Examining the South African Economic and Social Renaissance

A. Panko, G. Harke, J. Joubert, K. Hoo, P. Lee

Abstract

The following paper has a look at the historical context of South Africa while attempting to identify key economic drivers. To do this a Random Forest was fit to the data in order to leverage the power of machine learning and determine the feature importance. It was found that Investment GFCF is a key driver and it was confirmed by applying a linear regression in the traditional fashion.

Keywords: Economics, History, Risk Management, Machine Learning, South Africa

---

1. Introduction

South African history provides great context for a case study in risk management. Analysts have the opportunity to investigate the effects of sociopolitical and economic hardships during the apartheid regime followed by an African Renaissance that has been criticized for implementing so called neo-liberal economic policies (Mbeki 2016a and Mbeki 2016b).

The following sections are split up as follows: section 2 acts as a literature review by analyzing the historical contexts and economic drivers of South Africa's economy. Section 3 summarizes section 2 and dicusses the key economic drivers. Section 4 describes the data. Section 5 reviews the methodology of linear regression and random forests in the context of modeling financial time series. Section 6 discusses our findings. Section 7 Concludes.

2. Historical Context

2.1. The Impact of Apartheid

After South Africa's all-white National Party gained power in 1948, it began enforcing existing policies of racial segregation under a system called apartheid. Under apartheid, non-white South Africans were forced to live in separate areas from whites and use separate public facilities. Apartheid were implemented for 50 years despite strong oppositions within and outside of South Africa.

In the two decades following the rise to power of the National Party, whites (particularly Afrikaners) rose above all other ethnic groups in South Africa through their dominant and tactful performance in the labour market. Under apartheid system, black people (greater than 70% of the population) were pushed to the margins of their land through the imposition of the Land Act of 1913. In result; many blacks are unskilled, illiterate, and have low living standards. Apartheid also attracted sanctions and disinvestment which impacted the South Africa economy on a significant scale especially after mid-1980s.

Back in 1950, South Africa's GDP per capita (ranked 24th) were ahead of countries such as Japan (29th), Turkey (33rd). Based on World Banks data available from 1960, it is clear that South Africa's GDP per capita growth were lagging behind these comparable countries during its apartheid period:
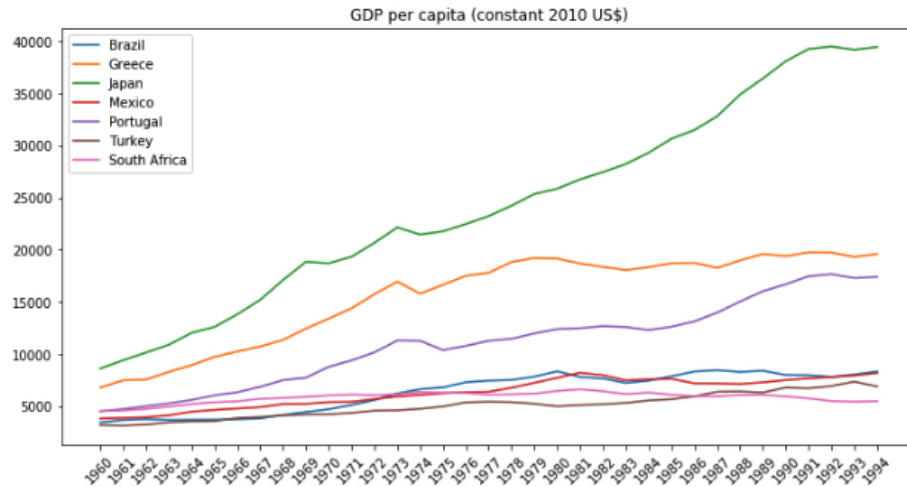

Figure 1: GDP per capita
WorldBank 2019a

2.1.2. Impact on Political Landscape

The apartheid government, under the influence of the ruling National Party, had, by the beginning of the 1980s, divided South Africa into five entities:

1. The Republic of Transkei
2. The Republic of Bophuthatswana
3. The Republic of Venda
4. The Republic of Ciskei
5. The Republic of South Africa

The South African government at the time considered the first four entities to be legally independent countries, but they never received international recognition of their 'statehood'. The international community regarded these four 'republics' as apartheid creatures, the only purpose of which was that of disenfranchising the majority of the citizens of South Africa. In terms of the National Party's ideology, Africans (who constituted close to 80% of the population of the old South Africa) were supposed to be citizens of one of these and other potentially 'independent' republics (e.g. one for Zulus in the old Natal Province). They were not entitled to representation in the national parliament.

2.1.3. Impact on Social Demographics

Based on Human Development Index (HDI) available from 1990, we are able to compare data from comparable countries for part of the apartheid period until 2008:
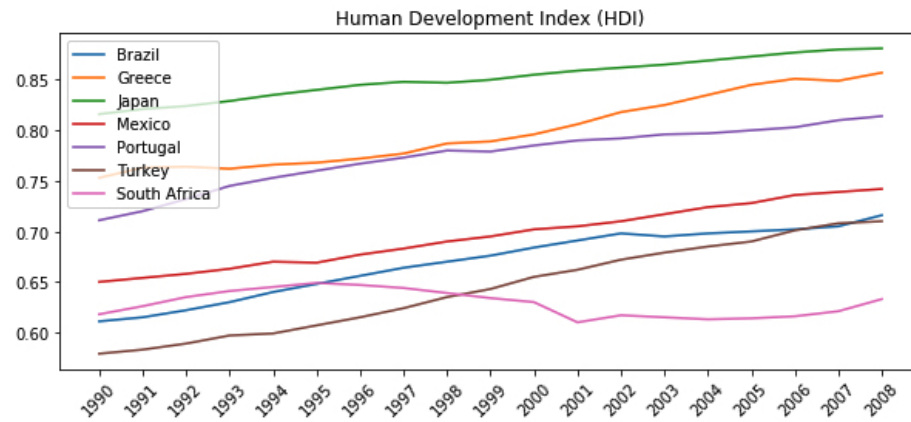

Figure 2: Human Development Index (HDI)
UNDP 2019

2.2. Cost of Sanctions

In 1986, South Africa's most important trading partners - (US, EC, Japan) imposed trade and financial sanctions. Specifically:

1. Europe and Japan sanctioned import of the Kruger rand and certain steel and iron products
2. Germany and Great Britain made recommendations and imposed no binding sanctions
3. US embargoed importing the Kruger rand and certain steel and iron products
4. OPEC imposed an oil embargo

5. Europe sanctioned new direct investments but left it to member states to declare if the sanctions would be binding where upon England and Germany did not impose binding sanctions
6. US sanctioned new direct investments and was the only country to impose sanctions on portfolio investments and credits/loans

Ultimately, the total cost of the trade/financial/oil sanctions were estimated to be 1.5% with trade sanctions accounting for 1.3%. During the sanctions period (4th quarter 1986 1st quarter 1991) South Africa suffered a net capital outflow of 16.2 billion Rands, which is equivalent to 2% of GNP (Hefti and Staehelin-Witt 2013).

Among the list of sanctions, the oil embargo represented the most painful measure for South Africa as it impacted the total population by raising the oil prices whereas it could either gain new trade partners (steel and coal) or modify production (gold bars instead of Kruger Rands) to alleviate the impact of other sanctions

On the other side, the relative meager impact of the financial sanctions was that reinvestment of profits was exempted. In fact, 80%+ of all FDI in South Africa during that period originated from reinvested profits and reinvested profits by 3 billion Rands during the sanctions which demonstrates investors had maintained a certain trust in South Africa. (Hefti and Staehelin-Witt 2013).

2.2.1. Analysis

"Even in the absence of sanctions, apartheid ultimately would have collapsed due to the economic stresses of a hugely inefficient system. Although sanctions may have hurried this process, they were not the driving force behind it. The fall of apartheid was not engineered by foreigners, nor was it primarily precipitated by foreign sanctions" (Lowenberg and Kaempfer 1998).

On its face, this was a human rights issue; in reality it may have been closer to culmination of socioeconomic realities and geopolitical evolution of the times. First, the decline of communist regimes lessened the political importance of the South African government as a hedge against communist stronghold in Africa.

Second, the socio-economic stimulus had been brewing with urbanization of the labor force and limitations of segregated development zones inherent in Apartheid. This limited the existing system as a viable economic solutions in an integrated economy.

As Game-theoretic models (Kaempfer and Anton 2007) suggest, success of sanctions depends on conflict expectations and levels of commitment of the participating nations. From this perspective, sanctions were limited as they were designed with domestic policies in mind and often continued ongoing economic trends. This is shown in that:

4

1. The net outflow of foreign investment was highest in 1985 which one year before imposition of economic sanctions.
2. US call for coal import sanctions among its partners would have benefited the US coal industry as US and South Africa were competitors for the European coal market.

2.3. Privatization in South Africa Post 1994

The post apartheid new government inherited big debts and over 300 state-owned enterprises and that guzzled subsidies and, for the most part, offered rotten service at extortionate prices. 1994 government was not in favor of privatization.Rather it followed Public Private Partnership model by selling equity to "strategic equity partners" and Black Empowerment Groups while retaining a majority interest (Jerome 2004).

Privatization aim was to reduce public borrowing, attract foreign investment, promote industrial competition and fuel economic growth.

2.3.1. Evaluation

Privatization had many obstacles. Most public firms were in a mess, over-staffed and deep in debt in 1994. The government reckons, for instance, that 27'000 jobs need to be lost at Transnet, the state transport company, and 10'000 at Telkom. Potential investors preferred the government to do this work. Due to high unemployment, government delayed the initiative. South Africa's powerful trade unions were another problem. Their frequent strikeswhich brought many public services to a halt and deter investors (Economist 1999).

Also government was having conflicting view and aims on privatization. One side, it was trying for privatization to raise money and bring in private-sector and foreign expertise. On other side it was seeking public firms to extend services to poor (usually black) areas, which the private sector might ignore(Hentz 2000).

Johannesburg Stock Exchange (JSE) supposed to contribute in structural shift to economy with help of financial institutions but due to liquidity and extensive cross holding it didn't contribute much.

In broader view, the program has not achieved much. Large number of state owned enterprises running businesses.The program didn't deliver much in the area of black economic empowerment.

Privatization in South Africa had been slow, with few visible results.

2.4. Reconstruction and Development Programme (RDP)

In 1994, Nelson Mandela government developed Reconstruction and Development Programme (RDP) framework in South Africa. A special ministry was established under Jay Naidoo to put initiative in place (O'Malley 1998).

2.4.1. Objectives

The RDP emphasized two objectives: the reconstruction of the country and the alleviation of poverty. The RDP's goal was to assist government in integrating growth with social development and economic reconstruction (Besada 2007).

The RDP framework proposed five suggestion for economic growth:

1. Human resource development
2. Meet basic necessities
3. Inject democracy
4. Public private partnership projects
5. Boost economy

2.4.2. Evaluation

The RDP displayed a dual character. First, as a policy framework, increase donor aid and shrink government's spending. It helped to addressing poverty in states. Also, less funding for military and more funds were allocated to resolve social inequalities in different areas.

At the same time, RDP fund was established to finance high-profile 'presidential projects', such as public works projects for unemployed youths, free medical care for pregnant mothers, and the electrification of homes of the poor in townships and rural villages. Between April 1994 and December 1998, government built approximately 500 new clinics, which served the health care needs of potentially five million individuals. Through vaccination programs, measles and polio were eradicated(Besada 2007).

2.5. Growth, Employment and Redistribution (GEAR)

GEAR is a long term, 5 year, macroeconomic policy used to achieve the goals set out in the basic policy framework of the RDP. The reason for the additional policy was because the RDP failed to boost economic growth due to its lack of skilled managers in government, and a narrow focus on fiscal prudence and the reallocation of existing revenues (SAHO 2014).

At the time (1996), projected economic growth was 3%, this would lead to a failure which:

1. Would not be able reverse the unemployment crisis
2. Provided inadequate resources for the expansion of social service delivery
3. Too slow for an equitable distribution of income and wealth

Table 1: Economic Forecast from 1996, Before GREAR

Base Scenario Projections: 1996-2000

Model characteristics	1996	1997	1998	1999	2000	Average
Fiscal deficit (% of GDP) (fiscal year)	5,1	4,5	4,0	3,5	3,0	4,0
Real government consumption (% of GDP)	19,8	19,5	19,1	18,6	18,1	19,0
Average tariff (% of imports)	10,0	9,0	9,0	8,0	8,0	8,8
Average real wage growth, private sector	0,8	1,5	1,7	1,3	1,4	1,4
Average real wage growth, government sector	4,8	0,4	0,4	0,3	0,0	1,2
Real effective exchange rate (% change)	-9,6	0,7	0,1	0,1	0,0	-1,8
Real bank rate	7,0	6,0	5,0	4,5	3,7	5,2
Real government investment growth	2,6	2,4	2,2	2,2	2,4	2,4
Real parastatal investment growth	3,0	2,5	2,5	2,5	3,0	2,7
Real private investment growth	6,3	4,2	4,4	5,8	7,1	5,6
Real non-gold export growth	9,6	7,5	6,4	5,5	5,3	6,9
Results	**1996**	**1997**	**1998**	**1999**	**2000**	**Average**
GDP growth	3,3	2,0	2,5	2,9	3,3	2,8
Inflation (CPI)	8,4	10,9	9,6	9,3	9,1	9,5
Employment growth (non-agricultural formal)	0,9	1,0	0,8	0,9	1,3	1,0
New jobs per year ('000s)	97	101	84	103	134	104
Current account deficit (% of GDP)	1,8	1,3	1,1	1,1	1,6	1,4
Real export growth, manufacturing	12,5	10,4	7,5	6,6	5,4	8,5
Gross private savings (% of GDP)	20,5	20,7	20,8	20,8	20,6	20,7
Government dissavings (% of GDP)	3,1	2,6	2,0	1,4	0,9	2,0

Manual, Marcus, and Ramos 1996

Hence the need for a long term macroeconomic policy to boost economic growth to 6% per annum. This would lead to an additional 400000 jobs per year until the year 2000.

The GEAR policy would focus on the following points mentioned in (Manual, Marcus, and Ramos 1996):

1. Accelerated growth of non-gold exports
2. Expansion in private sector capital formation
3. Public sector investment
4. Employment intensity of investment and output growth
5. Increase in infrastructure development and service delivery.

Table 2: Economic Forecast from 1996, With GREAR

Integrated Scenario Projections: 1996-2000						
Model characteristics	1996	1997	1998	1999	2000	Average
Fiscal deficit (% of GDP) (fiscal year)	5,1	4,0	3,5	3,0	3,0	3,7
Real government consumption (% of GDP)	19,9	19,5	19,0	18,5	18,1	19,0
Average tariff (% of imports)	10,0	8,0	7,0	7,0	6,0	7,6
Average real wage growth, private sector	-0,5	1,0	1,0	1,0	1,0	0,8
Average real wage growth, government sector	4,4	0,7	0,4	0,8	0,4	1,3
Real effective exchange rate (% change)	-8,5	-0,3	0,0	0,0	0,0	-1,8
Real bank rate	7,0	5,0	4,0	3,0	3,0	4,4
Real government investment growth	3,4	2,7	5,4	7,5	16,7	7,1
Real parastatal investment growth	3,0	5,0	10,0	10,0	10,0	7,6
Real private sector investment growth	9,3	9,1	9,3	13,9	17,0	11,7
Real non-gold export growth	9,1	8,0	7,0	7,8	10,2	8,4
Additional foreign direct investment (US$ m)	155	365	504	716	804	509
Results	**1996**	**1997**	**1998**	**1999**	**2000**	**Average**
GDP growth	3,5	2,9	3,8	4,9	6,1	4,2
Inflation (CPI)	8,0	9,7	8,1	7,7	7,6	8,2
Employment growth (non-agricultural formal)	1,3	3,0	2,7	3,5	4,3	2,9
New jobs per year ('000s)	126	252	246	320	409	270
Current account deficit (% of GDP)	2,2	2,0	2,2	2,5	3,1	2,4
Real export growth, manufacturing	10,3	12,2	8,3	10,5	12,8	10,8
Gross private savings (% of GDP)	20,5	21,0	21,2	21,5	21,9	21,2
Government dissavings (% of GDP)	3,1	2,3	1,7	0,7	0,6	1,9

Manual, Marcus, and Ramos 1996

GEAR was also focused on (SAHO 2014):

1. Reducing fiscal deficits
2. Lowering inflation
3. Maintaining exchange rate stability
4. Decreasing barriers to trade
5. Liberalizing capital flows

In many ways GEAR was successful in that consumption targets were almost met with the addition of greater macroeconomic stability, better reporting, accountability, and the management of public finances improved. It was also successful in that it reversed the the negative growth rate of the early nineties (SAHO 2014).

However GEAR was criticized largely by the Congress of South Africa's Trade Unions (COSATU) for its neo-liberal approach which they consider to be in stark contradiction to the RDP, which lead to the ANC party neglecting to implement their nationalisation policies.

GEAR resulted in low levels of economic growth and private investment which led to a disappointing unemployment rate. The policy also failed to redistribute

wealth which can be seen in South Africas shocking Gini Coefficient of 0.65, 2nd highest in the world! (WorldBank 2014)

In 2005 GEAR was replaced with the Accelerated and Shared Growth Initiative for South Africa (ASGISA)

2.6. President Thabo Mbeki (1999 - 2008)

Thabo Mbeki was South Africa's second post apartheid president, serving from 1999 up to 2008 at which point he resigned shortly before the end of his second term, as he was recalled by the National Executive Committee (NEC) of the ANC.

Mbeki was famous for implementing policies such as (NEPAD, OAU, and the AU) and became known as the quintessential African nationalist after Professor Adam Habib declared so in (The_Economist 2006). It was noted that his main goal was to establish the new South Africa as, first and foremost, a black African country (The_Economist 2006)

As mentioned previously GEAR was largely criticized by COSATO as being a neo-liberal structural adjustment program.

In 2016 Mbeki published a letter titled GEAR and Neo-Liberalism in which he defends his stance in a lengthy essay. He concludes the following: During the whole period when GEAR has been challenged from the left, the assertion was made that GEAR sought to replace the RDP. I am certain that there is no rational presentation that can be made to prove this assertion. Alternatively there is no credible argument that can be presented which shows that in the policies and programmes Government actually implemented, once GEAR was adopted it abandoned the pursuit of the RDP objectives. (Mbeki 2016a and Mbeki 2016b)

Mbeki was known for reducing state spending however in 1999 he became involved in the now famous arms deal which cost the South African tax payer dearly and was in contrast to his original stance of reducing government spending. (SAHO 2017)

Another scandal during this time was the German Frigate Consortium which saw the people of South Africa purchase 4 new war ships, each worth R4-billion. It was later revealed that a bribe of R130 million was paid to South African politicians.

2.7. Accelerated and Shared Growth Initiative for South Africa (ASGISA)

The Accelerated and Shared Growth Initiative for South Africa (ASGRISA) was launched by Deputy President Phumzile Mlambo-Ngcuka in February 2006. The target was halving unemployment and poverty between 2004 and 2014. This

could be achieved if the economy grew at an average rate of at least 4.5% in the period to 2009, and by an average of 6% in the period 2010 to 2014.

To achieve the objectives of ASGISA there was a need to address some key constraints in the economy. These included the relative volatility of the currency, the cost, efficiency and capacity of the national logistics systems, shortages of suitably skilled labour, and barriers to entry, limits to competition, the regulatory environment, and deficiencies in state organisation, capacity and leadership.

Some broad policy areas had been identified as main pillars for ASGISA to accelerate economic growth, including macroeconomic issues, infrastructure investment, education and skills development, industrial and sector strategies, second economy Initiatives, governance and State capacity Issues.

The start in 2006-2007 was promising: investments and public sector infrastructure expenditure increased. Many infrastructure projects were started. Unemployment rate was decreased due to creation of additional job places.

However, in 2007/2008 international financial crisis came. Additionally the administration in South Africa changed. These two factors made ASGISA being confined to the rubbish bin. Zuma administration simply closed ASGISA.

2.8. National Development Plan (NDP)

The National Development Plan 2030 is an important policy document designed by the National Planning Commission. The Commission is an advisory body first constituted in 2009 by President Jacob Zuma. The Plan consists of proposals which are supposed to eliminate poverty and reduce inequality by 2030.

The plan was created to create a country where everyone embraces their full potential, a country where opportunity is determined not by birth, but by ability, education and hard work without existed poverty and inequality. It is clear that the progress in any one area depends on development of another. This means that to achieve these goals the economy must grow faster and at the same time there must be no corruption and crime and so on. That is why the plan includes action across all sectors of South African society. There were nine primary challenges in the main focus:

1. Too few people work
2. The quality of school education for black people is poor
3. Infrastructure is poorly located, inadequate and under-maintained
4. Spatial divides hobble inclusive development
5. The economy is unsustainably resource intensive
6. The public health system cannot meet demand or sustain quality
7. Public services are uneven and often of poor quality
8. Corruption levels are high
9. South Africa remains a divided society.

The plan proposed reasonable actions for every of these areas, however, unfortunately, NDP went the same way as GEAR and ASGISA. All the ideas were great but have not been implemented and goals were not achieved. The plan was too broad, financial crisis played its role and as was written in the plan: "Its success will depend on all South Africans taking responsibility for the plan, led by the President and Cabinet". However, such type of assumptions are to strong for the real life. It's impossible to fight all problems at the same time being too optimistic about the people and the world itself.

2.9. Stock Market Developments

Based on World Bank data available from 1975 on value of stocks traded as % of GDP, we compared the South Africa's metrics against list of comparable countries (based on similar GDP per capita during 1950). As illustrated below, the stock market capitalization of South Africa grew exponentially post the apartheid era and became comparable to developed country like Japan in terms of its market capitalization to GDP ratio.
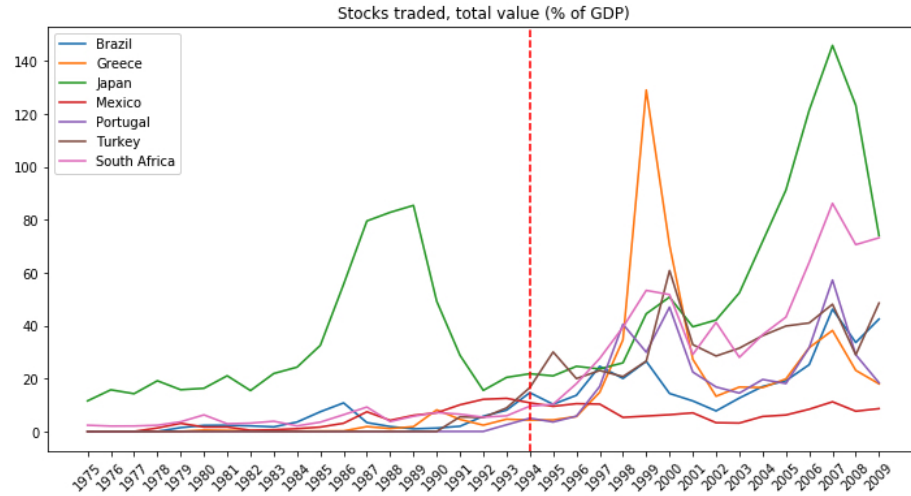

Figure 3: Stocks Traded as a Total Value Percentage of GDP
WorldBank 2019b

South Africa stock market was worth close to twice the country's GDP, which was larger in proportion to most of the emerging countries. The currency, bond and derivatives markets are also among the worlds twenty largest by turnover (Hassan 2013).

South Africa's Johannesburg Stock Exchange (JSE) was founded in response to the need for capital to fund the mining sector and it developed along with the growth and expansion of the mining industry. On top of equity, bonds, commodity, and foreign exchange, the JSE also permits trading in interest rate

derivatives. It has one of the largest number of product lines per exchange than any other derivatives exchange worldwide. Although the South African market is small comparing to the world's major financial centers, it is one of the more significant among the emerging countries.

2.10. Global Financial Crisis (GFC)

The GFC was a crisis in confidence that led to a global credit freeze. This was initiated by a credit boom and housing bubble - partly initiated by the Clinton agenda to increase home ownership during the 1990s - and worsened by financial product innovations which multiplied risks exponentially.

The chain of events that led to the GFC include:

1. Housing bubble
2. Asymmetric information/moral hazard
3. Securitization
4. Off-balance sheet vehicles which allowed institutions to skirt reserve requirements
5. Rating agencies

2.10.1. Impact of GFC on South African Economy

The South African economy was dependent on financial services and export of manufactured goods/primary commodities (gold, platinum and chrome). Ultimately, the GFC resulted in decreased production due to commercial credit/consumer demand reductions. This led to decline of export and import volumes ultimately leading to net capital outflows from the economy.

Financial Indicators reflected such a downturn: Current account balance as a percentage of the GDP increased from 1.1% 2008 Q4 to 5.8% 2009 Q2. GDP growth rate decreased from 1.8% 2008 Q4 to -3.2% 2009 Q2. Inflation rose to 8.9% with real unemployment of 32% (Ravinder and Malindi 2014).

2.10.2. Specific Sectors

Impact on mining sector was shown by decrease in global demand and prices for South African mineral products following the GFC due to tighter fiscal policies. For instance, gold price fell from USD 1030 to USD 750 in 2018. Mining decreased 12.8% between 2008 Q4 and 2009 Q1. Furthermore, any capital improvement was put on hold due to credit reduction (Nyanjowa 2008).

Macroeconomic imbalance was manifested in 1) decreased capital investment inflow that led to reduced credit availability which was felt most acutely in service sectors associated with debt-driven expenditure such as construction and automobiles, and 2) fall in manufacturing, mining, financial, business, real estate, and wholesale/retail trade that led to reduction in GDP. For instance,

GDP decreased 7.4% in Q1 2009 and 2.8% in Q2 2009. This led to the first recession in 17 years (ILO 2010).

However, recession was shallower than expected due to:

1. Lack of direct banking sector exposure to problem assets in the US and strong performance going into the GFC
2. Low oil prices
3. Growth in construction in preparation for the FIFA World Cup

Unemployment rate rose from 21.9% in 2008 and 25.2% 2010. This was felt most strongly in the low-skilled sectors leading to additional social costs. Ultimately, this lead to increased budget deficit and high public debt which recorded a 1% surplus in 2008 to -7.3% in 2010 (Gordhan 2010).

3. Key Economic Drivers

After 50 years of apartheid, South Africa's growth was severely limited in terms of economic, political, and social demographic development. The post apartheid period, despite Government efforts to privatize the inefficient, heavily indebted state-owned enterprises, proved more than challenging.

Both The RDP and GEAR policies, which were initiated during Mandelas presidency, aimed to boost economic growth and social development. Sadly both resulted in low levels of economic growth and disappointing unemployment rates - it also failed to redistribute wealth. During Mbekis presidency, ASGISA replaced GEAR with the objective to halve unemployment and poverty between the years 2004 and 2014. To accelerate economic growth, ASGISA targeted broad policy areas such as macroeconomic issues, infrastructure investment, education and skills development.

The start of 2007 was promising: investments in public sector infrastructure expenditure increased and many infrastructure projects were started. The unemployment rate was decreased due to the creation of additional jobs. However, the 2008 Global Financial Crisis came and ASGISA was closed after Zuma started his presidency in 2009.

South Africa was not as adversely impacted as other nations by the GFC due to decreased exposure to the toxic assets. Given the changes made on the social and economic frontiers, the South African economy appeared to have progressed forward.

3.1. Structural Advancements

South Africa has in the two decades since 1994 made large structural advances in 10 key areas, according to a study by Goldman Sachs (Coleman 2013)

1. Macro fiscal and monetary balances have improved
2. Government debt costs have trended lower and foreign reserves have risen
3. Overall cost of capital has declined
4. Corporate valuations have improved relative to global peers
5. Real asset ZAR returns have compared favourably
6. China and African trade rise has largely offset European trade decline
7. Disposable income of South Africans has risen
8. The rise of the black middle class has led to a structural boost in spending
9. Wage inflation and government grants have supported this trend
10. Per unit labour productivity has improved

Despite not fully achieving the intended results, the post-apartheid era policies mentioned in this paper are the key drivers which transformed South Africa to achieve a period of economic performance between 1994-2007 where it recorded average GDP growth rate of 3.6% (vs the 1.4% between 1980-1994) and lower average inflation of only 6.3% (vs the 14.3% between 1980-1994). These policies improved the macro fiscal and monetary balances, reduced government debt costs (from 49.7% of GDP in 1994 to 28.3% in 2007) and increased foreign reserves (from US$3.1bn in 1994 to US$39.7bn in 2009).

The corporate environment also improved as the overall cost of capital was reduced (from a lending rate of 15.6% in 1994 to 13.2% in 2007), stock valuations improved relative to the global peers (from a gap of 15x difference in forward P/E around 2000 period to equal P/E ratio in 2007), and trading with China increased significantly (exports to China increased from 1.5% to 12%, imports increased from 5% to 14%).

From a social economic perspective, the policies boosted the ratio of middle and upper class (from 48% in 2001 to 69% in 2010), and increased the disposable income of South Africans with impressive CAGR of 10% (from R665bn in 2001 to R1,494bn in 2009).

4. Data

The following data was taken from the Organisation for Economic Co-operation and Development (OECD) database, Quandl, and the World Bank.

Independent Variables

1. Investment (GFCF)
2. Current Account Balance
3. CPI (Inflation)
4. Long-term Interest Rates
5. M3 Money Supply
6. Unemployment

7. Ores and Metals Import
8. Crop Production (Wheat)
9. Exchange Rates (USD/ZAR)
10. Prime Energy

Dependant Variables

1. Gross Domestic Product (GDP)
2. Share Prices (JALSH Index)

4.1. Feature Engineering

As with all statistical modeling, variables need to be engineered to allow models to accurately forecast. Two transformations were applied to the data. The first order log difference was applied to:

1. Gold
2. Exchange Rate
3. Share Prices
4. Long term interest rates
5. M3 Money supply
6. Unemployment
7. GDP

The reason for this is to make the data stationary. Prices are log normally distributed and returns are normally. Next we apply Min Max Scaling to the following variables:

1. CPI
2. Ores metals
3. Prime Energy
4. Investment GFCF
5. Current Account Balance

5. Methodology

Machine learning is starting to build a good reputation in the financial literature as an effective tool. We made use of a Random Forest's (RF) feature importance to help identify the key drivers in the South African economy and Stock Market. After identifying the most important driver, we applied OLS linear regression to analyse the relationship. Our methodology combined the strengths of RF's non-linearity features and OLS's transparency in explaining the relationship between the variables.

This section will briefly provide a short explanation of both a RF's feature importance and the OLS Linear Regression technique.

5.1. Linear Regression: Ordinary Least Square (OLS)

Ordinary least Square is linear regression method. In OLS linear Regression, the goal is to find the line (or hyperplane) that minimizes the vertical offsets. We define the best-fitting line as the line that minimizes the mean squared error (MSE) between our target variable (y) and our predicted output over all samples i in our dataset of size n.

$$MSE = \frac{1}{n} \sum_{i-1}^{n} (y_i - \hat{y}_i)^2$$

Following are the approaches to perform a linear regression model using ordinary least squares:

1. Solving the model parameters analytically (closed-form equations)
2. Using an optimization algorithm (Gradient Descent, Stochastic Gradient Descent, Newtons Method, Simplex Method, etc.)

5.1.1. Normal Equation

The closed-form solution should be preferred for smaller datasets if computing matrix inverse is not a concern. For very large datasets, or datasets where the inverse of $X^\top X$ may not exist (the matrix is non-invertible or singular, e.g., in case of perfect multicollinearity), the GD or SGD approaches are to be preferred (Raschka and Mirjalili 2017). The linear function (linear regression model) is defined as:

$$y = w_0 x_0 + w_1 x_1 + ... + w_m x_m = \sum_{j=0}^{m} w^\top x$$

where y is the response variable, x is an m-dimensional sample vector, and w is the weight vector (vector of coefficients). Note that w_0 represents the y-axis intercept of the model and therefore $x_0 = 1$. Using the closed-form solution (normal equation), we compute the weights of the model as follows:

$$w = (X^\top X)^{-1} X^\top y$$

5.2. Assumption in OLS (DeFusco et al. 2007)

OLS has the next assumptions:

1. The regression model is linear in the coefficients and the error term
2. The error term has a population mean of zero
3. All independent variables are uncorrelated with the error term

4. Observations of the error term are uncorrelated with each other
5. The error term has a constant variance (no heteroscedasticity)
6. No independent variable is a perfect linear function of other explanatory variables

Traditionally, in quantitative finance linear regression based factor analysis is used to analyse the performance of the factors in different factor models (such as CAPM, French Fama five factor model). OLS is widely used to model the risk distribution. We believe this traditional approach of quantitative finance will help in our analysis to explain the dependent variables GDP share price. Also, we observed dependent variables, GDP and share price, are linearly correlated with explanatory variables and hence OLS will help us to find best fit values for coefficient to explain variability in dependent variables. To perform this analysis using linear regression, we have used statsmodels OLS api.

5.3. Random Forests Feature Importance

Random Forest is an ensemble methods that generates many classifiers and aggregates their results. Boosting (Shapire et al. 1998) and bagging (Breiman 1996) of classification trees are two common methods. It is used for classification and regression projects and adds randomness by searching for the most important feature among a random subset of features to split a node. Increased randomness can be produced by using random thresholds for each feature rather than searching for the best possible thresholds. A wide diversity results generally produces a better model.

The relative importance of each feature is measured by evaluating how much the tree nodes reduce impurity across all trees in the forest. The score is automatically computed for each feature after training and scaling the results such that the sum of all importance is equal to 1. Increasing the number of features will increase the risk of overfitting. Here, the algorithm randomly selects observations and features to build several decision trees and then averages the results.

Random forest variable importance measures may not be reliable if predictor variables vary in the scale of measurement or number of categories. As an example, if predictors include both sequence data and continuous variables. When subsampling without replacement is used, the resulting variable importance measures can be used reliably (Strobl et al. 2007). Estimate of error rate is quite accurate, given that enough trees have been grown (otherwise the OOB estimate can bias upward; see Bylander 2002).

The Hyperparameters in random forest can be used to: increase the predictive power of the model via maximizing number of features or minimizing number of leafs required to split an internal node. It can also make the model faster via increasing the number of processors used and assuring replicability when

17

given definite value of random state and the same hyperparameters as well as the same training data.

Interpretation of the result requires understanding the trade off between bias and variance. For instance, even though the model is predicting the right answer on average, higher variance may make it more difficult to trust the prediction. On the other hand, in a low variance-high bias situation, the predictions are biased but may be more closely clustered. The results may be more operationalised if the bias is understood given its consistency.

6. Results

We trained two RF models separately for each of the dependent variables: GDP and Share Prices. Both RF models point to **Investment GFCF** as the most important feature, with a greater than 0.3 feature importance. Note that all the values sum to 1.
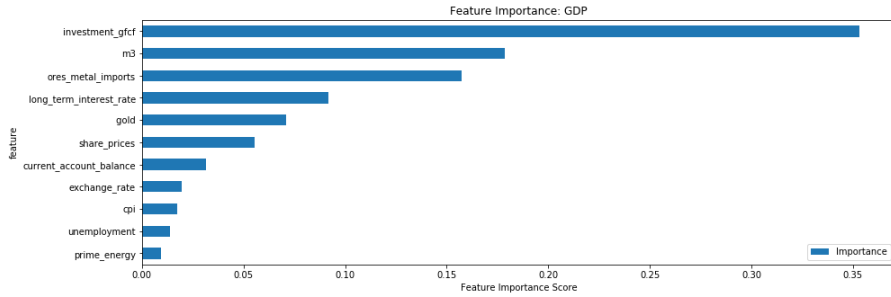

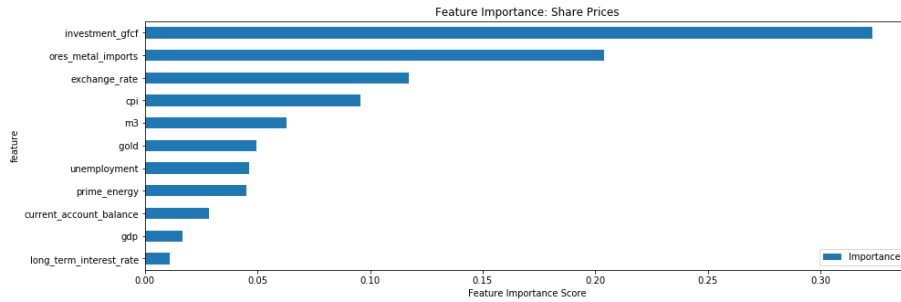Figure 4: Feature Importance for GDP

Figure 5: Feature Importance for Share Prices

6.1. Model GDP - Linear Regression

Next Investment GFCF is used in a traditional linear model. We first perform the analysis on a univariate model on GDP and then Share Prices. Next a

18

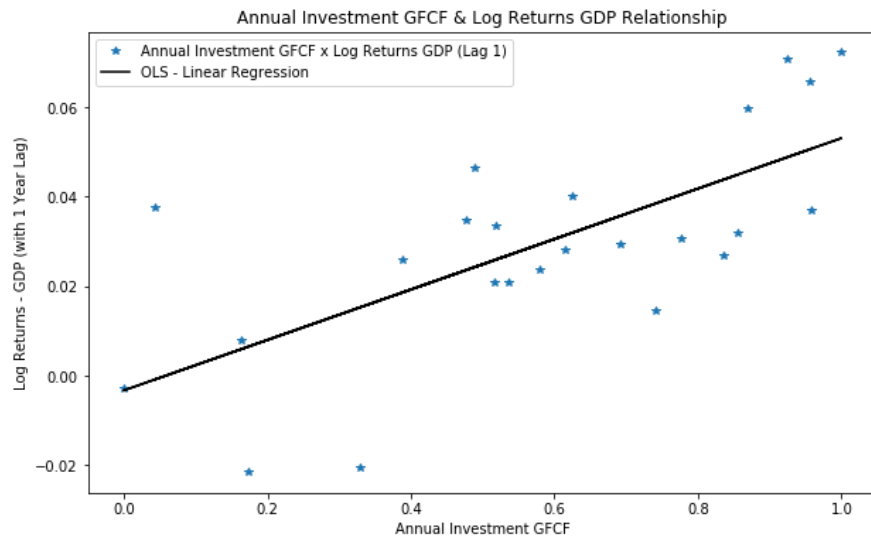multivariate analysis is performed.

Figure 6: Annual Investment GFCF x Log Returns GDP (Lag 1)

Table 3: Model GDP - Linear Regression

OLS Regression Results

Dep. Variable:	gdp	R-squared:	0.470
Model:	OLS	Adj. R-squared:	0.446
Method:	Least Squares	F-statistic:	19.50
Date:	Mon, 21 Jan 2019	Prob (F-statistic):	0.000219
Time:	12:18:47	Log-Likelihood:	63.476
No. Observations:	24	AIC:	-123.0
Df Residuals:	22	BIC:	-120.6
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.0033	0.008	-0.393	0.698	-0.021	0.014
investment_gfcf	0.0563	0.013	4.416	0.000	0.030	0.083

Omnibus:	0.031	Durbin-Watson:	1.385
Prob(Omnibus):	0.985	Jarque-Bera (JB):	0.078
Skew:	0.008	Prob(JB):	0.962
Kurtosis:	2.721	Cond. No.	4.76

As we can see from the graph above, the fitted line reflects the data rather good, for a single variable. Moreover, the F-statistic and p-value of the coefficient are highly significant. Adjusted R-squared of 0.446, meaning that 44.6% of the variability of the response data around its mean is explained by the model.

6.2. Model Stock Market - Linear Regression
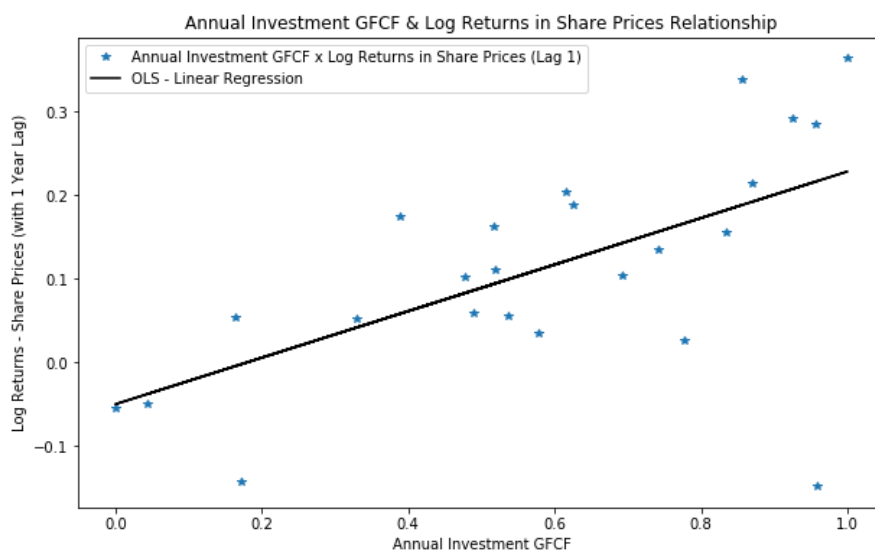
Figure 7: Annual Investment GFCF x Log Returns in Share Prices (Lag 1)

Table 4: Model Share Price - Linear Regression

OLS Regression Results

Dep. Variable:	share_prices	R-squared:	0.362
Model:	OLS	Adj. R-squared:	0.333
Method:	Least Squares	F-statistic:	12.47
Date:	Mon, 21 Jan 2019	Prob (F-statistic):	0.00188
Time:	12:20:10	Log-Likelihood:	19.744
No. Observations:	24	AIC:	-35.49
Df Residuals:	22	BIC:	-33.13
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.0503	0.052	-0.977	0.339	-0.157	0.057
investment_gfcf	0.2788	0.079	3.532	0.002	0.115	0.442

Omnibus:	17.699	Durbin-Watson:	2.103
Prob(Omnibus):	0.000	Jarque-Bera (JB):	21.801
Skew:	-1.565	Prob(JB):	1.84e-05
Kurtosis:	6.466	Cond. No.	4.76

For Share Price, however, the model is not as good as before. With the Adjusted R-squared at 0.333.

6.3. Model GDP with All Features

Table 5: Model GDP with All Features - Linear Regression

OLS Regression Results

Dep. Variable:	gdp	R-squared:	0.868
Model:	OLS	Adj. R-squared:	0.747
Method:	Least Squares	F-statistic:	7.165
Date:	Mon, 21 Jan 2019	Prob (F-statistic):	0.000981
Time:	12:21:22	Log-Likelihood:	80.149
No. Observations:	24	AIC:	-136.3
Df Residuals:	12	BIC:	-122.2
Df Model:	11		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.0326	0.012	2.620	0.022	0.005	0.060
gold	0.0688	0.025	2.733	0.018	0.014	0.124
exchange_rate	0.0848	0.042	1.999	0.069	-0.008	0.177
share_prices	-0.0714	0.026	-2.785	0.017	-0.127	-0.016
long_term_interest_rate	-0.0027	0.024	-0.112	0.912	-0.055	0.050
m3	0.2489	0.105	2.372	0.035	0.020	0.477
unemployment	-0.0566	0.050	-1.132	0.280	-0.165	0.052
current_account_balance	-0.0569	0.017	-3.301	0.006	-0.094	-0.019
cpi	-0.0236	0.028	-0.847	0.413	-0.084	0.037
ores_metal_imports	-0.0493	0.019	-2.622	0.022	-0.090	-0.008
prime_energy	-0.0077	0.013	-0.615	0.550	-0.035	0.020
investment_gfcf	0.0439	0.020	2.221	0.046	0.001	0.087

Omnibus:	1.518	Durbin-Watson:	2.347
Prob(Omnibus):	0.468	Jarque-Bera (JB):	0.923
Skew:	-0.480	Prob(JB):	0.630
Kurtosis:	2.954	Cond. No.	69.0

23

6.4. Model Stock Market with All Features

Table 6: Model Share Price with All Features - Linear Regression

OLS Regression Results

Dep. Variable:	share_prices	R-squared:	0.622
Model:	OLS	Adj. R-squared:	0.276
Method:	Least Squares	F-statistic:	1.797
Date:	Mon, 21 Jan 2019	Prob (F-statistic):	0.164
Time:	12:21:46	Log-Likelihood:	26.037
No. Observations:	24	AIC:	-28.07
Df Residuals:	12	BIC:	-13.94
Df Model:	11		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.0565	0.117	-0.484	0.637	-0.311	0.198
gold	0.0773	0.248	0.312	0.760	-0.463	0.618
exchange_rate	0.2393	0.412	0.581	0.572	-0.658	1.136
long_term_interest_rate	-0.1290	0.233	-0.554	0.590	-0.637	0.379
m3	1.1340	0.921	1.231	0.242	-0.873	3.141
unemployment	-0.6144	0.448	-1.371	0.196	-1.591	0.362
gdp	-2.5270	2.169	-1.165	0.267	-7.253	2.199
current_account_balance	-0.0697	0.165	-0.424	0.679	-0.428	0.289
cpi	0.0549	0.278	0.197	0.847	-0.552	0.662
ores_metal_imports	0.0365	0.165	0.221	0.829	-0.323	0.396
prime_energy	-0.0705	0.113	-0.625	0.543	-0.316	0.175
investment_gfcf	0.2260	0.216	1.045	0.317	-0.245	0.697

Omnibus:	21.044	Durbin-Watson:	2.734
Prob(Omnibus):	0.000	Jarque-Bera (JB):	32.274
Skew:	-1.704	Prob(JB):	9.81e-08
Kurtosis:	7.545	Cond. No.	139.

The following table does a comparison of multivariate and univariate r-squared values.

Table 7: Comparison of Multivariate and Univariate

	All	Investment GFCF
GDP	0.868	0.470
Share Prices	0.622	0.362

7. Conclusion

We applied Random Forest Regressor (RF) on the 10 selected key factors to model against GDP and Share Prices as proxies to portfolio risk. The RF identified Investment GFCF as the most significant feature (importance scores close to 0.35 for both GDP and Share Prices).

We then applied Ordinary Least-Squares (OLS) Regression on Investment GFCF as a factor on both GDP and Share Prices. For GDP (logged returns with 1 year lag), the OLS model achieved an R-Squared (i.e. the coefficient of determination) of 47% with just Investment GFCF as input factor. The Share Price model achieved a lower R-Squared score of 36.2%.

We also illustrated the relationship of the Investment GFCF factor and GDP, Share Prices returns in 2 separate charts together with the corresponding regression lines.

We compared the performance of our single-factor OLS models against models trained with all 10 initial factors. As seen in table 7 above, although the R-Squared scores are more impressive from the all-factors OLS models, our single-factor models still model relationships reasonably well for just a single variable (R-Squared of 36.2% for Share Prices and 47% for GDP)

References

[Bes07]   H Besada. "Fragile Stability: Post-Apartheid South Africa". In: (2007). [Online; accessed 04-January-2019]. URL: https://www.files.ethz.ch/isn/39546/WP_27.pdf.

[Bre96]   L Breiman. "Bagging predictors". In: *Machine Learning* 18 (1996). [Online; accessed 19-January-2019], 123fffdfffdfffd140. URL: https://link.springer.com/content/pdf/10.1007/BF00058655.pdf.

[Byl02]   T Bylander. "Estimating generalization error on twoclass datasets using out-of-bag estimates." In: *Machine Learning* 48 (2002). [Online; accessed 19-January-2019], 287fffdfffdfffd297. URL: https://www.r-project.org/doc/Rnews/Rnews_2002-3.pdf.

[Col13]   Colin Coleman. "Two Decades of Freedom - A 20 Year Review of South Africa". In: (2013). [Online; accessed 04-January-2019]. URL: https://www.goldmansachs.com/insights/archive/colin-coleman-south-africa/20-yrs-of-freedom.pdf.

[DeF+07]   R. DeFusco et al. "Quantitative investment analysis workbook. 2nd ed." In: (2007). [Online; accessed 19-January-2019], p. 65.

[Eco99]   Economist. "The painful privatisation of South Africa". In: (1999). [Online; accessed 04-January-2019]. URL: https://www.economist.com/international/1999/09/09/the-painful-privatisation-of-south-africa.

[Gor10]   Pravin Gordhan. " Budget Speech by the Minister of Finance him-
             self to the National Parliament". In: (2010). Cape Town: South
             Africa.

[Has13]   Shakill Hassan. "South African Capital Markets: An Overview". In:
             (2013). [Online; accessed 04-January-2019]. URL: `https://econrsa.`
             `org/system/files/publications/working_papers/working_`
             `paper_391.pdf`.

[Hen00]   J.J Hentz. "The Two Faces of Privatisation: Political and Economic
             Logics in Transitional South Africa". In: 38, 2 (2000). [Online; ac-
             cessed 04-January-2019]. URL: `https://www.jstor.org/stable/`
             `161649?read-now=1&seq=8#page_scan_tab_contents`.

[HS13]   E Hefti and E Staehelin-Witt. "Research summary PNR42+ Eco-
             nomic Sanctions against South Africa and the Importance of Switzer-
             land Ch". In: (2013).

[ILO10]   ILO. "South Africa's Response to the Crisis." In: *G 20 Briefs* (2010).

[Jer04]   A Jerome. "Privatisation and Regulation in South Africa: An Eval-
             uation". In: (2004). [Online; accessed 04-January-2019]. URL: `http:`
             `//citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.`
             `198.189&rep=rep1&type=pdf`.

[KA07]   William Kaempfer and D Anton. "The Political Economy of Eco-
             nomic Sanctions. Lowenberg". In: *Handbook of Defense Economics*
             2 (2007), pp. 867–911.

[Mbe16a]   Thabo Mbeki. "GEAR and Neo-Liberalism (Part 1)". In: (2016).
             [Online; accessed 04-January-2019]. URL: `https://www.sahistory.`
             `org.za/archive/gear-and-neo-liberalism-part-1-thabo-`
             `mbeki-21-march-2016`.

[Mbe16b]   Thabo Mbeki. "GEAR and Neo-Liberalism (Part 2)". In: (2016).
             [Online; accessed 04-January-2019]. URL: `https://www.sahistory.`
             `org.za/archive/gear-and-neo-liberalism-part-2-thabo-`
             `mbeki-28-march-2016`.

[MMR96]   Trevor Manual, Gill Marcus, and Maria Ramos. "Growth, Employ-
             ment and redistribution A Macroeconomic Strategy". In: (1996).
             [Online; accessed 04-January-2019]. URL: `http://www.treasury.`
             `gov.za/publications/other/gear/chapters.pdf`.

[Nya08]   W Nyanjowa. "The Impact of Global Financial Crisis on South
             Africa's Mining Industry." In: (2008). [Online; accessed 23-December-
             2018]. URL: `http://www.frost.com/sublib/dis-%20play-`
             `market-insight-top.do?id=149029949`.

[OMa98]   P O'Malley. "Reconstruction and Development Programme (RDP)".
             In: (1998). [Online; accessed 04-January-2019]. URL: `https://omalley.`
             `nelsonmandela.org/omalley/index.php/site/q/03lv03445/`
             `04lv03446/05lv03493.htm`.

[RM14]   Rena Ravinder and Msoni Malindi. "Global Financial Crises and
             its Impact on the South African Economy: A Further Update". In:
             Journal of Economics 5:1 (2014), pp. 17–25.

[RM17]   S. Raschka and V Mirjalili. "Python Machine Learning. 2nd ed".
          In: (2017). [Online; accessed 19-January-2019], pp. 311–322.

[SAH14]   SAHO. "South Africafffdfffdfffds Key economic policies changes (1994
          - 2013)". In: (2014). [Online; accessed 04-January-2019]. URL: https:
          //www.sahistory.org.za/article/south-africa%5C%E2%5C%
          80%5C%99s-key-economic-policies-changes-1994-2013.

[SAH17]   SAHO. "Thabo Mvuyelwa Mbeki". In: (2017). [Online; accessed 04-
          January-2019]. URL: https://www.sahistory.org.za/people/
          thabo-mvuyelwa-mbeki.

[Sha+98] R Shapire et al. "Boosting the margin: A new explanation for the
          effectiveness of voting methods." In: Annals of Statistics 26 (1998).
          [Online; accessed 19-January-2019], 1651fffdfffdfffd1686. URL: https:
          //projecteuclid.org/download/pdf_1/euclid.aos/1024691352.

[Str+07] C Strobl et al. "Bias in random forest variable importance mea-
          sures: Illustrations, sources and a solution." In: BMC Bioinformatics
          (2007). [Online; accessed 19-January-2019], pp. 8–25. URL: https:
          //bmcbioinformatics.biomedcentral.com/articles/10.1186/
          1471-2105-8-25.

[The06]   The_Economist. "Africa's hegemon". In: (2006). [Online; accessed
          04-January-2019]. URL: https://www.economist.com/special-
          report/2006/04/06/africas-hegemon.

[UND19]   UNDP. " Human Development Index (HDI)". In: (2019). [Online;
          accessed 04-January-2019]. URL: http://hdr.undp.org/en/
          content/human-development-index-hdi.

[Wor14]   WorldBank. "The World Bank In South Africa". In: (2014). [Online;
          accessed 04-January-2019]. URL: https://www.worldbank.org/en/
          country/southafrica/overview.

[Wor19a] WorldBank. " GDP per capita (constant 2010 US$)". In: (2019).
          [Online; accessed 04-January-2019]. URL: https://data.worldbank.
          org/indicator/NY.GDP.PCAP.KD.

[Wor19b] WorldBank. " Stocks Traded as a Total Value Percentage of GDP".
          In: (2019). [Online; accessed 05-January-2019]. URL: https://data.
          worldbank.org/indicator/CM.MKT.TRAD.GD.ZS?end=2017&
          start=1975.