4-28-2023

Vietnam's GDP: Re-assessing Growth Rate and Identifying an Alternative Indicator

My Linh D. Nguyen
University of Richmond, linh.nguyen@richmond.edu

Vietnam's GDP: Re-assessing Growth Rate and Identifying an Alternative Indicator

by

My Linh D. Nguyen

Honors Thesis

Submitted to:

Department of Economics
and Department of Mathematics & Statistics
University of Richmond
Richmond, VA

April 28, 2023

Advisors: Dr. Dean Croushore
Dr. Taylor Arnold

1 Introduction

Since the economic reform known as Doi Moi (Renovation) in 1986, Vietnam has changed from one of the world's poorest to a middle-income country in one generation (USAID, 2022). The country has consistently registered high and stable economic growth since the reform, averaging 6.3% from 1985 to 2021 (World Bank, 2022). High growth rate of gross domestic product (GDP) is good news, but it has also raised questions that go both ways. On one side, there is much speculation that the government of Vietnam has manipulated economic statistics, compared to the case of China and India. As quoted in Kinh Hoa (2017), Mr. Le, a Vietnamese economist, pointed out the discrepancies between the GDP growth rate and the growth rates of electricity use and net export, while Pesek (2019) raised doubts about Hanoi's GDP revision, claiming that it provided the government with perfect cover for increasing borrowing. Recently, Mr. Nguyen, former Director of the Central Institute for Economic Management, (quoted in Giang, 2021) questioned the country's 5.4% growth rate during the first six months of 2021 when many large cities and provinces were under social distancing and the majority of businesses closed. He argued that the consumption was lower than the same period in 2020 and net exports were negative. However, on the other side, evidences from the recent GDP revisions (Dinh, 2021) and the large and increasing informal economy (T. H. Nguyen, 2019; Cling, Razafindrakoto, & Rouband, 2011; Nghiem & Rouband, 2022) suggest that Vietnam's actual economic growth might be higher than reported.

The debate over whether economic figures are under-reported or over-reported is long-standing, especially among the Vietnamese public. It got more heated in 2021 when Vietnam's statistical reliability was again questioned after the country was accused of over-devaluing the Vietnam Dong to gain export advantages by the United States (US) Treasury (Dapice, 2020; Lawder & Lambert, 2021). Reliable statistics are the foundation for economic

analysis, which informs the actions of policymakers and businesses in Vietnam as well as its trading and foreign policy partners. Inaccurate statistics are problematic for all related parties and can undermine the whole economy if they lead to bad policy choices. Hence, it is important that this on-going debate is settled. However, there has not been any formal research done on this matter. This paper aims to fill that gap by answering the question of whether the Vietnamese economic statistics are accurately reported, and, if not, exploring the potential reasons. I focus primarily on GDP growth rate as it is not only the single best indicator of the economy's health but also the statistic that has been the subject of the most skepticism in recent years.

To answer the accuracy question, I evaluate the validity of reported GDP growth by using alternative indicators of economic activities, a method commonly used by financial analysts and scholars to evaluate the reliability of Chinese economic data, such as Rawski (1976), Le Keqiang (2007), and Koch-Weser (2013). I compare the growth rate of official real GDP with the growth rates of selected alternative indicators, including the number of air passengers, rail cargo volume, output of electricity, coal and cement, and the number of foreign direct investment (FDI) project licensed, from 1996 to 2021. Data availability is one reason for the choice of such indicators. However, the main reasons are that these indicators measure specific activities that closely track the growth of the economy and their units of measurement are unrelated to prices and currencies, which are prone to errors and manipulation.

The external validity check should provide some evidence on whether GDP growth rate accurately describes the real state of Vietnam's economic growth. My hypothesis is that Vietnam's GDP growth rate has been under-reported because official statistics do not adequately account for the country's informal economy (Thu Trang, 2019). This sector of the economy is not recognized in the country economic data because of difficulties with data collection. As my alternative indicators account for all economic activities, including informal

ones, it is reasonable that their growth rates are higher than the reported GDP growth rate.

Nonetheless, no single indicator on its own is particularly reliable due to influences from changes in the economy's structure and factors outside of the economy. Thus, the second part of this paper focuses on constructing an indicator that better captures economic growth. I use principal component analysis (PCA) to find the first principal components (PC1) of a potential set of individual indicators, consisting of growth rates of six alternative indicators and GDP growth. I construct 127 first principal components and compare them to the benchmark indicator, real imports. Imports is a suitable benchmark as it is highly correlated with true activities in the economy while it is not subject to systematic mismeasurement or misreporting[1]. For each PC1, I run a regression with imports growth as my dependent variable and independent variables being PC1 and the growth rate of the US dollar - Vietnam Dong annual exchange rate, controlling for non-activity channels that affect imports. The chosen indicator is the PC1 that is best fitted to real imports growth rate based on the coefficient of determination, R^2, and Bayesian/ Schwarz information criterion, BIC/ SIC.

The analysis reveals an inconclusive answer to the accuracy question with both under-reporting and over-reporting observed. On the other hand, the construction of an alternative index finds the combination of the number of air passenger and coal output to be the best in terms of goodness of fit. However, the six-indicator (all indicators except cement output)[2] can also be considered as an option. I also find that GDP is informative relative to other indicators and, together with the number of air passengers and coal output, can be the base for constructing another index in case more data and indicators are provided.

This paper is organized as follows. In section 2, I summarize the relevant literature. Section 3 details the data and methods used. Section 4 discusses the empirical results and section 5 concludes.

[1]For further explanation, readers may refer to Section 3.

[2]The combination consists of GDP, number of air passengers, rail freight, output of coal and electricity, and number of FDI projects licensed.

2 Literature Review

One available literature questioning the credibility of Vietnam's economic data is a report made by the United States General Accounting Office (GAO) in response to a request from the US Members of Congress Dana Rohrabacher and Zoe Lofgren in 1999. The GAO examined the availability, transparency, and quality of Vietnam's economic data and found that published economic statistics were subject to both underestimation and overestimation. In particularly, Vietnam's calculations of GDP failed to measure certain components of the economy, such as the large informal economy, small businesses, telecommunication, or the service sector. The report also criticizes the General Statistics Office of Vietnam (GSO) for having not published the methods used to collect and process economic and financial figures and identified potential data limitations or gaps. This issue has been partially addressed by the GSO as they published the methods used to collect, but not process, data in the Methodology section. On the other hand, it also points out that the quality of available data on Vietnam has improved over the years. Nonetheless, as this paper dates to 1999, its findings might still be relevant for some questions but they do not cover the most recent years of rapid growth. However, there has not been any research on this matter since, so this study is done with hope to fill in the lack of formal research on the credibility of Vietnam's economic statistics.

Even though there are no papers on the credibility of Vietnam's GDP, there are studies on the country's economic growth from other aspects related to my research. The empirical test results from H. H. Nguyen (2020) show that FDI has a positive and statistically significant influence on economic growth of Vietnam while T. X. Nguyen and Xing (2008) finds that a 1% increase in FDI inflows results in a 0.13% increase in Vietnam's exports, which, in turn, increases GDP. Studying the same topic on provincial level, Anwar and Nguyen (2010) finds a positive direct effect of FDI on economic growth in Vietnam, specifically for four regions:

Red River Delta, North East, South East, and Mekong River Delta. Another research done by Anwar (2016), in collaboration with Alexander, reveals a statistically significant long-run relationship amongst pollution, openness to trade, energy consumption, and real GDP in Vietnam.

Although to the best of my knowledge there are no papers that study the validity of Vietnam's national accounts data over the last 25 years, a number of papers examine the credibility and accuracy of the statistics of other countries. Magee and Doces (2015) and Martinez (2022) argue that authoritarian regimes manipulate economic statistics and overestimate GDP growth and both compare reported GDP data to the night-time lights recorded by satellites from outer space. Magee and Doces find that autocracies exaggerate annual GDP growth rates by $0.5 - 1.5\%$ while Martinez estimates an overstatement of as much as 35%. Using a similar method, Clark, Pinkovskiy, and Sala-i Martin (2017) finds underestimation instead of overestimation. They use satellite-recorded night-time lights to confirm the claim that China's economic statistics are reliable and finds China's GDP growth rates are not too far below the official reported statistics and might be considerably above them. Using night-time light is a common method to validate statistics. Henderson, Storeygard, and Weil (2012) develops a statistical framework to use satellite data on night lights to augment official income growth measures, which Pinovsky and Sala-i Martin (2016) follows to gauge growth in economic activity for a cross-section of countries. However, such method is scrutinized by Fernald, Hsu, and Spiegel (2019) to have massive measurement errors due to changes in the sensitivity and quality of satellites over time.

Fernald et al. (2019) is a part of a body of literature concerning the quality of China's economic figures, which has been an issue of concern among scholars for years. Results from this paper align with Clark et al. (2017), suggesting that Chinese statistics, including GDP, have become more reliable over time. Holz (2014) and Klein and Ozmucur (2002) have similar findings. Results from Holz (2014) suggest that there is no clear evidence of

data falsification and manipulations would be virtually undetectable. On the other hand, Klein and Ozmucur (2002) finds that movements of principal components of several strategic indicators suggested by basic social accounting principles, such as energy, labor, wage, and inflation, are consistent with the movements of China's official GDP. However, the authors warn that such results do not necessarily prove official measured GDP of China is correct, which aligns with most literature on this matter.

The majority of literature concerning China's statistics argue that Chinese data are not credible. Rawski was one of the first to question the reliability of Chinese figures, pointing out that the country's economic data suffer from substantial margin of error (Rawski, 1976). Many years later, Wang and Meng (2001) finds that China's industrial growth rate was over-reported by 4.5 percentage point in the 1990s and the GDP growth rate by 1.3 to 3.2% from 1978 to 1997. Similarly, results from Koch-Weser (2013) conclude that China's economic statistics are not as reliable as those of US and Europe with alternative indicators showing sharp drop unreflected in real data and huge divergences appearing among data reported by different levels of government. While other research span a variety of statistics, Groen and Nattinger (2020) focuses on official GDP growth rates, similar to this study, and finds that China's rates have been unrealistically smooth during the 2010s.

Besides China, another existing communist state whose economic data are in question is Cuba. In the country's volume in the series of country studies sponsored by the Department of the Army, Cuban economic statistics, especially GDP, are questioned for their reliability as Cuba had not been able to provide detailed statistics to support reported growth rates (Federal Research Division, Library of Congress, 2002). Pérez-López and Mesa-Lago (2019) agrees on this question and finds discontinuities and anomalies in Cuban GDP data. They find significant disparities in GDP growth rates from 1990 to 2000 that cannot be explained by official inflation and inconsistent changes from 2001 to 2007. One of the main reasons for Cuban poor statistics is the country's slow change in national account compiling methodology

from the Soviet's Material Product System (MPS) to the United Nations System of National Accounts (UNSNA)[3] (Pérez-López & Mesa-Lago, 2019; Pérez-López, 2020).

Nonetheless, studies on China are still a huge part of literature concerning countries' economics figures. Among these studies, there is a substantial amount of work done using alternative indicators, some of which I also use in this paper. Even Chinese senior government officials have acknowledged the usefulness of alternative indicators. In 2007, China's then Party Secretary of Liaoning and now premier Li Keqiang indicated electricity consumption, rail cargo volumes, and bank lending as his preferred measures of economic activity rather than the official GDP (Rabinovitch, 2012; Wagstyl, 2012). Among scholars, Rawski (2001) sets a precedent by using drops in airline travel and electricity consumption to conclude that China's official data is unrealistic, reflecting the government's objectives rather than actual economic outcomes. Despite having a relatively plausible conclusion, this paper lacks actual calculation as Rawski provides alternative figures for China's real GDP growth purely based on his guess. This problem was fixed by Wang and Meng (2001) in a similar attempt in the same year, calculating the physical output of 168 industrial commodities. Using a different approach, Groen and Nattinger (2020) puts proxy of imports, manufacturing activity to the imports and retail sales in sparse partial least square regression. In Klein and Ozmucur (2002) which finds contrary results to the last two papers, the authors use a set of 15 indicators, in which transport and energy overlap with this study. Energy also overlaps with Koch-Weser (2013) which uses electricity production as one of six alternative indicators in the same approach as the first part of this research, graphical comparison.

Besides China, alternative indicators are also used in literature concerning other countries. Marcellino and Sivec (2021) uses fuel sales data from petrol station, google trends,

[3]One major difference between the MPS and the SNA is that the MPS did not include market transactions, while the SNA considers both market and non-market activities. The MPS also had a stronger emphasis on meeting production targets set by the government, while the SNA allows for a more decentralized approach to economic activity, with market forces playing a greater role in determining the allocation of resources (United Nations Economic Commission for Europe Secretariat, 1981).

short-term state aid data and new car registration data to nowcast GDP growth of Luxembourg. Barr and sharp (2006) examines the revision of real GDP in South Africa and propose an indicator (a measure of real earning flows based on the Johannesburg Stock Exchange Industrial and Financial Index) that overcome the problems highlighted.

My work is most closely related to Fernald et al. (2019) which constructs an activity index alternative to China's GDP from alternative indicators. I take a similar approach in the second part of this research, using principal component analysis to construct a proxy for GDP growth rate and then, examining the principal components' fit to imports to determine which one is chosen. Principal component analysis is also used in Klein and Ozmucur (2002). However, instead of regressing the principal components against imports, the 2002 research runs a linear regression of percentage change in GDP on the principal components of annual change in several strategic indicators suggested by basic social accounting principles.

3 Data & Methods

3.1 Data

My data set consists of nine variables: GDP, six alternative indicators (including the number of passengers transported by air, rail cargo volume, coal output, electricity output, cement output, and the number of FDI projects licensed), imports, and exchange rate. I choose the time period to be from 1995 to 2021 on the basis of data availability. 1995 marked the end of the US's 30-year trade embargo against Vietnam and the beginning of Vietnam's consistently high GDP growth while 2021 is the latest year that all of the data are available[4]. All raw data are level data in annual terms and collected in October and November 2022.

The primary variable of interest in this study is real GDP with base year of 2015, which

[4]2022's data are available for only GDP, imports, and exchange rate.

is from the World Bank's Development Indicators (WDI). This series is not compiled from a Vietnamese source because the publicly available data by the General Statistics Office of Vietnam only goes back to 2009. The WDI is also the source for the official exchange rate between Vietnam Dong and US dollar (EXR), which is determined as an annual average based on monthly averages.

Another externally collected data is imports (IMP), which is retrieved from the International Monetary Fund (IMF)'s Direction of Trade Statistics (DOTS) in form of exports to Vietnam reported by trading partners[5]. This series is in millions of current US dollars and is deflated for analysis. As the DOTS reports imports in US dollars converted with market exchange rates, a dollar-based deflator is needed. The US's export price index fits this bill as it measures the change in monthly prices of goods and services of exports from the US using US dollars and is often used to deflate trade statistics. The annual average of the index is retrieved from FRED with 2015 as the base year to be consistent with real GDP.

The other six variables are collected from the General Statistics Office of Vietnam (GSO), including air passengers (AIR), railway freight carried (RAIL), electricity output (ELEC), cement output (CEM), coal output (COAL), and number of foreign direct investment (FDI) projects licensed (FDI). Apart from data availability[6], the reasons these indicators are chosen is that they measure very specific activities in the economy and use numbers that are unrelated to prices and currencies. The paragraphs that follow will define six indicators and explain why each of them is a good alternative indicator.

Air passengers is the total real number of passengers transported by air carriers registered in the country. It measures the actual number of passengers are on the plane at time of departure, rather than the number of tickets sold. This number also disregard the travel

[5]Assuming that all trade flows are accurately recorded and reported by both the exporting and importing countries and data follows the principle of double-entry accounting.

[6]Other alternative indicators, such as area of housing floors constructed and output of iron ore) are also available on GSO, but do not cover all of the years in the specified timeframe.

distance. That is there is no weight on air miles covered by the flight: 100 passengers are counted the same for a 389-mile flight from Hanoi to Da Nang and a 968-mile from Hanoi to Ho Chi Minh City. Number of air passengers transported is an indicator of Vietnam's level of economic activity as countries with large numbers of air passengers tend to be more integrated into the global economy and have stronger trade and tourism industries.

Railway freight is the total volume of cargoes transported by rail regardless of travel distance. It is calculated by the actual weight of goods carried, including packaging, and only measured after the completion of transportation to the destination as mentioned in contracts and finishing delivery procedure. As rail is often used to move goods over long distances, the volume of rail freight indicates the amount of goods produced and the level of industrial production and trade activity in the country.

Coal, electricity, and cement outputs are simply aggregate coal, electricity, and cement produced in the country within a year. They are all essential measure of a country's economic activity. Coal is a fossil fuel used extensively in power generation, steel production, and other industrial processes, making it a vital indicator of a country's energy sector and industrial production. Meanwhile, electricity is a fundamental input for economic activity, powering industries, households, and transportation. Thus, electricity output is also an useful gauge of the country's energy sector and industrial production. On the other hand, cement is a crucial building material and is utilized in construction projects such as buildings, bridges, and roads. Consequently, cement output offer valuable insights into the levels of construction activity taking place within the country.

The number of FDI projects licensed is the total number of investments made by foreign entities into either a company or a new project in Vietnam. This number is a crucial metric that measures the level of interest and confidence foreign investors have in Vietnam's economic potential and of the competitiveness of the country's business environment, such as tax policies, regulations, and infrastructure. Therefore, it provides substantial insights

into Vietnam's attractiveness to foreign investors and its level of economic development.

Table 1 presents summary statistics for all data describe above. Alongside unit of measurement, the mean and median values, standard deviation (SD), minimum and maximum values for each series are reported. GDP exhibits a positive (right) skew with mean GDP ($173,852.17$ million 2015 USD) much higher than median ($158,101.90$ million 2015 USD). The wide variation in GDP across the sample period is also notable. Other variables exhibit a wide range of variability. The number of air passengers, output in the electricity, coal, and cement industries, the number of licensed FDI projects, and real imports all have high variability. FDI stands out with a relatively high standard deviation compared to the mean, indicating a large variation in the number of projects across the years. On the other hand, there is a low variability in rail freight and the exchange rate. RAIL has a relatively balanced distribution with the mean value of $6,640.93$ thousand tons only slightly higher than the median of $6,525.90$ thousand tons. EXR has a narrow range, with a minimum of VND $11,032.58$ and a maximum of VND $23,208.37$ per one US dollar. Overall, the level of volatility observed in the data is not surprising and is typical of such data.

3.2 Empirical Methods

The analysis consists of two parts. The first part calculates the growth rates for all variables in the data set, and then, graphically compares the real GDP growth with each of the alternative indicators' growth rates. The comparison should provide a relative picture of the performance of Vietnam's official GDP growth rate. In the second part, an indicator alternative to GDP growth is constructed utilizing principal component analysis. The idea is to identify which first principal component of the seven indicators' growth rates (GDP and six alternative indicators) or combination thereof are particularly informative in terms of correlation with the benchmark indicator, imports.

To start the analysis, the year-on-year growth rates of all variables are calculated using the following formula

$$gX_t = \frac{X_t - X_{t-1}}{X_t} \times 100, \tag{1}$$

where gX_t is the growth rate of variable X in year t and X_t is the value of variable X in year t. Real GDP growth is, then, compared with the growth rates of the six alternative indicators by plotting the time series of GDP growth against each indicator. This comparison should provide some evidence in whether GDP growth rate accurately describe Vietnam's economic growth. Nevertheless, we must account for the fact that no single indicator on its own is particularly reliable. The performance of indicators can be influenced by factors other than economic growth and changes in the economy's structure. For instance, decrease in electricity consumption, which affect the demand of electricity output, can be attributed to gains in energy efficiency while the use of heavy industry inputs such as coal might be relatively less indicative of the overall economy as light industry and services have been assuming a larger share of GDP. In addition, these indicators are likely to correlate with GDP growth. Thus, a better indicator that capture economic growth more adequately should be considered.

Hence, the goal of the following part of the study is to identify such indicator. In order to accomplish this, a benchmark that is highly correlated with actual activity but is not subject to systematic mismeasurement or misreporting is required. I argue that imports best satisfy these requirements for three reasons. First, countries often accurately measure imports for tariff purposes. Second, data on imports are retrieved from external sources (IMF's Direction of Trade Statistics) as the sum of self-reported exports to Vietnam by trading partners. Third, Fernald et al. (2019) discover that import growth closely tracks GDP growth for countries with highly reliable statistical system.

To identify a better indicator of Vietnam's economic growth, principal component analysis is employed to find the first principal components of a potential sets of individual

indicators[7]. PCA is a common statistical technique to reduce a high-dimensional data set into a smaller set of dimensions while retaining many of the most crucial features of the original data (Abdi & Williams, 2010; Jolliffe & Cadima, 2016). It does this by creating new variables, known as principal components, that explains the most significant amount of variation in the base data. To compute principal components, PCA uses a matrix of data values representing the variables being analyzed. The first principal component (PC1) is calculated by finding the direction in the data that captures the most significant amount of variance, capturing and explaining the largest part of the data set's inertia (Abdi & Williams, 2010). This direction is found by identifying the eigenvector with the highest eigenvalue of the covariance matrix of the data (Abdi & Williams, 2010). Once PC1 is constructed, subsequent principal components are calculated by finding the directions that capture the most variation, subject to the constraint of being orthogonal to the previous principal components (Abdi & Williams, 2010). Each principal component is a linear combination of the original variables, with weights determined by the eigenvectors of the covariance matrix[8].

PCA is the chosen method because it enables the identification of the principal components capture the key common information in the indicators with first component explaining the maximum variation in the data set. This would help create an index that integrates many alternative indicators, resulting in a more informative picture of the actual GDP growth in Vietnam. It is probable that the alternative indicators' time series do not closely track Vietnamese imports because they narrowly focus on specific areas of the economy. However, the first principal components of subsets of indicators should be able to accomplish that. In addition, since GDP is meant to be the most comprehensive indication of economic activity, it is included as one of the indicators in this part of the study. By doing so, the variation

[7]This include seven indicators: GDP, number of air passengers (AIR), rail freight (RAIL), output of electricity (ELEC), cement (CEM) and coal (COAL), and number of FDI projects licensed (FDI).

[8]For a more in-depth understanding of PCA and how to compute principal components, readers may refer to Abdi and Williams (2010) and Jolliffe and Cadima (2016).

in GDP is accounted for in the measures, resulting in a more accurate representation of the overall economic growth in Vietnam.

The process starts with constructing the first principal component of all possible subsets of the growth rates of seven indicators, which involve considering a total of 127 combinations. At one extreme, seven combinations would consist of only a single indicator while the other extreme considers a combination of all indicators. For each subset, a regression of imports growth on PC1 is run, using exchange rate as the control variable. Thus, the baseline OLS specification is

$$gIMP_t = \beta_0 + \beta_1 PC1_t + \beta_2 gEXR_t + \epsilon_t. \tag{2}$$

$gIMP_t$ is the growth rate of imports, $PC1_t$ is the value of the first principal component from the year-on-year growth rates in the subset of indicators, $gEXR_t$ is the growth in the Vietnam Dong – US dollar exchange rate, and ϵ_t is the error term. After running the regression in 2, the coefficient of determination, R-squared or R^2, and the Bayesian/ Schwarz information criterion, BIC/ SIC, are reported to determine the PC1's goodness of fit to imports. The best index is one with the highest R^2 and the lowest BIC/ SIC.

The exchange rate is chosen as the control variable to control for non-activity factors that might create a change in imports. Theoretically, the demand for imports depends on domestic demand toward foreign goods and demand for imported intermediate inputs that are employed in production process and later, exported. If, for instance, the Vietnam Dong depreciated, then there would be a decline in demand for imported goods as they would become more expensive. Similarly, because imported inputs would be more costly, their demand would decrease as producers would switch to domestically produced intermediates that are now less expensive. Thus, I allow the regression to control for such non-activity channels that might alter imports by controlling for the exchange rate between Vietnam Dong and US dollars, which is normally used in trade.

4 Results & Discussion

4.1 Graphical comparison

Figure 1 and Table 2 tell a story from different perspectives about the comparison between GDP growth and the growth rates of six alternative indicators. Figure 1 shows plots of GDP growth in red against each alternative indicator's growth rate in blue. Meanwhile, Table 2 provides the summary statistics for all growth rates with similar descriptive statistics as Table 1, including mean, median, standard deviation, minimum and maximum. In general, there is much variability in the growth rates of alternative indicators compared to the relatively smooth GDP growth and the answer to the question of whether Vietnamese GDP is over-reported or under-reported in quite inconclusive.

One thing the graphs clearly show is that Vietnam's GDP growth rate has been impressively consistent prior to COVID-19, which is backed up by the standard error of only 0.96% from 1996 to 2019. During this period, the growth rate was consistently high around the range of 5 to 7% exception of the 1997 Asian Financial Crisis and the 2008 Global Financial Crisis. The 1997 crisis saw the country's GDP growth decline from 9.34% in 1996 to 8.15% and 5.76% in 1997 and 1998, respectively, while the growth rate decreased from 7.13% in 2007 to 5.66% in 2008 and 5.40% in 2009 during the Great Recession. After 2009, Vietnam's GDP growth slowed down to around 6% from 2010 to 2015. It started to pick up again in 2016 with growth rate of 6.69% and has continued to grow at a steady pace ever since. In 2020 and 2021, despite COVID-19 pandemic, Vietnam still managed to achieve growth rates of 2.94% and 2.59%, respectively, making the average 6.36% and increasing the standard error to 1.39% for the whole sample period.

It is also worth noting that among indicators, electricity output follows the trend of GDP the closest with the highest long-term correlation of 0.65. Both took a big dive during the 1997 Asian Financial Crisis following by a recovery in the early 2000s. The global pandemic

in 2020 also had a significantly negative impact on both indicators with both taking an identical drop in the year of 2020. Despite such sync, GDP and electricity output growths took different turns during the Great Recession with GDP having a gradual decrease while electricity output having a rapid recovery in the following years.

In comparison with all alternative indicators, GDP growth and indicators' growth rates have shown positive trends, but there are differences in the magnitude and pace of growth among series. The growth rates of air passengers, output of coal, electricity and cement, and the number of FDI projects licensed have been generally higher than the GDP growth rate while rail freight growth has been slightly lower, which have been captured by both Figure 1 and Table 2.

The differences in the size of growth rates, particularly before the COVID-19 pandemic, with that of GDP being lower, may suggest an under-reporting in GDP growth rate. This is also supported by the fact that some alternative indicators showed increased growth during period of declining GDP growth [9]. One possible explanation is that the informal sector[10] of the Vietnamese economy were underrepresented in official GDP statistics, leading to an underestimation of the overall size and growth of the economy. Several studies (Cling et al., 2011; T. H. Nguyen, 2019; Nghiem & Rouband, 2022) reveal that the informal sector is estimated to be anywhere from 15.6% to 32.7%[11] of the country's GDP, although it is unclear how much of this is already included in the national accounts. Cling et al. (2011) also finds that 82% of employment in Vietnam are informal employment[12].

[9] In 1997: rail freight, cement output and number of FDI projects licensed;
In 2009: rail freight, output of coal, electricity and cement, and number of FDI projects licensed;
In 2021: all except number of air passengers and cement output.

[10] The informal sector is defined as all private unincorporated enterprises that produce at least some of their goods and services for sale or barter, are not registered (no business licence) and are engaged in non-agricultural activities (Cling et al., 2011).

[11] 20% in Cling et al. (2011), 15.6 to 27.9% in T. H. Nguyen (2014), and 32.7% in Nghiem and Rouband (2022).

[12] Informal employment is defined as employment with no social security (social insurance). All employment in the informal sector is thus considered to be informal employment, as is part of the employment in the formal sector(Cling et al., 2011).

On the other hand, the disparities in pace of growth between GDP and six alternative indicators, especially during economic downturns, may indicate an over-reporting of Vietnam's GDP growth rate. In the event of the Asian Financial Crisis of 1997, similarly to GDP, the number of air passengers and output of coal and electricity all declined in 1997 while all alternative indicators dropped in 1998. Notably, all indicators experienced sharper drop in growth rates than that of GDP. The same was observed during the Great Recession where all indicators experienced a much steeper decline in growth rates than GDP. In addition, it is noteworthy that rail freight and coal output experienced a decrease in growth rates in 2007 and 2010, respectively, which were not reflected in real GDP growth. Similarly, in 2011, the growth rates of the number of air passengers, rail freight, electricity and cement output, and the number of licensed FDI projects decreased, but GDP growth rates remained unaffected. During the COVID era, a steeper decline was observed for all indicators in 2020 and those that continued to decrease in 2021, which includes the number of air passengers, and cement output. There is no definitive explanation for this and I also cannot say that the data has been manipulated under political pressure to maintain high levels of GDP growth as there is no hard evidence.

Therefore, the answer to the question of whether Vietnamese GDP is over-reported or under-reported in inconclusive. Even though all indicators follow the general positive trend of GDP, there are differences in the magnitude and pace of growth among series. The results indicate both under-reporting and over-reporting in the growth rate of GDP in Vietnam, suggesting limitations in data collection and reporting process. It is worth noting that Vietnam has made efforts in recent years to improve the accuracy and transparency of its economic statistics, including revising its GDP calculation methods in 2019 (United Nations Viet Nam, 2019). However, there is still room for improvement in terms of data quality and accuracy, particularly in measuring economic activity in the informal sector.

4.2 Identifying an alternative indicator

4.2.1 Relative performance of individual indicators

Table 3 gives the summary of the performance of individual indicators' PC1 in the regression in equation 2[13]. The coefficient of determination, R-squared, and the Bayesian/ Schwarz information criterion, BIC/ SIC, are reported. The table is arranged from best to worst in terms of R^2 and BIC/ SIC. Column 1 provides the ranking from the sorting and column 2 the abbreviation of the indicator while columns 3 and 4 give R^2 and BIC/ SIC, respectively.

My best performing individual indicator is the number of air passengers, AIR, with R^2 of 0.2137 and BIC/ SIC of 196.0537, outperforming other indicators but not by a lot. Following the number of air passengers are GDP and electricity output with R^2 of 0.1809 and 0.1793 and nearly equal BIC/ SIC of 197.1173 and 197.1673, respectively. Then, the midfield includes FDI, coal output and rail freight in that order. The worst performing indicator is output in the cement industry. Based on the individual performances, I would expect number of air passengers, GDP, and electricity output to be in the mix of the best performing combinations in the next subsection.

It is worth noting that the reported R^2 correlates with BIC/ SIC in that the higher R^2, the lower BIC/ SIC. Despite the goodness of fit, all of the individual indicators' estimates are statistically insignificant[14] with high standard error and p value. This proves the hypothesis that the individual standing on their own are not reliable.

[13] Other functional forms other than linear, such as quadratic, spline, and non-parametric, are also examined. However, linear regression still produces the best results and thus, is kept for final results.

[14] $H_0 : \beta_1 = 0$.

4.2.2 Performance of combinations of indicators

Table 4 shows the results for top 20 best performing combinations of indicators based on R squared and BIC/ SIC with the same columns as Table 3. Here, I observe the inverse correlation between R^2 and BIC/ SIC consistent with one observed in Table 3. Notice that coefficient estimates for top combinations of indicators, besides the last combination of GDP and numbers of air passengers and licensed FDI projects, are all statistically significant[15] with relatively small standard error and p value.

My best combination is a combination of two indicators: number of air passengers and coal output with R^2 of 0.2621 and BIC/ SIC of 194.4019. It is expected that number of air passengers is in the mix as it is the best performing individual indicator. Nevertheless, it is interesting that this best performing combination includes two indicators on two different economic activities: services for number of air passengers and energy production for coal output. Comparing to the best-performing individual indicator, number of air passengers alone, adding coal output into the mix improve R^2 by 0.0484 and BIC/ SIC by 1.6518.

However, adding more indicators does not necessarily mean improving the goodness of fit. The next-best performing combination adds GDP and number of FDI projects licensed into the mix, which decreases R^2 by 0.0119 and BIC/ SIC by 0.4186. The last one in the top three top performing indicator combinations swaps GDP out for rail freight and electricity output[16] and decreases R^2 to 0.2482 and increases BIC/ SIC to 149.8876.

Notice the combination of all indicators except for cement output[17] is ranked sixth on the list with six indicators and is the only six-indicator combination in top 20. Its R^2 is 0.2455, which is 0.0166 lower than that of the best combination, and its BIC/ SIC is 194.9829, only

[15] $H_0 : \beta_1 = 0$.

[16] It consists of number of air passengers, rail freight, number of FDI projects, and output in the coal and electricity industries

[17] Thus, it consists of GDP, number of air passengers, rail freight, output of coal and electricity, and number of FDI projects licensed.

0.581 worse than the best one. The trade-off in goodness of fit between the first-ranking combination and this one is not too big that it can be chosen as the desired index to avoid concerns about choosing indicators that just happened to fit well in the current sample.

Besides the first one, all other top combinations are combinations of at least three indicators, which institutes to the theory that it is essential to use a combination of indicators to get a more comprehensive understanding of the economy's performance. It is also interesting to see that all combinations in the top 20 include either number of air passengers or coal output with all of top 10 have both indicators. To be more specific, 90% include coal output while 80% consist of air passenger number.

The next most-seen indicator is, surprisingly, GDP, which appear in 14 out of 20 combinations. Combining with the fact that GDP is also one of the best fitted individual indicators, it proves that GDP is either informative on its own or adds additional information on economic activities that my alternative indicators do not address, such as retail or credit activity. This is reasonable and logical. As my alternative indicators by construction focuses on specific areas of Vietnam's economy, it is possible that the times series of Vietnam imports does not follow those of many of my alternative indicators exactly. GDP is supposed to be the broadest measure of the economy and, so by including GDP, its variation is included, making some differences to the explanatory power of my preferred index. It is also worth noting that in eight out of 14 times that GDP appears in the mix mentioned above, there is a present of the best performing combination, number of air passengers and coal output, and one of which is the combination of the three indicators themselves.

The complete list of indicator combinations arranged by R^2 and BIC/ SIC is in Table 5 in the Appendix. Note that the overall best sets of indicators fails to include cement output. Indeed, output in the cement industry is never chosen in a best-fitting combination that does not include all indicators. The all indicators set with cement output included is ranked 31 with a RMSE of 8.14. This is modestly worse than the six-indicator combination (all

indicators except cement output) which has a RMSE of approximately 8 and a ranking of six. Table 3 and Table 4 shows that not only does cement on its own is not as informative as other indicators, it also provides no additional information on activities that are not already in other variables.

Overall, the comparison of Tables 3, Table 4, and Table 5 highlight a number of results. First, the qualitative set of individual indicators that perform best is relatively stable. Indicators such as GDP, number of air passengers, and coal output seem to be ones that one would always want to include if further similar studies with more data are conducted. Second, GDP is relatively informative compared to alternative indicators. Finally, the six-indicator combination (all indicators except cement output) does well enough that one could rationally chose to use it, letting the data speak solely through the weights chosen in generating the principal component for the activity index. The main advantage of this combination is that, while it might not be optimal, it avoids concerns about choosing indicators that just happened to fit well in a particular sample. [18]

5 Conclusion

Vietnam's economy has undergone significant growth since the Doi Moi economic reform in 1986, with a consistently high GDP growth rate averaging 6.3% from 1985 to 2021. However, this growth has been subject to scrutiny and debate regarding the accuracy of reported economic statistics. Some experts suspect that the government manipulates economic figures for political reasons, while others argue that the actual growth might be higher than reported. This paper addresses this issue by aiming to answer the question of whether Vietnamese economic statistics, specifically GDP growth rate, are reported accurately.

[18] I also run the regression in equation 2 without the control variable, growth rate of exchange rate. Even though some of the rankings are altered, the best-performing individual indicator and the best-performing combination remain the same. The resulting R squared and BIC/ SIC are both lower.

In order to do so, I conduct external validity check by comparing the growth rate of official real GDP with the growth rates of six alternative indicators, including the number of air passengers, rail freight volume, output of electricity, coal and cement, and number of FDI project licensed. I plot the time series of GDP growth rate on each of the alternative indicators' growth rates and finds inconclusive answer to the aforementioned question. The results suggest both over-reporting and under-reporting in Vietnamese GDP growth rate which has been impressively smooth. This raises the concern of limitations in data collection and reporting process, urging enhancement of the quality and accuracy of economic statistics, particularly in accounting for the informal sector.

However, no single indicator is entirely reliable, so I propose constructing a better alternative indicator. The second part of the paper focuses on constructing this indicator using PCA. I find the PC1 of 127 combinations of growth rates of indicators, including six alternative indicators and GDP, and regress each PC1 on growth rate of the benchmark indicator, real imports, with growth rate of exchange rate as the control variable. R squared and BIC/ SIC are reported to assess the goodness of fit. My best-performing individual indicator is the number of air passengers while my best combination consists of number of air passengers and coal output. I recommend using the six-indicator combination (all indicators except cement output) as the desired index alongside my best-performing combination to avoid concerns about selecting indicators that only fit well in a particular time period.

Furthermore, I observe that GDP is informative relative to other indicators as it is one of the best-fitted individual indicators and appears in 70% of the top combinations. It also provides additional information on economic activities that my alternative indicators, by construction, do not address, such as retail or credit activity. Additionally, GDP consistently accompanies the number of air passengers and coal output in eight out of its 14 in the top combinations, suggesting that these indicators may be the ideal base for constructing another alternative index with more data and indicators provided.

To conclude, it is important to note several caveats regarding the data of this research. First, imports may not capture the full picture of economic activity as they may underweight certain activities such as services and other non-tradable sector. As such, an alternative approach for future research would be using exports as the benchmark indicator. It can be a good benchmark for similar reasons as imports. First, not only do exports represent the production and sale of domestically produced goods and services to foreign buyers, but they also reflect the demand for these goods, which in turn drives economic growth. Second, exports can also be retrieved from external sources as trading partners' reported imports from Vietnamese.

Second, the data set is not ideal. The problem with annual economic data is that it provides a limited picture of economic activity, as it cannot capture the nuances and fluctuations that occur within a year. In addition, having only one data point per year makes it difficult to compare and construct a new index accurately. Unfortunately, the General Statistics Office of Vietnam only reports the alternative indicators in annual terms, making it challenging to conduct more detailed analyses. Moreover, the set of alternative indicators used in the study lacks information on certain aspects of the economy, such as retails and credit activity. This lack of data is either due to a lack of historical data, especially before 2009, or a lack of the whole series. Therefore, the availability of quarterly data on alternative indicators and more diverse economic data being reported would not only benefit this research but also other studies on the Vietnamese economy.

Furthermore, the results from this paper open some other avenues for future research. One potential area of investigation is to delve further into the sources of data limitations and potential biases in the collection and reporting process of economic data in Vietnam. Another avenue is to explore the impact of the informal sector on the accuracy of economic data in Vietnam. The informal sector is known to play a significant role in the country's economy, but it is often difficult to measure and track due to its informal nature. Additionally, it

would be interesting to investigate the impacts of the misreporting on economic forecast and policy of Vietnam. Finally, there is a need for more research on the development of reliable and accurate alternative indicators of economic activity that can complement official GDP data for all countries in general and for Vietnam in particular. These avenues of research would contribute to a better understanding of the Vietnamese economy and help to inform policy decisions for sustainable economic growth.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *WIREs Computational Statistics*, *2*(4), 433-459.

Anwar, S., & Nguyen, L. P. (2010). Foreign direct investment and economic growth in vietnam. *Asia Pacific Business Review*, *16*(1-2), 183-202.

Barr, G., & sharp, L. (1996). The problem of revised GDP figures in south african alternative indicator of economic activity. *South African Journal of Economics*, *64*(4), 224-231.

Clark, H., Pinkovskiy, M., & Sala-i Martin, X. (2017). China's GDP growth may be understated [NBER Working Papers]. (23323).

Cling, J.-P., Razafindrakoto, M., & Rouband, F. (2011). The informal economy in vietnam. *International Labour Organization*.

Dapice, D. (2020, November 10). Debunking vietnam's currency manipulation.

Dinh, D. S. (2021, June 17). GDP cua viet nam phai vuot moc 500 ty USD [Vietnam's GDP must exceed the 500 billion USD landmark]. *Vietnam Net*.

Federal Research Division, Library of Congress. (2002). *Cuba: a country study* (4th ed.; R. A. Hudson, Ed.).

Fernald, J. G., Hsu, E., & Spiegel, M. M. (2021). Is china fudging its GDP figures? Evidence from trading partner data. *Journal of International Money and Finance*, *110*, 102262.

Giang, T. (2021, July 2). Thay gi sau buc tranh tang truong GDP [What is behind the high GDP growth rate]. *Vietnam Net*.

Groen, J. J. J., & Nattinger, M. (2020, October). Alternative indicators for Chinese economic activity using sparse PLS regression. *Economic Policy Review*, *26*(4), 39-68.

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *The American Economic Review*, *102*(2), 994–1028.

Holz, C. A. (2014). The quality of China's GDP statistics. *China Economic Review*, *30*,

309-338.

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2065), 20150202.

Kinh Hoa. (2017, November 6). Đang sau con so tang truong lac quan cua GDP [Behind the optimistic growth rate of GDP]. *Radio Free Asia*.

Klein, L. R., & Ozmucur, S. (2016). The estimation of China's economic growth rate. *Journal of Economic and Social Measurement*, *28*(4), 187-202.

Koch-Weser, I. N. (2013). The reliability of China's economic data: An analysis of national output. *US-China Economic and Security Review Commission Staff Research Project*.

Lawder, D., & Lambert, L. (2021, July 19). Vietnam pledges not to devalue currency in agreement with U.S. Treasury.

Magee, C. S. P., & Doces, J. A. (2015). Reconsidering regime type and growth: Lies, dictatorships, and statistics. *International Studies Quarterly*, *59*, 223-237.

Marcellino, M., & Sivec, V. (2021). Nowcasting GDP growth in a small open economy. *National Institute Economic Review*, *256*, 127–161.

Martinez, L. R. (2022). How much should we trust the dictator's GDP growth estimates? *Journal of Political Economy*, *130*(10), 2731-2769.

Nghiem, T., & Rouband, F. (2022). Measuring the non-observed economy in Vietnam: A focus on informal economy. *World Economics*, *23*(2), 1-26.

Nguyen, H. H. (2020). Impact of foreign direct investment and international trade on economic growth: Empirical study in Vietnam. *Journal of Asian Finance*, *7*(3), 323-331.

Nguyen, T. H. (2019). How large is Vietnam's informal economy? *Economic Affairs*, *39*(1), 81-100.

Nguyen, T. X., & Xing, Y. (2008). Foreign direct investment and exports: The experiences

of Vietnam. *Economics of Transition*, *16*(2), 183-197.

Pesek, W. (2019, September 2). Vietnam's surprise GDP revision risks damaging economic credibility. *Nikkei Asia*.

Pinovsky, M., & Sala-i Martin, X. (2016). Lights, camera, ... income! Illuminating the national accounts-household surveys debate. *The Quarterly Journal of Economics*, *131*(2), 579-631.

Pérez-López, J. (2020). Reflections on Cuba's national income statistics. *University of Pittsburgh Press*, *49*, 136-146.

Pérez-López, J., & Mesa-Lago, C. (2019). Cuban GDP statistics under the special period: Discontinuities, obfuscation, and puzzles. *Cuban in Transition*, *19*, 153-167.

Rabinovitch, S. (2012, July 13). China figures prompts doubts over accuracy. *Financial Times*.

Rawski, T. G. (1976). On the reliability of Chinese economic data. *The Journal of Development Studies*, *12*(4), 438-441.

Rawski, T. G. (2001). What is happening to China's GDP statistics? *China Economic Review*, *12*(4), 347-354.

Thu Trang. (2019, September 10). Dieu tra, đanh gia lai GDP đe thay ro buc tranh kinh te Viet Nam [Investigating, reassessing GDP to see the full picture of Vietnam's economy]. *The Ministry of Finance*.

United Nations Economic Commission for Europe Secretariat. (1981, Dec). Comparison of the System of National Accounts (SNA) and the Material Product System (MPS).

United Nations Viet Nam. (2019, Dec 13). Viet Nam launches revised GDP.

United States General Accounting Office. (1999). Vietnam economic data: Assessment of availability and quality.

USAID. (2022). Improving governance for economic growth.

Wagstyl, S. (2012, July 11). Measure for measure: Oil vs. electricity as a alternative for

chinese economic growth. *Financial Times*.

Wang, X., & Meng, L. (2001). A reevaluation of China's economic growth. *China Economic Review*, *12*(4), 338-346.

World Bank. (2022). World development indicators [Data set].

Tables and Figures

Series	Unit	Mean	Median	SD(*)	Min	Max
GDP	million 2015 USD	173,852.17	158,101.90	81,462.63	66,841.78	331,131.56
AIR	thousand people	15,923.78	10,200.00	15,378.43	2,435.00	55,079.56
RAIL	thousand tons	6,640.93	6,525.90	1,518.22	4,041.50	9,153.20
ELEC	million kilowatt-hour	96,968.40	70,960.00	73,854.98	14,665.00	244,864.00
CEM	thousand tons	46,851.34	40,009.00	32,726.13	5,828.00	114,685.98
COAL	thousand tons	31,669.75	38,778.00	14,308.10	8,350.00	48,307.65
FDI	projects	1375.70	1186.00	951.79	285.00	4028.00
IMP	million 2015 USD	106,481.77	75,851.27	95,099.90	11,854.64	300,508.46
EXR	VND per USD 1	17,695.34	16,302.25	3,921.35	11,032.58	23,208.37

Table 1: Summary statistics for base data
(*) *SD = Standard deviation*

Series	Mean	Median	SD	Min	Max
GDP	6.36	6.56	1.39	2.59	9.34
AIR	9.73	12.68	20.07	-53.11	44.64
RAIL	1.34	1.06	9.65	-22.33	21.61
COAL	7.71	8.94	12.55	-17.50	41.60
ELEC	11.48	12.25	3.13	3.51	17.00
CEM	12.43	10.50	8.08	-3.29	31.41
FDI	8.11	6.71	21.67	-35.20	56.43

Table 2: Summary statistics for growth rates of variables
(unit of measurement for all series is percentage)

Rank	Indicator	R squared	BIC/ SIC
1	AIR	0.2137	196.0537
2	GDP	0.1809	197.1173
3	ELEC	0.1793	197.1673
4	FDI	0.1709	197.4333
5	COAL	0.1619	197.7125
6	RAIL	0.1065	199.3779
7	CEM	0.0885	199.8958

Table 3: Relative performance of individual indicators

Rank	Indicators	R squared	BIC/ SIC
1	AIR, COAL	0.2621	194.4019
2	GDP, AIR, COAL, FDI	0.2502	194.8205
3	AIR, RAIL, COAL, ELEC , FDI	0.2482	194.8876
4	GDP, AIR, RAIL, COAL, FDI	0.2476	194.9101
5	AIR, RAIL, COAL, ELEC	0.2456	194.9796
6	GDP, AIR, RAIL, COAL, ELEC, FDI	0.2455	194.9829
7	AIR, COAL, ELEC, FDI	0.2453	194.9870
8	GDP, AIR, RAIL, COAL	0.2445	195.0152
9	GDP, AIR, COAL, ELEC, FDI	0.2441	195.0300
10	AIR, COAL, FDI	0.2422	195.0961
11	GDP, AIR, COAL	0.2406	195.1486
12	AIR, COAL, ELEC	0.2392	195.1963
13	GDP, RAIL, COAL, FDI	0.2375	195.2561
14	GDP, AIR, RAIL, COAL, ELEC	0.2373	195.2626
15	GDP, RAIL, COAL, ELEC, FDI	0.2354	195.3262
16	GDP, AIR, COAL, ELEC	0.2304	195.3752
17	GDP, COAL, FDI	0.2332	195.4002
18	GDP, AIR, ELEC, FDI	0.2314	195.4632
19	GDP, COAL, ELEC, FDI	0.2304	195.4967
20	GDP, AIR, FDI	0.2299	195.5119

Table 4: Top 20 best performing indicator combinations

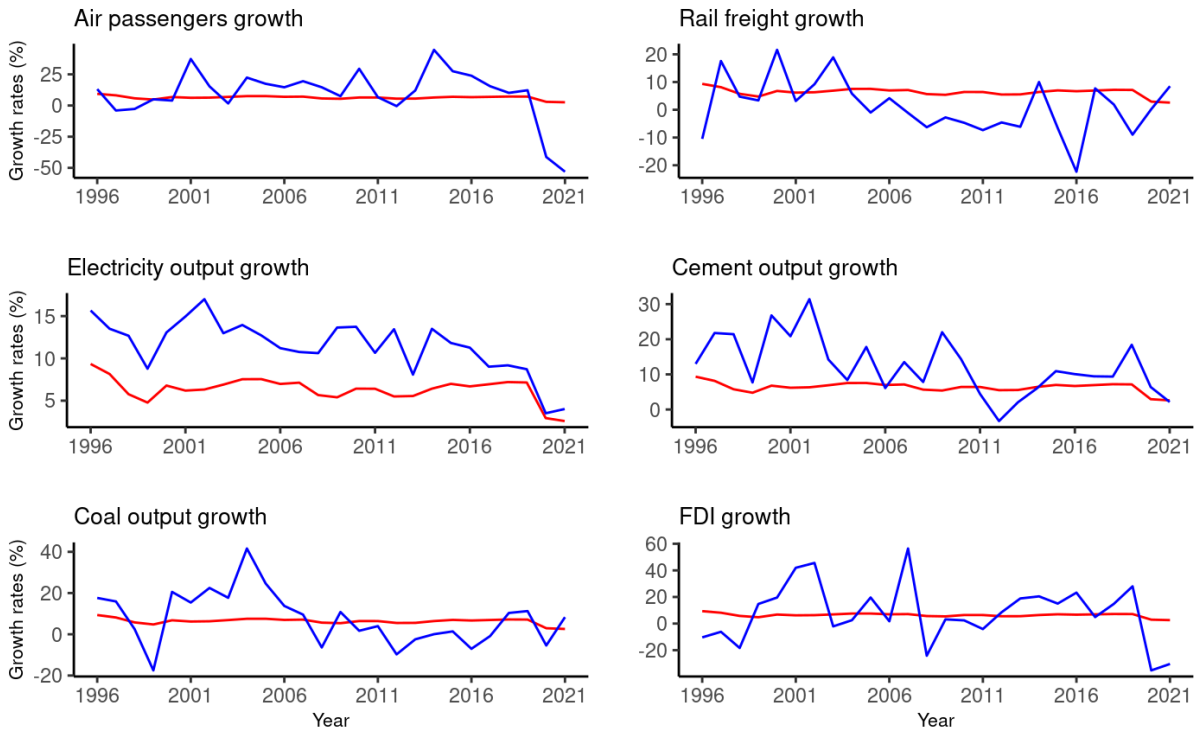

Figure 1: Graphical comparison of GDP growth
and growth rates of alternative indicators
The red series is GDP growth. The blue series is the alternative indicators' growth rate.

Appendix

Table 5: The complete list of indicator combinations arranged by R^2 and BIC/ SIC

Rank	**Indicators**	**R squared**	**BIC/ SIC**
1	AIR, COAL	0.2621	194.4019
2	GDP, AIR, COAL, FDI	0.2502	194.8205
3	AIR, RAIL, COAL, ELEC , FDI	0.2482	194.8876
4	GDP, AIR, RAIL, COAL, FDI	0.2476	194.9101
5	AIR, RAIL, COAL, ELEC	0.2456	194.9796
6	GDP, AIR, RAIL, COAL, ELEC, FDI	0.2455	194.9829
7	AIR, COAL, ELEC, FDI	0.2453	194.9870
8	GDP, AIR, RAIL, COAL	0.2445	195.0152
9	GDP, AIR, COAL, ELEC, FDI	0.2441	195.0300
10	AIR, COAL, FDI	0.2422	195.0961
11	GDP, AIR, COAL	0.2406	195.1486
12	AIR, COAL, ELEC	0.2392	195.1963
13	GDP, RAIL, COAL, FDI	0.2375	195.2561
14	GDP, AIR, RAIL, COAL, ELEC	0.2373	195.2626
15	GDP, RAIL, COAL, ELEC, FDI	0.2354	195.3262
16	GDP, AIR, COAL, ELEC	0.2304	195.3752
17	GDP, COAL, FDI	0.2332	195.4002
18	GDP, AIR, ELEC, FDI	0.2314	195.4632
19	GDP, COAL, ELEC, FDI	0.2304	195.4967
20	GDP, AIR, FDI	0.2299	195.5119
21	COAL, ELEC, FDI	0.2285	195.5611
22	RAIL, COAL, ELEC, FDI	0.2276	195.5898
23	GDP, AIR, RAIL, ELEC, FDI	0.2268	195.6188

Rank	Indicators	R squared	BIC/ SIC
24	AIR, ELEC, FDI	0.2256	195.6565
25	COAL, FDI	0.2233	195.7361
26	GDP, RAIL, ELEC, FDI	0.2218	195.7843
27	GDP, AIR, ELEC	0.2213	195.8005
28	AIR, RAIL, COAL, FDI	0.2208	195.8173
29	GDP, AIR	0.2205	195.8291
30	GDP, ELEC, FDI	0.2201	195.8416
31	GDP, AIR, RAIL, COAL, ELEC, CEM, FDI	0.2194	195.8644
32	GDP, AIR, RAIL, COAL, CEM, FDI	0.2194	195.8645
33	AIR, ELEC	0.2188	195.8854
34	AIR, RAIL, ELEC, FDI	0.2176	195.9261
35	GDP, AIR, COAL, ELEC, CEM, FDI	0.2172	195.9375
36	GDP, AIR, COAL, CEM, FDI	0.2172	195.9379
37	GDP, AIR, RAIL, ELEC	0.2167	195.9557
38	GDP, FDI	0.2159	195.9821
39	RAIL, ELEC, FDI	0.2154	195.9972
40	GDP, RAIL, COAL, ELEC	0.2153	196.0029
41	AIR	0.2137	196.0537
42	GDP, AIR, RAIL, FDI	0.2133	196.0687
43	AIR, FDI	0.2128	196.0839
44	GDP, AIR, RAIL, ELEC, CEM, FDI	0.2104	196.1624
45	GDP, COAL, ELEC	0.2100	196.1779
46	GDP, AIR, ELEC, CEM, FDI	0.2098	196.1844
47	AIR, RAIL, COAL, ELEC, CEM, FDI	0.2092	196.2042
48	ELEC, FDI	0.2080	196.2430
49	AIR, RAIL, ELEC	0.2078	196.2501
50	GDP, AIR, CEM, FDI	0.2074	196.2627

Rank	Indicators	R squared	BIC/ SIC
51	AIR, COAL, ELEC, CEM, FDI	0.2073	196.2645
52	GDP, AIR, RAIL, COAL, ELEC, CEM	0.2070	196.2767
53	GDP, AIR, RAIL, CEM, FDI	0.2054	196.3278
54	GDP, AIR, COAL, ELEC, CEM	0.2050	196.3418
55	GDP, RAIL, ELEC	0.2034	196.3928
56	GDP, AIR, RAIL, ELEC, CEM	0.2000	196.5021
57	GDP, AIR, COAL, CEM	0.1997	196.5134
58	GDP, COAL	0.1996	196.5163
59	AIR, COAL, CEM, FDI	0.1995	196.5195
60	AIR, RAIL, ELEC, CEM, FDI	0.1991	196.5337
61	GDP, RAIL, FDI	0.1988	196.5437
62	GDP, RAIL, COAL	0.1985	196.5516
63	GDP, AIR, ELEC, CEM	0.1984	196.5546
64	GDP, ELEC	0.1979	196.5708
65	COAL, ELEC	0.1977	196.5787
66	AIR, ELEC, CEM, FDI	0.1975	196.5835
67	GDP, AIR, RAIL, COAL, CEM	0.1973	196.5905
68	AIR, RAIL, COAL, CEM, FDI	0.1964	196.6210
69	GDP, AIR, CEM	0.1953	196.6554
70	GDP, AIR, RAIL, CEM	0.1945	196.6814
71	GDP, RAIL, COAL, ELEC, CEM, FDI	0.1937	196.7064
72	GDP, COAL, ELEC, CEM, FDI	0.1919	196.7665
73	GDP, AIR, RAIL	0.1907	196.8044
74	RAIL, COAL, ELEC	0.1903	196.8156
75	AIR, COAL, ELEC, CEM	0.1890	196.8579

Rank	Indicators	R squared	BIC/ SIC
76	AIR, CEM, FDI	0.1865	196.9375
77	AIR, RAIL, COAL, ELEC, CEM	0.1862	196.9489
78	GDP, RAIL, ELEC, CEM, FDI	0.1859	196.9576
79	AIR, RAIL, CEM, FDI	0.1851	196.9824
80	AIR, RAIL, ELEC, CEM	0.1833	197.0416
81	GDP, ELEC, CEM, FDI	0.1829	197.0540
82	AIR, RAIL, FDI	0.1819	197.0868
83	AIR, ELEC, CEM	0.1815	197.0991
84	GDP	0.1809	197.1173
85	GDP, COAL, CEM, FDI	0.1802	197.1379
86	GDP, RAIL, COAL, CEM, FDI	0.1794	197.1652
87	ELEC	0.1793	197.1673
88	GDP, RAIL, COAL, ELEC, CEM	0.1766	197.2531
89	RAIL, COAL, FDI	0.1759	197.2761
90	GDP, COAL, ELEC, CEM	0.1748	197.3096
91	RAIL, ELEC	0.1747	197.3124
92	RAIL, COAL, ELEC, CEM, FDI	0.1718	197.4051
93	COAL, ELEC, CEM, FDI	0.1715	197.4152
94	FDI	0.1709	197.4333
95	GDP, RAIL, CEM, FDI	0.1690	197.4911
96	GDP, CEM, FDI	0.1668	197.5621
97	GDP, RAIL, ELEC, CEM	0.1664	197.5732
98	GDP, ELEC, CEM	0.1630	197.6789
99	COAL	0.1619	197.7125
100	RAIL,ELEC, CEM, FDI	0.1597	197.7821

Rank	Indicators	R squared	BIC/ SIC
101	AIR, COAL, CEM	0.1570	197.8638
102	ELEC, CEM, FDI	0.1568	197.8716
103	AIR, CEM	0.1563	197.8865
104	GDP, COAL, CEM	0.1559	197.8976
105	GDP, RAIL, COAL, CEM	0.1551	197.9230
106	RAIL, COAL, ELEC, CEM	0.1508	198.0551
107	COAL, ELEC, CEM	0.1491	198.1079
108	COAL, CEM, FDI	0.1481	198.1376
109	RAIL, COAL	0.1480	198.1423
110	RAIL, COAL, CEM FDI	0.1458	198.2086
111	AIR, RAIL, COAL, CEM	0.1356	198.5164
112	GDP, CEM	0.1334	198.5840
113	RAIL, ELEC, CEM	0.1316	198.6368
114	GDP, RAIL, CEM	0.1304	198.6733
115	ELEC, CEM	0.1274	198.7613
116	CEM, FDI	0.1268	198.7804
117	RAIL, CEM, FDI	0.1265	198.7883
118	RAIL, COAL, CEM	0.1253	198.8245
119	AIR, RAIL, COAL	0.1199	198.9857
120	COAL, CEM	0.1196	198.9923
121	RAIL	0.1065	199.3779
122	AIR, RAIL	0.1057	199.4002
123	GDP, RAIL	0.1000	199.5649
124	RAIL, CEM	0.0995	199.5802
125	RAIL, FDI	0.0970	199.6536
126	AIR, RAIL, CEM	0.0953	199.7021
127	CEM	0.0885	199.8958