE the Argentina model forecast performance is best and United States model has worst ability to forecast.

Insert Table 1 Here (Appendix-1)

Figure 1: Graph of the Retention Variables in LASSO for Growth Modeling



The figure above shows the total significance of variables in LASSO modeling for all countries. In focus variables the foreign direct investment (FDI) found significant 11 times out of 43 regressions. The gross fixed capital product (LNGCF) got significant 21 times out of 43 regressions. While the total labor force (LNTLF) 25 times found significant out of 43 regressions. It means in focus variables the LNTLF is highly significant in repeated modeling and FDI got low significance.

In case of auxiliary variables, the export of goods and services (LNREXP) found highly significant in modeling 30 times out of 43 regression and total debts (TDebts) got less significance 15 times out of 43 regressions.

Conclusion

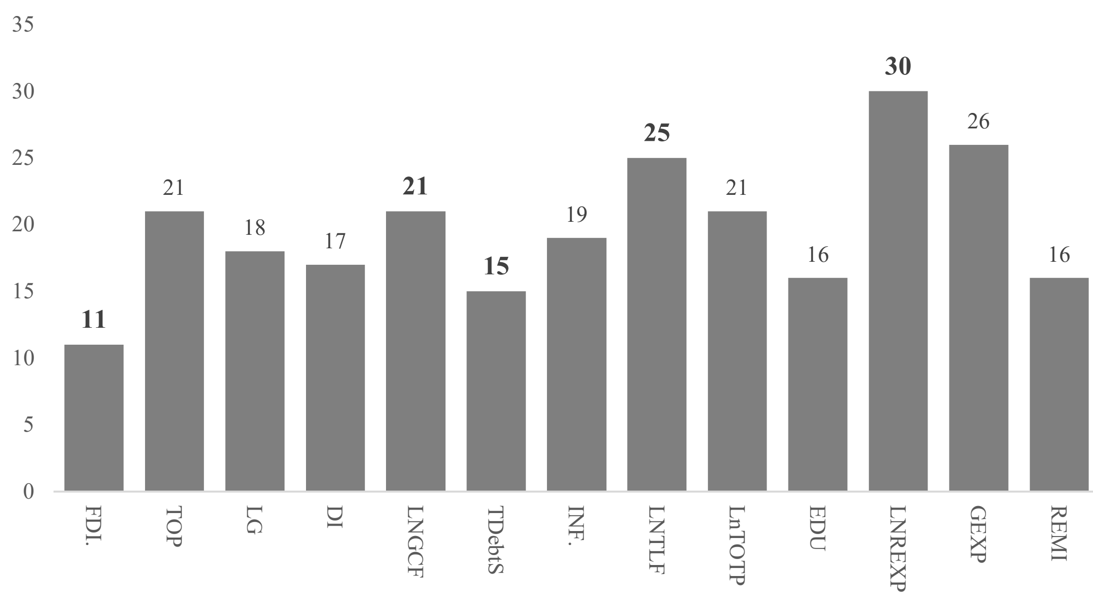
The main goal of this study was to compare the performance of different 43 countries for the same growth model to find the performance of potential determinants by using Least Absolute Shrinkage and Selection Operator (LASSO). There are some points of agreement as well. The variable LNREXP, LNGEXP, LNGCF, LNTLF, TOP. LNTOTP traced out to most of the time significant determinants by applying the Shrinkage procedure LASSO. While the best country based on least FRMSE shows that forecasted value of Argentina is lowest 0.01 while the forecast RMSE for United States growth model is highest 3.57. So, according to the LASSO modeling based on forecast RMSE the Argentina model forecast performance is best and United States model has worst ability to forecast.

References

- Abou-Ali, H. and Kheir-El-Din, H. (2009), "Inflation and Growth in Egypt: Is There a Threshd Effect?", Middle East Development Journal, 1(1): 59-78
- Anwar, S., & Nguyen, L. P. (2010). Foreign direct investment and economi Therefore, by using these six models we have tested them for non-nested and nested encompassing through Cox test and F-test respectively.c growth in Vietnam. Asia Pacific business review, 16(1-2), 183-202.
- Awan, M. Z. (2010), "A Nexus between foreign direct investment & Pakistan's economy", International Research Journal of Finance and Economics, 52.
- Ahmad, N., Hayat M.F., Luqman M., Ullah S. (2012). "The Causual Links Between Foreign Direct Investment and Economic Growth in Pakistan", European Journal of Business and Economics, Volume 6, pp. 20-21
- Anyanwu, J.C. (2014), "Factors Affecting Economic Growth in Africa: Are There any Lessons from China", African Development Review, 26(3):468-493.
- Barro, R. J. (1996). Determinants of economic growth: A cross-country empirical study.
- Bundell R., Dearden, L., Meghir, C. and Sianesi, B. (1999), "Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy", Fiscal Studies, 20(1): 1-23.
- Bassanini, A. &Scarpetta, S. (2001), "Does Human Capital Matter for Growth in OECD Countries? Evidence from Pooled Mean-Group Estimates", OECD Economics Department Working Papers, No. 282.
- Bajona, C., Kehoe, T. J., &Ruhl, K. J. (2010). Trade liberalization, growth and productivity, University of Nottingham, UK, unpublished paper.
- Biwott, P. K., Moyi, E. D., &Khainga D. (2013).Trade Liberalization and Economic Growth: The Role of Regulatory Policies. Journal of World Economic Research, 2(3), 2013, 45-57.
- Calero, Carla (2008) Remittances, Liquidity Constraints and Human Capital Investments In Ecuador. World Development (www.elsevier/locate/worlddev).
- Gupta, S., Pattillo, C.A. and Wagh, S. (2009), "Effect of Remittances on Poverty and Financial Development", World Development, 37(1)
- Giuliano, P., & Ruiz-Arranz, M. (2009). Remittances, financial development, and growth. Journal of development economics, 90(1), 144-152.
- Hendry, D. F. (1988). The encompassing implications of feedback versus feedforward
- mechanisms in econometrics. Oxford Economic Papers, 40(1), 132-149.
- Hendry, D. F., & Richard, J. F. (1983). The econometric analysis of economic time
- series. International Statistical Review/Revue Internationale de Statistique, 111-148.
- Jackson,

- Hoover, K. D., & Perez, S. J. (1999). Data mining reconsidered: encompassing and the general-to-specific approach to specification search. The Econometrics Journal, 2(2), 167-191.
- Mamo, F. (2012). Economic Growth and Inflation: A panel data analysis
- Silaghi, M. P., & Medeşfălean, R. (2014). Some insights about determinants of economic growth in Romania. An empirical exercise. Theoretical & Applied Economics, 21(6).
- Uneze, E. (2013). The relation between capital formation and economic growth: evidence from sub-Saharan African countries. Journal of Economic Policy Reform, 16(3), 272-286.

Figure 2: Graph of Total Significance of Variables in LASSO for Growth Modeling



Appendix-I

Table 1: The Results of Least Absolute Shrinkage and Selection Operator for Growth

The Results of Least Absolute Shrinkage and Selection Operator for Growth Mo

Variables	Constant	LnGCF	LnTLF	FDI	TOP	LG	DI	TDebts	INF	LnTOTP	Ed
Country Name											
Argentina	16.224	..	-0.016	..	-11.983	2.015	..	-0.003	..	0.533	..
Australia	-1.381	..	-4.394	..	..	-8.438	-8.145	..	-6.903	1.881	-3.
Austria	27.256	0.095	-0.011	..	-11.378	..	-0.005	..	-0.0002	..	..
Bangladesh	0.781	0.134	..	..	..	..	..	..	..	0.646	..
Belgium	-3.623	9.406	-5.4	..	..	..	..	..	-3.18	2.253	..
Bhutan	-2.727	..	..	..	..	-6.874	..	..	-8.246	3.64	..
Bulgaria	1.873	3.143	..	..	..		..	..		0.775	..
Brazil	-2.641	0.679	0.045	..	..	0.432	..	-0.036	..	..	..
Canada	21.498	0.194	..	..	..	..	..	-0.025	..	..	..
China	-99.42	0.707	..	0.015	1.1069	..	..	..	..	5.156	..
Chili	6.997	..	-0.002	..	..	0.357	-0.002	..	-0.002	..	..
Denmark	51.781	-0.105	..	0.001	..	-1.262	..	..	0.001	-3.183	0.
France	7.562	..	..	..	..	..	..	..	..	..	..
Germany	-291.952	0.788	-1.154	-0.003	-288.285	-4.205	..	-0.179	-0.029	16.481	0.
Ghana	20.68	-0.099	..	0.075	..	..	..	..	..	..	..
Hungary	13.898	..	0.079	..	..	..	..	..	..	..	-0.
India	14.089	..	..	..	..	0.155	0.298	-0.016	..	..	..

Variables **Country Name**	**Constant**	**LnGCF**	**LnTLF**	**FDI**	**TOP**	**LG**	**DI**	**TDebts**	**INF**	**LnTOTP**	**Edu**	**LnRexp**	**REMI**	**LnGexp**	**RMSE**
Indonesia	7.992	..	..	..	..	..	..	..	0.003	..	..	..	..	..	**0.605**
Iran	22.423	..	..	..	..	..	0.021	-0.021	..	..	..	..	..	0.138	**0.621**
Japan	4.456	6.014	4.026	-2.431	-1.906	4.305	-8.508	-1.993	-3.044	-2.334	-1.623	-1.633	-3.977	4.336	**0.707**
Luxembourg	-11.82	0.439	-0.05	..	..	-0.356	-0.172	-0.731	..	1.911	0.1296	..	0.001	..	**0.195**
Malaysia	1.042	0.514	-0.044	-0.0253	..	-0.247	0.043	-3.946	..	..	-0.045	-0.122	-0.011	0.237	**0.296**
Maldives	-3.913	..	-5.82	1.246	..	9.244	1.027	2.781	..	4.697	1.831	..	-3.806	1.572	**0.774**
Mexico	7.799	0.283	0.04	..	..	..	..	..	-0.0006	0.089	..	..	..	0.285	**0.136**
Morocco	2.095	3.938	-1.679	..	-1.036	..	-1.989	..	-2.805	..	..	5.052	..	..	**0.034**
Nepal	20.904	..	-0.009	0.013	-8.409	-2.017	-0.0773	..	..	..	..	0.3501	0.004	0.009	**0.081**
Netherland	9.926	..	-0.152	..	-0.365	..	..	..	-0.007	..	..	0.101	..	0.001	**0.082**
New Zealand	21.659	..	..	..	-9.833	..	..	..	-0.002	..	..	0.811	..	..	**0.069**
Norway	17.235	..	..	..	..	..	-0.033	..	..	..	..	0.42	..	..	**0.159**
Pakistan	8.261	..	..	..	-7.531		0.001	-0.008	..	0.714	-0.016	0.007	..	..	**0.063**
Peru	1.637	3.304	-7.044	4.494	-8.903	-5.039	..	..	-1.696	2.5202	..	2.687	-6.965	1.354	**0.072**
Paraguay	1.747	8.064	..	..	-1.048	-1.367	..	-3.083	-4.048	1.834	3.026	3.898	..	1.814	**0.024**
Philippines	15.372	0.081	..	-0.003	-10.259	-0.004	-0.003	-0.0114	..	0.441	..	0.071	..	0.301	**0.025**
Portugal	6.362	..	-0.029	..	..	..	..	..	-0.007	..	..	0.287	..	..	**0.067**
Qatar	46.382	..	0.012	..	-10.877	..	-0.01	..	..	0.007	..	..	..	..	**0.178**
South Africa	2.305	..	1.063	..	-1.151	..	..	..	..	..	4.65	5.083	..	-3.964	**0.011**
Sri Lanka	2.84	0.144			-9.459			-0.001		0.986	-0.009	0.189		0.156	**0.032**

Variables **Country Name**	**Constant**	**LnGCF**	**LnTLF**	**FDI**	**TOP**	**LG**	**DI**	**TDebts**	**INF**	**LnTOTP**	**Edu**	**LnRexp**	**REMI**	**LnGexp**	**RMSE**
Switzerland	23.217	..	-0.066	..	-10.896	..	..	..	-0.002	..	0.009	0.314	..	0.002	**0.023**
Sweden	2.583	..	-1.455	..	-1.189	-2.365	7.653	2.882	..	..	..	9.517	..	..	**0.017**
Turkey	9.972	..	..	3.337	-9.95	..	-3.336	..	-8.731	7.193	-6.375	2.153	1.154	1.886	**0.06**
United States	19.446	0.1	0.063	..	..	1.099		..	-0.014	..	-0.603	..	0.793	..	**3.57**
United Kingdom	20.773	..	-0.003	..	-10.541		..	..	..	..	..	0.475	..	0.007	**0.018**
Uruguay	-2.989	0.174	-0.002	..	-8.77	..	..	..	..	1.36	0.003	0.367	-0.165	0.007	**0.073**
Retention Frequency		**21**	**25**	**11**	**21**	**18**	**17**	**15**	**19**	**21**	**16**	**30**	**16**	**26**	