Exploring Unbalanced Growth Theory by Linear and Nonlinear Methods: Case of Indonesia

Andisheh Saliminezhad

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Economics

Eastern Mediterranean University
February 2018
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Assoc. Prof. Dr. Ali Hakan Ulusoy
Acting Director

I certify that this thesis satisfies the requirements as a thesis for the degree of Doctor of Philosophy in Economics.

Prof. Dr. Mehmet Balcılar
Chair, Department of Economics

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Doctor of Philosophy in Economics.

Prof. Dr. Fatma .G. Lisaniler
Supervisor

Examining Committee

---

1. Prof. Dr. Işıl Akgül ____________________
2. Prof. Dr. Fatma Doğruel ____________________
3. Prof. Dr. Fatma Güven Lisaniler ____________________
4. Assoc. Prof. Dr. Çağay Coşkuner ____________________
5. Assoc. Prof. Dr. Gülcay Tuna Payaslioğlu ____________________

ABSTRACT

This study seeks to test the Hirschman's theory of unbalanced growth and the viability of adopting a nonlinear model in Indonesia. It also studies the growth pattern in Indonesia by employing a suite of variable ranking algorithms to find the most significant leading sector during the study period from 1995 to 2015. To this end, an Input-Output framework is applied to detect the high linkage sector(s) (key sectors) of the Indonesian economy. Then the linear and nonlinear relationships between the extracted key sectors and GDP growth are covered with two different approaches specifically, Multiple Linear Regression and Multi-Layered Perceptron (MLP) Artificial Neural Network (ANN). Whereas detection of sector ranking is crucial for preparing a proper development plan; in the same vein, we apply two types of feature ranking methods (namely, Stepwise Regression and Ant Colony Optimization (ACO)-MLP based).

Empirical results from linear and non-linear models show that the effects of different sectors on growth in GDP in Indonesia are consistent with the structure of unbalanced growth theory. In general, we found that manufacturing sector is the most strategic sector in Indonesia since it has been selected both by linear and nonlinear forms as the first rank. Therefore, its development path firstly could be reinforced by more investment in this leading sector and then followed by investment in construction, hotels and restaurants, and agriculture.

Keywords: Economic Development, Neural Networks, Input Output Analyses, Feature Selection Methods

ÖZ

Bu çalışmada, Hirschman'nın unbalanced kalkınma teorisi Endonezya için nonlinear model kullanılarak test edilmiştir. Bu çalışmada, 1995 ile 2015 arasında Endonesya'nın kalkınma trendini değişken sıralama algoritmaları paketi kulanılarak incelenmiş ve kanlınma için en önemli sektörler belirlenmiştir. Son olarak, girdi çıktı taploları kullanarak Endonezya ekonomisi için en önemli sektörler belirlemiştir. Ekonomik kalkınma ile başlıca sektörler arasındaki doğrusal ve doğrusal olmayan ilişkiyi incelemek için iki farklı yaklaşım kulanılmıştır basit doğrusal regresyon ve MLP yapay sinir ağı modelleri. Beklenen kalkınma palnı hazırlanabilmesi için sektorlerlerin önem listesinin belirlenmesi çok önemlidir, bu doğrultuda, iki farklı ranking tekniği kulanılmıştır (isimleri, Stepwise Regression and Ant Colony Optimization (ACO).

Doğrusal ve doğrusal olmayan modellerin bilimsel sonuçları gösterimiştir ki Endonezyada farklı sektörlerin ekonomik kalkınmaya etkileri unbalance kalkınma theroisinin yapısı ile tutarlıdır. Genel olarak, imalat sektörü Endonesyadaki en stratejik sektör olarak seçilmiştir doğrusal ve doğrusal olmayan sıralama sonuçları sonrası. Böylelikle, ekonomik kalkınma ilk olarak bu sektöre yatırımdan geçmekte, diyer önemli sektörler sırasıyla, inşaat, hotel ve resotrant, ve tarımdır.

Anahtar Kelimeler: Ekonomik Kalkınma, Sinir Ağı, Girdi Çıktı Analizi, Özellikli Seçilmiş Modeller

To My Dear Mom

&

My Lovely Daughter

ACKNOWLEDGMENT

Firstly, I would like to express my deepest appreciation to my supervisor, Prof. Fatma Güven Lisaniler, who patiently guided me from the initial phases of this study. Her continuous supervision motivated me and put me back on track when I occasionally went astray, making this task an interesting learning experience. Without her encouragement and supports, it would not have been possible to complete this thesis in such a short time.

A special thanks must go to my parents for their love and support. Without their encouragement, I would have never given myself the chance to continue my education.

I would also like to thank Dr. Pejman Bahramian for everything he has done for me in the past 7 years and accompanying me in all the steps of my graduation path.

Finally, thanks to all the academic and administrative staff of the Economics Department of the Eastern Mediterranean University for their friendship and for providing good environment and facilities to study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

In decades after the World War II, development attracted a lot of attention and any given economy could take alternative routes to obviate this issue. In the theory of economic development, the leading sector investment strategies for achieving rapid growth as opposed to the relative efficacy of balanced is one of the longest and controversial debates of concern.

Therefore, this study focuses on investigating unbalanced growth theory of economic development proposed by Hirschman. It aims to scan the validity of unbalanced growth in Indonesia as a case study and answers to a question whether this theory is the best strategy to promote growth in Indonesia[1] as a LDC. According to the report of World Bank (2017), considering the pressing need by the Indonesian government for sectoral development and increased focus on improving the regulatory environment and stimulating sectoral spending, hence Indonesia was an interesting case for examining the unbalanced growth theory.

As respects, the application of unbalanced growth can be implemented using either linear or nonlinear models; in that regard the viability of adopting both linear and nonlinear models of development are also examined to define an accurate and the best fitted model of development for this country.

[1] With attention to some characteristics of the economy of this country, it was an interesting case to investigate the unbalanced growth theory. Some of the important features of Indonesian economy are available in Appendix C.

Moreover, since the inefficient sectoral rankings for specialization may mislead decision-makers in attaining a sustainable growth pattern and therefore may drive them to get an output which can be inferior to the current situation; the significance and priorities of growth determinants in Indonesia are determined by employing a suite of variable ranking algorithms.

Therefore this study contributes to the literature in the following ways:

I. Proposing a novel approach for examining the unbalanced growth theory.
II. First study on the investigation of the unbalanced growth model in Indonesia.
III. First study on the application of variable ranking methods in the context of investment priorities.
IV. First study in introducing ACO as a feature selection technique in economic applications.
V. First study which applies both linear and nonlinear models in economic growth framework and proves the outperformance of the nonlinear models.

The structure of the thesis is organized as follows: Chapter 2 provides the literature review. Chapter 3 describes data and methodology of the study. Chapter 4 presents and discusses the empirical findings and the last chapter deals with the conclusion.

Chapter 2

LITERATURE REVIEW

2.1 Theoretical Literature

Rosenstein-Rodan (1943), Nurkse (1953), and Scitovsky (1954) propose in their balanced growth theory that with attention to the existence of economic interrelationships and complementarities, all sectors of the economy should be developed simultaneously. It is also referred to how the vicious circle of poverty which is the characteristic of less developed economies can operate on both supply and demand sides of capital formation and how difficult is breaking the vicious circle.

Nurkse (1953) states that, by applying the balanced economic growth model, the vicious circle of poverty can be broken by investing in a large number of industries simultaneously. Moreover he points out that, it is the vicious circle operating in the less developed countries (LDC), which makes obstacle in the way of their economic development.

Obviously low demand for industrial goods leads to low profitability and so investment is discouraged which is the reason for operating the vicious circle of poverty on the demand side of capital formation. If the vicious circle can be broken subsequently economic development will occur afterwards.

Poverty in LDCs and thus low per capita income is an important factor which limits the demand and the size of the market. So for those of plant and machinery that only large scale production is economical and thus possible, the entrepreneurs are discouraged from investment. As a result, capital formation in the country is discouraged. Due to limited physical capital, productivity per worker is low, leading to low per capita income, which translates in to poverty. This is the way of the operation of the vicious circle of poverty in the LDCs.

In order to remove the poverty in such economies; the solution is simultaneous investment in the various sectors of the economy which can increase the aggregate demand due to the employment provided for a large group of people in the various industries to produce diverse commodities. So their income will increase and they will be able to afford the consumer goods made by one another, therefore the aggregate demand or in fact the size of the market will be expanded.

This is the way through which supply can build its own demand; as Say's Law[2] acknowledges in the strategy of balanced growth. If one industry expands then it helps the other industries boom and consequently the overall growth happens. Accordingly, the obstacle to economic growth derived from the limited demand or the small size of the market is cleared and the investment would further be induced more. Hence, in this way the trap of under-development equilibrium and following that the vicious circle of poverty can be broken through this path of balanced growth.

[2] "It is worthwhile to remark that a product is no sooner created than it, from that instant, affords a market for other products to the full extent of its own value (Say, 1803: pp.138–9)."

In LDCs people disburse their incomes mostly on food. Hence, investment in agriculture will be essential to meet demand for agricultural products and to boost balanced growth. If this circle is broken once then it will switch from poverty to balanced growth and to overall development of the economy since there is circular connection. In this way, the circle can benefit the economy but since LDCs' financial capital is limited, it is impossible to provide resources simultaneously for various sectors and to create an environment for massive investment and balanced growth.

In fact, Nurkse (1953) states that the small size of the market can cause difficulty which is related to individual investment stimulations. A number of industries may be unprofitable if taken separately so that the private profit can't motivate to induce investment in these industries. Nevertheless, a balanced increase in production would expand the size of each firm's market so that harmonized fulfillment would become beneficial. This pattern of capital investment in different industries is entitled by Nurkse as "balanced growth (1953: 56)".

Now there can be one thing to be questioned as which industries should be chosen for investment? Nurkse (1953) offers that investment should be made simultaneously in industries that consumers have higher demand or in those industries which they spend their income most. Only by exerting a coincident investment in a large number of complement industries, production or supply will make its own demand; so that the ones employed by them become each other's customers.

Another question arises that how we can trust that simultaneous investment is actually accomplished in a large number of industries? Nurkse (1953) answers that entrepreneurs and industrialists can be motivated to make investment simultaneously in different industries but if there aren't, then the government is responsible to implement the balanced growth itself. The government can make simultaneous investment in several industries. Therefore increasing the use of capital goods in large amounts will raise the level of productivity and increase people's incomes. Consequently, there will be an extensive increase in the aggregate output of consumer's goods and services. Ultimately the poverty of the people will be eliminated.

The two prominent American economists, Hans Singer (1958) and Albert Hirschman (1958), discussed about balanced growth doctrine of Nurkse (1953) and debated about the impossibility of that in LDCs. They claim that, a strategy of well-advised planned unbalanced growth is needed, not a balanced growth. According to Singer (1958), the problem of the LDCs neither can be solved by balanced growth, nor do they possess the requisite resources to attain balanced growth. Singer explains the theory of balanced growth for LDCs as: "how may hundred flowers grow whereas a single flower would wither away for lack of nourishment? Where are the resources to grow hundred flowers? ". (Singer 1958:132)

Singer believes that the basal trouble with the balanced growth doctrine is the problem of shortage of resources in LDCs. "The Think Big is an advice to LDCs, but Act Big is unwise guidance if it provokes them to bite more than they can possibly chew (Singer 1958:145)". The author also notes that, if the LDCs want to perform a large and varied package of industrial investment and ignore the agricultural productivity it may cause trouble. As the income goes up at the initial stages of development then the food and also agricultural sectors will face with relatively greater demand. Therefore, the agricultural productivity would have to be greatly raised in order to sustain industrial investment. So during the transition to industrialization, the country should implement a big push in industry coinciding with a big push in agriculture as well, not to involve in shortage of foodstuffs and agricultural raw materials. But when we are dealing with varied investment package for industry and agriculture at the same time, the capacity of LDCs to follow the balanced growth path would run into serious doubts.

Marcus Fleming (1955) states that the balanced growth doctrine assumes complementarity between the most industries, but the restriction of supply assures that the relationships are competitive for the most part. In the same vein, Singer (1958) adds that in fact a country possessing such resources would in fact not be less developed since the resources required for performing such policy are of such an order of magnitude. Any type of investment necessarily induces some additional investment and some other productive activities.

In the opinion of Singer (1958:52), “the expansion of social overhead capital and the growth of consumer goods industries and improvement of production techniques in them to raise productivity cannot take place simultaneously, because the LDCs have only limited capabilities of making use of their resources.” Moreover, in every LDC, there are some highly desirable programs of investment reflecting previous investments at any point of time which are not balanced investment packages but represent unbalanced investments to supplement existing imbalances. Singer (1958) proposes that in order to complement the available imbalance, the favorable investment programs always can be found in those countries that represent unbalanced investment. Such investments conduce to another balancing investment by creating a new imbalance. Consequently, the rate of growth in one sector will always be more than another one, and since investments must complement existing imbalance so the necessity for unbalanced growth will continue.

Singer (1964:143) asserts that “the balanced growth doctrine is premature rather than wrong and is applicable to a subsequent stage of self-sustained growth.” Like Singer, the most noteworthy opponent of the balanced growth school, Hirschman (1958:581) argues that “balanced growth theory requires huge amounts of precisely those abilities which have been identified as likely to be very limited in supply in the LDCs”. Hirschman characterizes the balanced growth doctrine as the application to underdevelopment of a therapy originally devised for an underemployment situation (Keynes, 1936).

After all, Hirschman acclaims that it was not balanced growth to propel many industrialized countries to where they are now. For instances, if one looks at the economy of the United States in 1950 and in 1850; will see the growth of many things, but will find different growth rates throughout the whole century (Hirschman, 1958).

Hirschman (1958) affirms that although the problem in LDCs is not only the absence of resources and the capabilities required running the balanced growth but that balanced growth itself cannot be favorable. He believes that if the country’s economic growth is to be quickened, it will be attained by unbalanced growth through creating imbalances in the economy which in turn will make incentives and pressures that will persuade development in the private sector.

As stated by Hirschman (1958), except the real scarcity in the resources in LDCs, they are also not able to bring them into play. He divides the initial investment into two related activities: one group of activities is so-called Directly Productive Activities (DPA), and the other is called Social Overhead Capital (SOC). With attention to this classification of activities; LDCs encounter with two alternatives to run the method of unbalanced growth; they can undertake initial investment either in SOC or the DPA.

Hirschman (1958) contends that both SOC and DPA cannot be expanded simultaneously because of the limited ability to utilize resources in most LDCs. Having stated this, the planning problem should focus on specifying the sequence of expansion that will maximize the induced decision-making. In this vein Hirschman (1958) recommends investment either in SOC or in DPA in order to create imbalances.

SOC includes all basic services; and the other productive activities such as primary, secondary and tertiary are dependent on them and cannot function without them. SOC investment comprises investment in education, public health, electricity, transport and communications, irrigation, drainage etc. The two paths of unbalanced growth can occur through two different frameworks. One pattern is development via shortage of SOC and the other path deals with development through excess of SOC. The intrinsic feature of the two directions is that they can result in additional investment and output (Gupta, 2009).

The first case takes place when the country creates the imbalances in the economy through investing in activities such as manufacturing, industries and construction which are DPAs. If investment in them is preferred as the main program then there will be need to expand SOC. If the country expands investment in DPA with no endeavor and considering expansion in SOC; then those industries will incur higher cost of production due to insufficient access to overhead facilities. In this situation, the government has to exert pressures and intervene in such that undertakes investment in SOC to create the infrastructure required for all-round development of the economy.

In the second case, SOC expands and so provides cheap inputs to agriculture and industry, therefore makes a country more attractive to DPA investors and induces investment in DPA. Since the major objective of the economy is to achieve increasing output of DPA, large investment in SOC will encourage investment in DPA by reducing the cost of services. According to Hirschman (1958:425) “development with excess SOC capacity is essentially permissive while that via shortages is an instance of disorderly compulsive sequence”.

What matters is that in LDCs, balanced growth of DPA and SOC is neither attainable nor an eligible policy since it does not make the incentives and the necessary pressure for induced investment decisions. Since poor countries cannot afford to be economical and they may lack of the ability to employ the resources, Hirschman assumes in his unbalanced growth that a country invests either in DPA or SOC.

When the progression of development happens from SOC to DPA, then this is interpreted as development through excess capacity. So when the second progression occurs from DPA to SOC, then this is named as development through shortages. If the first progression of development (excess capacity of SOC) happens; the sequence will be more continuous and smooth toward development comparing to the second progression (shortages of SOC). To quote Hirschman (1958: 123); "Investment in SOC is advocated not because of its direct effect on final output, but because it permits and in fact invites DPA to come in some SOC investment as a prerequisite of DPA investment."

Hirschman (1958) also underlines interdependencies and complementarities between sectors. As regards a LDC may not have sufficient resources to make massive investments in all sectors simultaneously, investing in one or a few key sectors could have the effect of pulling up other interdependent sectors. In the following Hirschman clarifies a precise theoretical formulation of his unbalanced growth strategy.

In LDCs some joint features like low levels of GNI per capita and slow GNI per capita growth, low levels of productivity, large income inequalities and widespread poverty, high rates of population growth and dependency burdens, high unemployment and underemployment; they all drive a country to encounter with rare resources or inadequate infrastructure to utilize these resources. Cash flows also cannot be diverted into different sectors that affect balanced economic growth due to the lack of investors and entrepreneurs. Hirschman (1958) asserts that intentionally unbalancing of the economy is the best method as the development policy, and its task is to conserve disproportions, tension and disequilibrium to keep the economy moving ahead. The protection of existing imbalances is preferred to the balanced growth considered as a goal. Therefore, an ideal pattern for development is precisely the sequence that leads away from equilibrium.

The foregoing discussion leads us to conclude that the LDCs should depend largely on judiciously-planned unbalanced growth to achieve rapid economic development and the balanced growth doctrine cannot be the solution for their path toward development.

The path of unbalanced growth can be described through three phases (Cornwall and Cornwall, 1994):

I. Complementarity: occurs when rise in production of one good or service will create demand for the other good or service. This demand will lead to imports of the second good or results in higher domestic production of the second product; otherwise rise in demand materializes as a political pressure.

II. Induced investment: this concept operates such as a multiplier, because a series of subsequent events can be triggered by each investment meaning that investment in one industry or sector can lead to stimulate investment in others through complementarity.

III. External economies: are created by preceding ventures and often appropriated by new projects which develop external economies that also may be employed by the latter ones. Hirschman states that the projects with a larger input of external economies toward the output, must be net beneficiaries of external economies (Hirschman, 1958).

Hirschman's theory of unbalanced growth is based on the following propositions (Hirschman, 1958:342):

A. Degree of complementarity is different in miscellaneous industries. Therefore, the formation of those industries where these complementarities happen to be the greatest should be the aim for the program of the economic development.

B. According to the program of balanced growth, income propagation may be achieved a once and later, the economy will be stabilized at a higher level with no further progress. The aim of development is that the process of increase in national income must continue year after year which can be required through creation and maintenance of deliberate imbalances in the economy.

Hirschman (1958) was the sole economist to present the idea of linkages in order to guide a deliberate strategy of development. Every investment project has both forward linkages and backward linkages and the role of development policy is detection of projects with the maximum total linkage. Investment in projects with maximum linkage affects the demand and supply positions in other projects in an economy which translates in to their expansion as well.

When investment in a specific project stimulates investment in latter phases of production, this process is creating the forward linkage which measures the degree of interdependence through the sales with the other industries as the buyers of the product of the given industry.

Conversely, the backward linkage is established when investment in a project induces investment in earlier projects which facilitate the process of the production of the given project and similarly can measure the interdependence through the purchases of that particular project from the products of the earlier ones. Since, actually projects create both forward and backward linkages therefore, the focus should be directed toward investment in projects with greater total number of linkages; and the awareness about project linkages can be prepared via input and output studies.

Altogether, to build an optimal roadmap for economic development strategies in LDCs; the importance of unbalanced growth theory is a crucial fact. In many LDCs, concerning particular challenges that make it difficult for them to stimulate and sustain economic growth; inducement mechanisms by key sectors can help to overcome the various obstacles to development.

2.2 Empirical Literature

The notion of unbalanced growth theory has led many scholars to focus on the validity of the theory as well as its prerequisites. In that regard, this chapter covers a summary of the most significant studies in this era.

First, we review the available studies which are related to this study either in theory or in terms of methodology used. We start with a review on inter-industry analysis for development since it helps to detect the most strategic sectors in a country. We also review the studies which focus to find the engine of growth in the economy which has a significant effect on growth comparing to the other sectors. In addition, we take a look at the available studies which investigate unbalanced growth theory either for testing or in modeling and further we briefly review the studies which prove the outperformance of nonlinear methods over its linear counterparts.

2.2.1 Inter-Industry Analysis

Since we applied input-output tables in this study so we firstly try to introduce it and give a short literature in this part.

After 1941 when the primal I-O tables were introduced for the American economy by Leontief, the input-output analysis became as vital and inherent tool for examining various aspects on mutual intertwinements of sectors in an economy.

Hence, early Rasmussen (1958) and, Chenery (1960) began to use the input-output tables for appointing the linkages between sectors of the economic system. The linkages were perused on the supply side to individual sectors where through backward linkages are correlated to the inputs, as well as on the side of demand of an individual sector to other sectors which through forward linkages affect the side of outputs (Pfajfar and Lotric Dolinar, 2000).

According to Chenery (1960) the coordination of investment plans in interrelated sectors of an economy necessitates the use of inter-industry framework for development policy. He firstly applied the Static Leontief Model in order to calculate the direct linkage effects with Rasmussen's indirect linkage effects.

Chenery (1960:36) believes that the inter-industry methodology is still in an experimental phase. As maintained by him, "the accumulation of input-output data over time and the more systematic exploitation of technological information is the most needed, especially for new types of production."

Kaldor (1986) specified the effect of the manufacturing sector's production on the overall growth of an economy. He underlined that due to the interaction between all the products and sectors of the economy, a permanent and stable economic growth will require the establishment of new and complementary products and sectors.

Leontief (1947) thinks that the world economy can be visualized as the economy of a single country with a system of interdependent processes through which certain outputs are generated and a specific combination of inputs are absorbed. Whenever the output of one becomes an input of the other, it leads to the direct interdependence between the two processes. The structure of a two-way input-output table can be instrumental to show the flows of goods and services among multiple sectors and can easily describe the state of a specific economic system.

2.2.2 Sectoral Effect on GDP

Since this study tries to detect the relationship between the sectors and growth in order to find the engine sector therefore, in this regards we try to mention some studies which deal with finding the engine sector among the main sectors in the economy.

The link between the expansion of services sector and economic growth has been tested by Dutt and Lee in 1993. By using the regression analysis between the growth rate of GDP and share of services in employment, they suggested that relative rise in the services share in employment is in line with a fall in the output growth rate.

Kaldor (1986) accomplished the empirical research on the sectoral effect on economic growth. He detected positive and statistically significant correlation between the growth rate of the manufacturing sector and growth rate of output in UK.

In another study, Necmi (1999) investigated Kaldor's law (on manufacturing as an engine of growth) in developing countries. He used the cross country data covering the period 1960-1994. The results indicated that growth rate of manufacturing output is exogenous as Kaldor claimed, and also adoptable by most of the developing world.

Wilber (2002) examined the nexus between service sector enlargement and growth rate of output. He used panel data for 25 OECD countries from 1960 to 1994. He discovered that causal link exists from services to growth. In this way the relative expansion of the whole service sector was associated with a decline in the growth rate of total output. Whereas disaggregated analysis showed that consumer and government services have a negative effect on growth, however producer services have positive impact.

Singh et al (2005) examined whether service sector will be the new engine of economic growth in India. They estimated six different simple linear growth equations and found that all the equations indicate high correlation between sectoral and overall growth. Unfortunately, only four of those equations satisfactorily passed the various diagnostic tests relating to manufacturing and service sector. They also found high correlation between agricultural and GDP growth rates, but not correlated as high as the manufacturing sector.

Linden et al (2007) analyzed how sectors' shares are related to economic growth. They used panel data for Schengen region for the period of 1970-2004 to investigate the long run nexus between sectors' share and economic growth. By employing a dynamic econometric modeling they estimated the product function and also used co-integration and error correction model in addition to conducting the Granger

Causality test in panel setting. They discovered that turbulence in sector share relationships has risen out of the shocks in industry sector share which slowly corrects to equilibrium. In addition, the Granger Causality test results revealed the existence of unidirectional causality from the growth of GDP per capita to agriculture share growth while a bi-directional causality detected between growth rate of GDP per capita and industry share growth. Similarly, between services share growth and growth rate of GDP per capita was also a bi-directional relationship. In general, their findings depicted that even in the presence of a complex link between GDP per capita growth rate, agriculture and service shares still the engine of economic growth was the industry sector.

Obasan et al (2010) examined the effect of industrial sector on economic development in Nigeria. They adopted an endogenous growth model estimated by Ordinary Least Square (OLS) method. The real gross domestic product (RGDP) was the dependent variable of their specified model and the independent variables (as exogenous variables) were exchange rate, inflation rate, interest rate, government expenditure and manufacturing output was included as a proxy for industrial sectors. They discovered that except the exchange rate and government expenditure there is positive relationship between the RGDP and the other three exogenous variables.

Hussain and Khan (2011) for the purpose of ascertaining the role of agriculture in the process of economic growth in Pakistan empirically examined the relationship between GDP growth rate and agriculture during the period of 1961–2007. They specified a linear growth equation with GDP growth rate as dependent variable and the agriculture growth rate as the independent variable, and employed the OLS estimation method. The results showed that agricultural growth rate and GDP growth rate were positively related. Therefore, they concluded that structural changes in the agricultural sector should be made by the government in order to ensure that agriculture leads to overall growth in Pakistan.

Enu et al (2013) examined the contribution of the agricultural, service and industrial sectors in the achievement of higher economic growth in Ghana. The OLS estimation method was used for the analysis over the period 1966 to 2011. They discovered the positive relationship between these sectors and GDP growth whereas the results showed that the agriculture sector contributed most to the overall growth.

2.2.3 Applications of the Unbalanced Growth Hypothesis

Streeten (1963: 670) states that "insofar as unbalance growth model does create desirable attitudes, the crucial question is not whether to create unbalance, but what is the optimum degree of unbalance, where to unbalance and how much in order to accelerate growth; which are the 'growing points', where should the spearheads be thrust, on which slope would snowballs grow into avalanches?"

The applicability of Hirschman's hypothesis was detected by Yotopoulos and Nugent's study where they tested unbalanced growth hypothesis using quantitative analysis that connects linkage coefficients to economic growth, in a cross-country analysis. They picked six developed countries and five LDCs and consolidated the input-output (I-O) tables of all into one I-O table for developing countries and one for LDCs. They examined the relationship between linkage coefficients and growth rate by using the "Hirschman-compliance index"[3]. The results rejected "unbalanced growth hypothesis; countries which emphasized high-linkage sectors were able to

[3] An Index which covers the correlation between sectoral growth rates and their total linkage index Yotopoulos & Nugent (1973).

achieve higher growth rates than countries that emphasized low-linkage sectors (1973:124)". The study confirmed a balanced-growth version of the linkage hypothesis.

At the late 1980s and early 1990s, formal models of unbalanced growth hypothesis were introduced by the work of Murphy et al. (1989). The paper depicted that under what conditions, the balanced growth doctrine might be applied. They demonstrated that to ensure the occurrence of balanced growth; the industrializing firm must be capable to impress aggregate income not only via its profits but also other factors are significant. If the only way to affect aggregate income by firm's industrialization is profit then this is not possible. Therefore Murphy et al looked at three ways in their paper in which aggregate income may be affected by firm's industrialization; sectoral wage differentials, the inter-temporal model that varies income between periods and the declines in cost via infrastructure investment.

Their model is not fitted to explore the ideas lied behind unbalanced growth. All firms are considered similar and the only question is that whether the industrialization happens in one firm then it will possibly generate more demand for other firms via a rise in aggregate income.

As opposed to Murphy et al (1989), the model presented by Krishna and Pérez (2005) is structured to illuminate these key questions. Their aim has been to formalize the concept of unbalanced growth in a simple way. They design an economy with a hierarchy of goods in which only the good at the top is consumed and each good only uses the good below. This vertical structure of the model is causing to reduce the price of the final good and so augments its demand. Hence, this formulation of the model affects the demand for all intermediates and pushes them to rise and propels the profits upwards. Linkages describe the conditions by which profitability of sectors above and below a sector is affected with industrialization by that particular sector.

The model presented by Krishna and Pérez conforms to the complementarities and cumulative processes in the literature. When a sector invests and acquires a better technology then it leads to decline in the price of the consumption good and consequently a corresponding rise in demand for other sectors' output which adds to the profitability of their potential investments. Since the new technology is associated with fixed costs no firm finds it profitable to invest independently. If the sector cannot benefit from investment complementarities, so an underdevelopment trap of low demand -no investment -low demand emerges and remains as a vicious circle. The unbalanced growth approach is basically turning a vicious circle into a virtuous one. If industrialization takes place in the sector having the most extensive positive effect on other sectors' investment decision then it will be the best efficient way to reach this critical mass. This sector is defined as the leading sector in the system. They found that the leading sector gravitates to be the sector with the utmost upstream linkages when each other's outputs are used by sectors as inputs relatively intensive. The converse case happens when the leading sector is downstream implying the low intensity use of other's output.

Although the studies on unbalanced growth hypothesis are limited, recently it has resurfaced in studies ranging from a U-shaped pattern of sectoral diversification during the development process (Imbs and Wacziarg, 2003) to industrial policy in the literature.

For instance Cohen's study in 2007, distinguishes between horizontal (framework) industrial policy including many or all sectors, and vertical (sector-specific) industrial policy which aims at particular sectors (details are available in Livesey, (2012)). He asserted that unbalanced growth hypothesis comports squarely into the latter category.

Hausmann et al's study in 2008, targeted how to conduct an industrial policy in a low- to middle-income country. Their manifest classified an industrial policy into two different groups; namely small and large. In the small class, the aim is to improve the performance of existing industries through identification of the available roadblocks, whilst the large category implies on the prioritization of the sectoral policy. For the latter case, the linkage component of the unbalanced growth hypothesis would prepare an examination for making the optimal sectoral priorities.

However it is noteworthy to stress that, while industrial policy focuses on optimizing growth, Hirschman's (1958) concern is rather about entering a growth trajectory into the development process at an early stage. This point also appears in recent studies. For example Rodrik (2010) states that "good industrial policy attempts to enhance the relative profitability of non-traditional products that face large information externalities or coordination failures, or which suffer particularly strongly from the poor institutional environment (p. 93)". Hirschman (1958) might have found that large information externalities are faced by all products associated with industrialization. Whereas Rodrik (2010) recommends that "the government is only focused on providing complementary inputs to the market (p. 25)", Hirschman considers a bigger role for the government which comprises involvement in directly productive activities.

Holz (2011) considers linkage coefficients in the basis of state involvement in the economy and measures the potential nature of profit opportunities. He focuses on the test of two specific components of the unbalanced growth theory in China. First, he studies the impact of linkage effects on economic growth and second, the extent that the Chinese government considers these effects in decision makings about the distribution of state ownership across sectors. In order to achieve these, the linkage coefficients are calculated in the first step. Second, linkage coefficients are correlated with ownership data across sectors and lastly, economic growth is related to linkage coefficients and ownership data. While in earlier literature there was no achievement to support Hirschman's unbalanced growth hypothesis through quantification of linkage effects however Holz unambiguously confirms the unbalanced growth hypothesis. Holz proved that in case of greater degree of linkage in a Chinese province, that province will have more rapid economic growth and profit linkage is playing a significant role rather than the output linkage. The evidence suggests that "however the degree of linkage matters in generating economic growth in China; province-specific withdrawal strategies for the state sector have no effect on economic growth (2011:46)."

Widodo et al. in 2014, in the context of sectoral specialization investigated the effect of agglomeration economies on productivity growth of manufacturing industries in Indonesia during the 2000s. Their findings supported a positive specialization effect and negative diversity effect on productivity growth in manufacturing.

Chapter 3

DATA AND METHODOLOGY

3.1 Data Description

We combined different datasets to derive the largest dataset concerning the unbalanced growth model. The samples for input-output tables used in the computation of the linkage coefficients are annual data from the World Input-Output database covering the period 1995-2014[4]. These tables are constructed for 35 sectors in a clear conceptual framework based on officially published input-output tables in conjunction with national accounts and international trade statistics. The updated version of the database for GDP per capita[5], import and export values (US dollar), and sectoral investment shares (percentage of GDP) of the key sectors (the ones extracted from input-output tables) are drawn by DataStream database from 1995 to 2015 on a monthly basis.

3.2 Methodology

This study combines different methodologies to test the validity of the unbalanced growth theory in Indonesia. First, we identified the high linkage sector(s) of the Indonesian economy by means of Input-Output framework [backward and forward linkages (Hirschman-Rasmussen, 1958)] for the period 1995-2014[6]. Then two

[4] Available at http://www.wiod.org/home

[5] GDP per capita is used in constant USD dollar (2005).

[6] It is noteworthy to mention that, after defining the key sectors on an annual basis, in order to highlight the effect of the leading sector on GDP growth and provide a comprehensive point of view, the monthly frequencies of sectoral investment shares for the extracted key sectors have been applied.

different approaches for capturing the linear and nonlinear relationships between the extracted key sectors and GDP growth were examined.

For the linear part, we used a linear regression as a common statistical data analysis technique. While, in order to capture the nonlinear phenomenon of economic growth in Indonesia, this study considered a Multi-Layered Perceptron (MLP) Artificial Neural Network (ANN) as a non-parametric device which does not assume any particular parametric form. Second, in order to perform a rank-order analyses for detecting the most significant leading sector between all the extracted strategic sectors throughout the study period; two groups of feature selection methods (namely, Stepwise Regression and Ant Colony Optimization (ACO)-MLP based) where investigated.

3.2.1 Input-Output Analysis (Hirschman-Rasmussen Index)

This study detected the yearly key sector in Indonesia by using input-output tables via the Hirschman Rasmussen index. Analyses were carried out with MATLAB for 35 different sectors in the period 1995-2014.

Input-output analysis pioneered by Wassily Leontief (1906-1999) is one of the techniques that can be used to investigate the magnitudes of bi-directional dependencies. As an analytical model, it interprets the interrelationships that exist in an economy in terms of growth linkages in a simple and meaningful way (Roux, 1991).

Input-output table constructs a model to collect, categorize and analyze data based on the sectoral structure, meantime has the capability to examine the interdependences between the various sectors. Hence, it has been applied extensively in multiple forms of economic effect measurements so far.

Input-output analysis represents sectoral input and output linkages through measuring inter-relationships between various industries. It is noteworthy that input-output analysis can be considered as one of the main contributions to economics in the 20th century in which theory, data and application can support one another.

3.2.1.1 The Concept of Economic Growth Linkages

As mentioned in Kristiansen (2003), the foremost primary ideas about the conceptual applications of growth linkage model are belonged to Hirschman's and Perroux's works (1958, 1955). From Perroux's point of view, the theory of linkages was substantially dealt with the establishment of industries with the ability of driving the economic development. Later, Hirschman (1958) introduced the backward and forward linkages effects through which developed the linkage concept. Also, Hirschman (1956) displayed that backward and forward linkages detected through input-output analysis, operate as investment forces due to providing an economy with growth momentum through their chain effects.

An input-output table comprises of value added sections and final demand. In the economic system, the process of sale of inputs takes place from the buyer producer to the industrial sectors. Production is for the households, government and international trade which constitute the final demand sector. These units' demand and the magnitude of their purchases from each industrial sector are generally unrelated to the amount produced in the other sectors (Miller and Blair, 2009).

An industry production in the framework of the input-output model has increased demand and supply effects on latter industries in the system; when industry *i* increases its production this leads to requesting more demand for inputs from other industries, furthermore an increase in production by other industries results in additional output required from industry *i* to supply inputs to meet the increased demand.

In the input-output model, the demand in the former case is referred to as backward linkage and the supply function in the latter case is referred to as forward linkage. When the expansion of an industry causes other induced activities and thereby makes more benefit for the economy than other industries; that industry has higher backward linkages than other industries. Similarly, an industry with relatively more sensitive production to changes in other industries' output has higher forward linkages than other industries.

After introducing the classical concept of linkages by Hirschman (1958) in regional development economics, he was willing to understand how economic system grows and sustains itself overtime (Drejer, 2003). Hirschman noticed that when activities in other sectors stimulate others to involve in new economic activities; linkage is in operation. He also understood that interdependencies between sectors or firms rise out of their supply of and demand for intermediate products.

Heretofore Rasmussen (1956) had proposed how to measure the forward linkages and backward indices. The total of backward and forward linkages build the net output of the industry multiplied by the probability of building each supplying industry because of the additional activities of one industry or a group of industries (Drejer, 2003).

Nevertheless, if the markets for other industries as the starting consequence do not rely on the supplies of the industries whose activities have increased, it is not adequate to define forward linkages (Drejer, 2003).

Since linkage effects are operational, the Hirschman's model remains dynamic and the existent industries drive the growth of the other ones in the system; implying that economic systems with strong casual linkage effects and a high degree of interrelatedness are more dynamic (Drejer, 2003).

3.2.1.2 Key Sector

Hirschman (1958) endeavored to realize the main sources of development and economic growth in the long run. He pointed that the country should build its economic planning with attention and identification of vital economic sectors which prepare incentives for growth and efficient resource allocation.

The evidence has shown that the interaction of the firms determines their process of entry and exit and economies are induced by innovative and adaptable firms. Consequently, this has been considered as the source of long-term growth of productivity. Such a scheme at the sectoral level is denoted by key sectors through which the economy is driven to increase the interdependencies and the levels of income (Cuello and Mansouri, 1992).

Clearly, the necessity of the concept of the key sector is associated with the unbalanced development notion. Hirschman (1958) reasons that the entire economy is driven on the path of efficient growth through the unbalanced development of main final demand sectors like that of a competitive economy. According to Amores (2012), the countries like Taiwan, Japan and South Korea which implemented Hirschman's strategy could achieve the highest prosperity in their development path.

The backward linkage indicator (BL) measures changing effect of the final demand in a specific sector on the total production of the economy, whereas the forward linkage indicator (FL) analyses if there is a global change in the final demand of all sectors then what will be the effect on the production of a specific sector. These indicators make the determination of the key sectors possible in an economy. Due to generating a high multiplier and fostering effect of the key sectors on production, they can help for defining the development strategies as part of the economic policy.

3.2.1.3 Open Leontief Model

To illustrate, consider an input-output economic system with *n* inputs and *n* final demand (*F*). The sum of the demand for inputs (x_{ij}) and final demand (F_i) gives the total output (x_i) for sector *i*. Hence,

$$X_i = \sum_{j=1}^{n} x_{ij} + F_i \qquad i = 1,2,.....n \qquad (1)$$

To assume input a_{ij} as a coefficient for input ij that contributes to the total input i, is needed to generate one unit *j* insofar as for producing *X* products, an industry needs $a_{ij}X_j$ of input i, where x_{ij} represents input sector *i* required by industry j. So, the input coefficients can be shown as:

$$a_{ij} = \frac{x_{ij}}{X_j} \qquad (2)$$

This equation can be also written as:

$$X_i = \sum a_{ij}X_j + F_i \qquad (3)$$

The matrix form of equation (3) can be written as

$$X = A \cdot X + F \qquad (4)$$

Equation (4) introduces the basic equation to the open Leontief economic system stating the gross output X as the total output of intermediate products AX and the final demand F. This equation can be solved for X: (I - A) X = F where I is the identity matrix and the matrix IA indicates the technology matrix. If (I - A) ≠ 0 meaning that it is not a singular matrix then the inverse matrix $(I - A)^{-1}$ exists and each output is measured as following:

$$X = (I - A)^{-1} F \qquad (5)$$

The gross output level of X's essential to sustain a given vector of final demand F in the input output model, is calculated by equation (5). Inverse matrix $(I - A)^{-1}$ is identified as the Leontief inverse matrix. The entire production system (ΔX) is affected by changes in final demand (ΔF). This can be shown by the following equation:

$$\Delta X = (I - A)^{-1} \Delta F \qquad (6)$$

3.2.1.4 Rasmussen Method

In the input-output framework two types of economic impacts exist by the production of a particular sector, on other sectors of the economy. Rasmussen (1956) looks at such relationship as the interconnection between industries. Backward linkages of a sector indicate the type of internal relations through which inputs are purchased in order to be combined in the production process to produce outputs. On the other side forward linkages indicate the sales of a sector's output to other sectors as inputs (Miller and Blair, 2009). In the primary input sector, it is assumed that there is no backward linkage, similarly in the final demand sector, no forward linkage and high backward and forward linkages belong to the intermediate sectors of the economy.

The method gauges that if industries use the incremental output of the other industry as incremental intermediate inputs then what will be the effect on the other industries' output and to what extent other industries can increase their outputs (Los, 2002; Rasmussen, 1956). Nørregaard Rasmussen by using the column sums of the Leontief inverse $(I - A)^{-1}$ formulated the method (e.g. Cai and Leung, 2004; Drejer, 2003), through which one can gain idea "about the economic structure of sectors within an economy (Sonis, et al., 2000: 401)", and also his method helps to identify those sectors which have above average multiplier effects (Cuello and Mansouri, 1992).

Rasmussen (1956) had indicated that utilizing the inverted Leontief matrix can denote inter-industry linkages. The size of the dependency between production and consumption is measured through the analysis of the strength of the relationship between subsectors of a sector.

The column sums of the Leontief-inverse from the demand side of the input-output model gives the backward linkages, and the row sums from the supply side defines the forward linkages. The relationship of inputs purchased along each row of the (i) in column (j) of the inverted Leontief matrix introduces a backward linkage, indicating the relationships in the purchase of inputs from various sectors in the same column, directly and indirectly.

In fact, the relationship between the rows in a matrix inverse is shown through the backward relationship which is also associated as supplier relationships in the industry. The greater than unity in average index value of the backward and forward linkages can signify the strong ties that exist between sectors or key sectors.

The analysis of the elements for $(I-A)^{-1}$ would specify the structure of the economy like that of the industry. If the components of the $(I-A)^{-1}$ matrix are denoted by (K_{ij})'s then the sum of the column components of the $(I-A)^{-1}$

$$\sum_{i=1}^{m} K_{ij} = K_j \qquad (7)$$

illustrates the total input required for an additional unit in the final demand of the j^{th} sector. Similarly the sum of the row elements

$$\sum_{j=1}^{m} K_{ij} = K_i \qquad (8)$$

states the additional output of sector number i required to tackle with a unit rise of final demand for all the industries. Rasmussen interprets the averages.

$$\frac{1}{m} K_j \qquad (\mathrm{j} = 1, \ldots, \ ) \quad (9)$$

"As an estimate of the direct and indirect rise in output provided by an industry if the final demand for the products of industry number j (j = 1, . . . , m) increases by one unit." [7]To the set of averages

$$\frac{1}{m} K_i \qquad (\mathrm{i} = 1, \ldots, n) \qquad (10)$$

Rasmussen interprets in a similar way. The key sector conceptually can be considered as a sector that can motivate the economy and drive the country into the development through bounds and leaps. A key sector refers to industries which have strong ties in both sides; that is, for the user of the products and the seller of inputs.

[7]. Refer Rasmussen, Studies in Inter-Sectoral Relations, 1956. chap. 8, page 133.

Inter-industry linkage indexes determine the strengths of these ties which for the key sectors backward and forward inter-industry linkages are above the average.

Moreover, in order to define the key sectors; direct input coefficients and direct output coefficients of sectors in form of total requirement matrix from the demand- and supply-driven model are employed (see Oosterhaven, 2008). For making inter-industrial comparisons and easily detecting the key sectors, these indices are not suitable and for this purpose the set of averages in (9) and (10) are normalized by the overall average defined as

$$\frac{1}{m^2}\sum_{j=1}^{m}\sum_{i=1}^{m}K_{ij} = \frac{1}{m^2}\sum_{j=1}^{m}K_j = \frac{1}{m^2}\sum_{i=1}^{m}K_i \quad (11)$$

Moreover, we consider the following indices which are termed by Rasmussen

$$U_j = \frac{\frac{1}{m}K_j}{\frac{1}{m^2}\sum_{j=1}^{m}K_j} \quad (12)$$

the "Index of Power of Dispersion", which describes how an increment in final demand for a sector is pervaded across the entire economic system, is used to measure Hirschman's backward linkages by using input-output tables (Drejer, 2003).

Though, as Cai and Leung (2004) address the assumption of dispersion index of Rasmussen, uniform final demand differs for all the sectors despite of any changes across sectors, limits the power of his index. The use of uniform demand changes is justifiable provided that the differences between sectors do not substantially influence the measurement of backward linkages (Cai and Leung, 2004).

$$U_i = \frac{\frac{1}{m}K_i}{\frac{1}{m^2}\sum_{i=1}^{m}K_i} \quad (13)$$

the "Index of Sensitivity of Dispersion" which can also be interpreted as measure of Hirschman's forward linkage. Likewise, Rasmussen measured the industry forward linkages via sensitivity of the dispersion index. The author acknowledged that through this index one can find the incremental effect of the increase in the output of an industry on the final demand of one or more industries in the economy (Drejer, 2003; Sonis *et al*., 2000).

These indicators make the determination of the key sectors possible in an economy. Drejer (2003) argues that for countries in which input-output tables have a lot of non-available elements; the sensitivity of the dispersion index is likely more helpful.

As mentioned above, the averages $1/m\, K_j$ have been interpreted as the requirements of inputs if the final demand of industry number j increases by 1 unit; For the case of $U_j > 1$, the industry draws heavily on the rest of the system and vice versa in the case of $U_j < 1$.

Similarly if $U_i > 1$ then the industry number i will have to increase its output more than others for an additional increase in final demand from the entire system. The indices in equations (12) and (13) are based on the method of averaging. However, according to the theory of statistics averages are sensitive to extreme values and may give illusive results; it is possible that an increase in the final demand for the product of a particular industry, characterized by a high index of power of dispersion, may not affect other industries. Such a case would arise if a particular industry only draws heavily on one or a few industries. Consequently, the indices in (12) and (13) do not completely describe the structure of a particular industry.

To surmount this difficulty a variability measure must be defined and the indices of coefficient of variation are as following:

$$V_j = \frac{\sqrt{\frac{1}{m-1}\sum_{i=1}^{m}(K_{ij} - \frac{1}{m}\sum_{i=1}^{m}K_{ij})^2}}{\frac{1}{m}\sum_{i=1}^{m}K_{ij}} \qquad (14)$$

and

$$V_i = \frac{\sqrt{\frac{1}{m-1}\sum_{j=1}^{m}(K_{ij} - \frac{1}{m}\sum_{j=1}^{m}K_{ij})^2}}{\frac{1}{m}\sum_{j=1}^{m}K_{ij}} \qquad (15)$$

A high V_j can indicate that a particular industry draws heavily on one or a few sectors and a low V_j can be interpreted as an industry drawing evenly from the other sectors. The similar interpretation is for V_i's.

A key sector can be specified as one observing conditions under which (a) both U_j and U_i are greater than unity (U_j> 1, U_i> 1), and (b) both V_j and V_i are relatively low [e.g. less than one]. Alternatively one can easily interpret these in terms of Hirschman's definition; the sector is a key one in the case that has a high forward as well as backward linkage. Since U_j and U_i have already been defined as backward and forward linkages Hirschman defines that any industry in which both U_j and U_i are bigger than one, can be detected as a key sector.

It is notable that since no restriction is stipulated on the values of V_j and V_i in Hirschman's definition of the key sectors, therefore he ignores the spread effects of the development of an industry. These spread effects are increasingly important from the perspective of industrial diversification and economic development.

However, in 1992 Cuello and Mansouri describe the drawbacks related to the Rasmussen model; one problem is with the average measure which gets affected by maximum values. If a sector adopts above average multiplier it does not warrant that the sector has higher number of linkages with other sectors. Indeed the system may possess only few but strong linkages. The other problem arises when the probability of inducement of investment is failed to be accounted by the computed indices. In fact, the investment stimulated by key sectors relies on how much demand exists for inputs and how much supply exists for inputs respectively translated into backward and forward linkages. Furthermore they pointed that the indices do not represent anything about the generated potential investment opportunities or the occurrence of the induced investment. Finally the tendency of the indices for being uniform across sectors makes them unreliable.

3.2.2 Artificial Neural Networks

After detecting the key sectors over the sample period, growth model of Indonesia was developed as follows:

$$\gamma_t = \mathcal{F}\ (x_1, x_2, x_3, x_4, x_5) \tag{20}$$

where γ_t abbreviates GDP Growth; x_1, x_2, x_3, x_4 represent the investment share (percentage of GDP) of the acquired strategic sectors (manufacturing, agriculture, construction, hotel and restaurants) in different years for the whole period; and x_5 represents trade openness (ratio of total trade: export plus import to GDP). The traditional approaches (e.g. linear regression) assume the linear mechanisms and are limited by their linearity; therefore, they may be not suitable when we face with a process which is nonlinear.

In the real word, systems are often nonlinear (Zhang et al., 1998), thus the inherently nonlinear structure of neural networks is useful to capture the complex underlying relationship in many real world problems. Artificial neural network (ANN) as a class of flexible nonlinear models is an information processing technique to imitate the function of the biological neural systems. Their applications have expanded over the years in different fields and began to be the popular option over other methods.

These networks may be applied as a substitute for linear regression, multivariable regression, and other statistical analysis and techniques (Singh et al., 2009). As many problems in economics can be considered in terms of classification or regression, neural networks might be capable of providing effective tools for their solution (McNelis, 2005). "One of the most significant advantages of ANN models is their ability to approximate a large class of functions with a high degree of accuracy. (Zhang & Su, 2002)"

Their power over other approaches such as traditional model-based methods arises from modelling the nonlinear processes of a system and processing the information without any prior assumption in the nature of the modelling process (Khashei & Bijari, 2010).

Instead, the characteristics of the data largely determine the network model. ANN consists of a set of interconnected processing components similar to neurons linked with weighted connections analogous to synapses. It is a computational system with the ability of adaptive biological learning. Learning happens through training of a true set of input/output data. The training algorithm adjusts the connection weights iteratively to reduce the error. Connection weights conserve the needful knowledge for solving specific problems. The schematic of typical neural network neurons is shown in Figure 1.

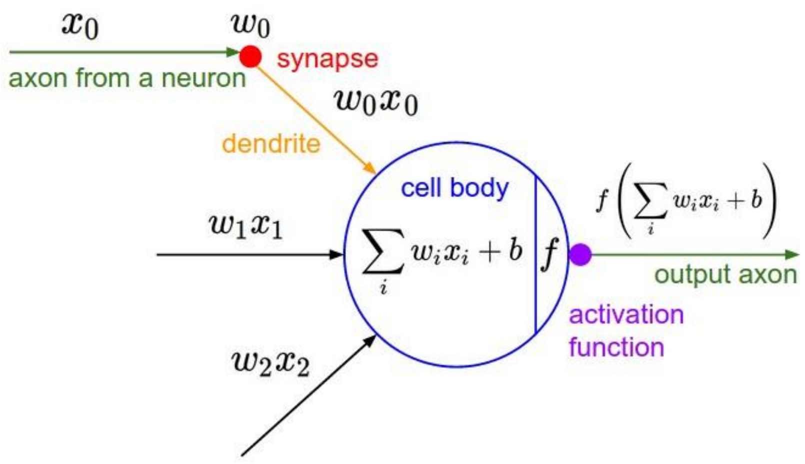

Figure 1. Neural Network Neurons (Bishop, 1995)

The feedforward networks are the most popular ANN structures with a hierarchy of neurons, organized as a series of layers. In this type of ANNs, information moves in one direction. The input layer feeds the network whereas the output layer makes the overall mapping of the network. Middling the input and output layers may lay hidden layer(s) to do more remapping or computing. The process of learning linear or non-linear boundaries is performed using the activation function. A popular choice of activation function is logistic function known as the sigmoid function.

$$F(x) = \frac{1}{1+e^{-x}} \qquad (21)$$

Using the logistic function in the single layer network makes it similar to the logistic regression model used extensively in statistical modeling. This function has a continuous derivative which can be employed in backpropagation. It is preferably chosen since its derivative is easily calculated.

The feedforward network begins the training process with propagating all the input values through the network and determines the output units. Thereafter the error of the actual output and the desired output is back propagated from the output layer to the previous layer(s) and proceeds until the input layer. Along this path, the error is reduced through the adjusted weights. Decreasing an error during the process is called Stochastic Gradient Descent. Multilayer perceptron (MLP) as a common type of feedforward artificial neural networks has been broadly utilized for classification or regression problems (Loh & Tim, 2000; Kenneth et al, 2001; Cohen & Intrator, 2002).

Due to these advantages, this study applied a Multi-Layered Perceptron (MLP) Artificial Neural Network (ANN) as an advanced valuation technique in order to capture the complex and nonlinear phenomenon of economic growth in Indonesia. MLP uses a supervised learning technique with an error back propagation algorithm (Rosenblatt, 1961).

The inherent advantage of MLP networks is their capacity to manage nonlinear functions to simulate nonlinear and dynamic systems in modelling. Generally the MLP network is formed with an input and an output layer, and one or more hidden layers of neurons between the input and the output layers. When the pattern of observations within an input layer is recognized, it can be used to predict the position of the consequent observations through the output layer. Every neuron of the hidden layer maps a weighted average of the inputs with a sigmoid function.

In a modelling application, each input neuron deputes one of the independent variables, while the output neuron illustrates dependent variable. The network is trained by modifying weights (w_i), which are attached to each link. The design of MLP can be indicated as:

$$\hat{y}_i = f(x_i) = w_2 g(w_1 x_t + w_b)\, w_2 g_t \quad (22)$$

Where $\hat{y}_i$ stands for the predicted value of y (GDP growth) and $x_t = [x_1, x_2, x_3, x_4, x_5]'$ is the vector of 5 independent variables. g as a nonlinear function displays the hidden layer with n sigmoid neurons. w_1, w_2 and w_b respectively represent the weight matrices of the hidden layer, output layer and the weight vector for connecting the bias to the hidden layer.

3.2.3 Variable Ranking Algorithms

One drawback of the above models is referred to the redundancy problem in the estimation process. When there are different leading sectors, the model might not work as expected (Copas, 1983; Nathans et al., 2012). Therefore, the important issue here is how to detect the optimal degree of investment and specify the most significant sector with the existence of different leading sectors.

To answer this question, we applied two different variable ranking algorithms (Feature selection methods) to define the optimal weights of each leading sector.

For many regression problems, using a higher number of features (variables) does not provide higher accuracy; indeed, in some cases it even decreases the speed and predictive accuracy of the model. Therefore, feature selection can serve as a pre-processing tool to reduce the number of irrelevant features (variables). Moreover, a good feature selection method can reduce cost of feature measurement while increasing predictor efficiency and estimation accuracy.

This algorithm is of considerable significance in pattern classification, data analysis, machine learning, and data mining applications. Several methods have already been applied to apply feature selection on training and testing data, as examples;

sequential search algorithms (Pudil et al., 1994), mutual information (Battiti, 1994), genetic algorithms (Raymer et al., 2000), and tabu search (Zhang et al., 2002).

A feature selection algorithm such as a search technique proposes new feature subsets through measuring the scores of different feature subsets. The evaluation metrics which heavily influence the algorithm, distinguish between the categories of feature selection algorithms. Usually the variable elimination methods were classified into three main groups: filters (open-loop approaches) wrappers (close-loop approaches), and hybrid approaches (Liu & Yu, 2005).

Filter method acts free of any learning algorithm and ranks the features with some criteria, thereafter, the highly ranked ones are applied as predictor. The wrapper method which is involved with the learning algorithm, selects the features based on their learning performance of a particular learning model. Wrapper approach performs better in terms of solution quality compared with filter approach with considering the fact that running learning algorithm for each subset in wrapper approach is too time-consuming with high computational costs. While, hybrid approach attempts to enjoy the characteristics of both filter and wrapper methods (Kashef & Nezamabadipour, 2015).

The present study benefits from stepwise regression methodology in order to detect the most significant key sector in Indonesia in the linear form, while in the nonlinear form, we utilized the Ant Colony Optimization (ACO) as one of the evolutionary feature selection methods for the training of feed-forward artificial neural network (ANN). Regarding the application of ACO in compliance with the stepwise regression; it is worth mentioning that number of desired variables included in the model is specified by one in both algorithms. Therefore, in the presence of capacity constraint (i.e. the number of feature is pre-determined by the user) our hypothesis definition is defined as a binary optimization problem.

3.2.3.1 Stepwise Regression

Stepwise regression is one of the approaches for selecting variables in a regression model. It detects the subset of predictor variables used in the final regression model by analysing the data. Finding this subset of independent variables seeks to fulfill two different objectives.

First, the regression model should be complete and realistic as much as possible; every regressor even if has a remote relation with the dependent variable should be included. Second, the regression should consist of fewer variables; since each regressor which is irrelevant would decrease the precision of the estimated coefficients and the predicted values. In addition, further variables augment the complexity of data collection and maintenance of the model.

The goal of variable selection is to achieve a balance between simplicity (as few regressors as possible) and fit (as many regressors as needed).

Stepwise regression is a combination of forward (Step-Up) selection and backward (Step-Down) elimination techniques.

Forward selection is often applied to administer a primal sieving of the candidate variables in the existence of a large group of variables. This procedure is also advantageous when multi-collinearity is a problem. This method begins with no predictors in the model. The variable that has the lowest p-value is selected. Adding variables is stopped when none of the remaining variables is significant. Once a variable enters the model, it cannot be deleted.

In contrast to forward selection method, backward selection model commences with all candidate variables in the model. At each step the variable which is the most insignificant is eliminated, and this process continues until no non-significant variable remains. The significant level threshold is set by the user. Since this method has a downward working principle instead of upward; a large value of R-Squared is retained. This procedure may include unnecessary variables that lead to less popularity.

In the mentioned algorithm, we employed two different approaches. First, by using forward stepwise as the selection method, we identified the vital sector for growth based on the normal criteria. Then, to trace the effects of linear combination of those sectors on real GDP growth, we adopted combinatorial approach as another selection method. Combinatorial method evaluates each potential combination of variables. Selection of an optimal combination with higher R-squared, guarantees the generality of the method.

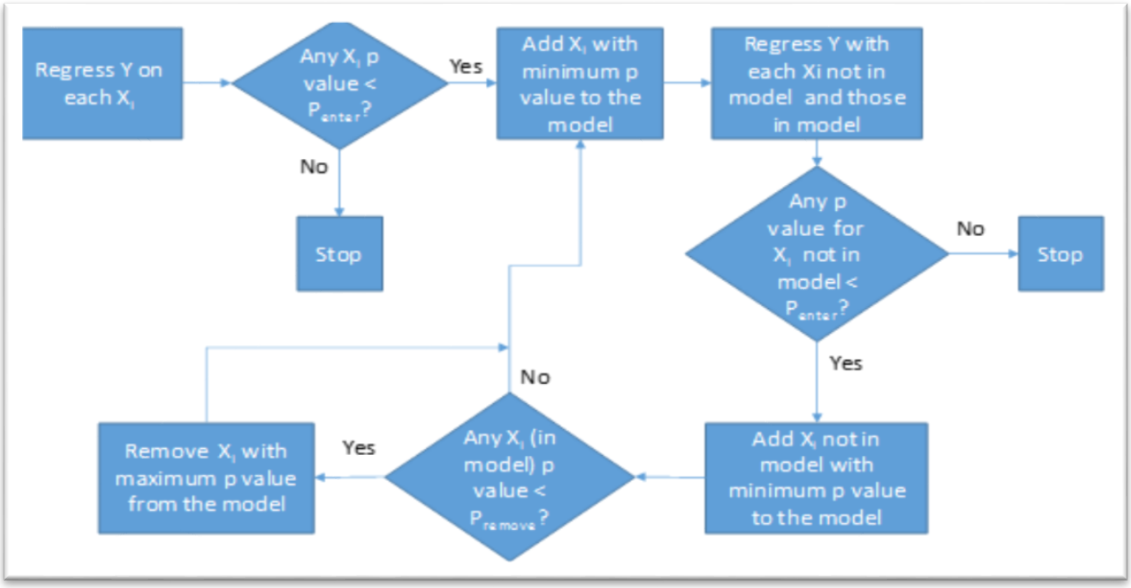

Figure 2. Stepwise Regression Flowchart (https://www.spcforexcel.com)

3.2.3.2 Ant Colony Optimization (ACO)

Similar to the linear examination, shortening the nonlinear model provided by MLP is substantial. Hence, ACO is applied to perform a rank-order analysis. ACO is a probabilistic technique for solving hard combinatorial optimization problems. It has been unexpectedly successful for solving problems in recent years such as the feature selection problem (Ani, 2005). This algorithm was initially developed by Marco Dorigo in 1992. He was inspired to propose the algorithm to search for an optimal path in a graph based on the ant's social behavior.

In the natural world, ants seek a path between their colony and a source of food and communicate information by leaving pheromone in order to mark some favorable paths that should be followed by the other ants. The ants select their path with regard to probabilities depended on pheromone trails previously have been left by them. Deposited pheromone will be evaporated soon, but it remains on the ground in short-time as trail of the ants.

Once one ant discovers a short path from a food source to the colony, the other ants do not travel randomly anymore. However, they follow that short discovered path with higher expectancy and the pheromone density becomes higher on that short path. The collective behavior eventually leads to positive feedback to all the ants following a single path. Therefore generally three important reasons lead to find the shortest path by the ants; 1- deposited pheromone by ants, 2- evaporation of deposited pheromone, and 3- probability in selecting the path.

The idea of the ant colony algorithm models the behavior of simulated ants as a searching process for finding a minimum cost solution. In ACO, several generations of artificial ants construct good solutions. Each ant of each generation provides a solution; the ant which detects a proper solution signs its path by leaving some amount of pheromone on the edges of its path. Then the algorithm walks around the graph representation of the problem and explores the best solution. The visual algorithm for ACO is shown in Figure 3.

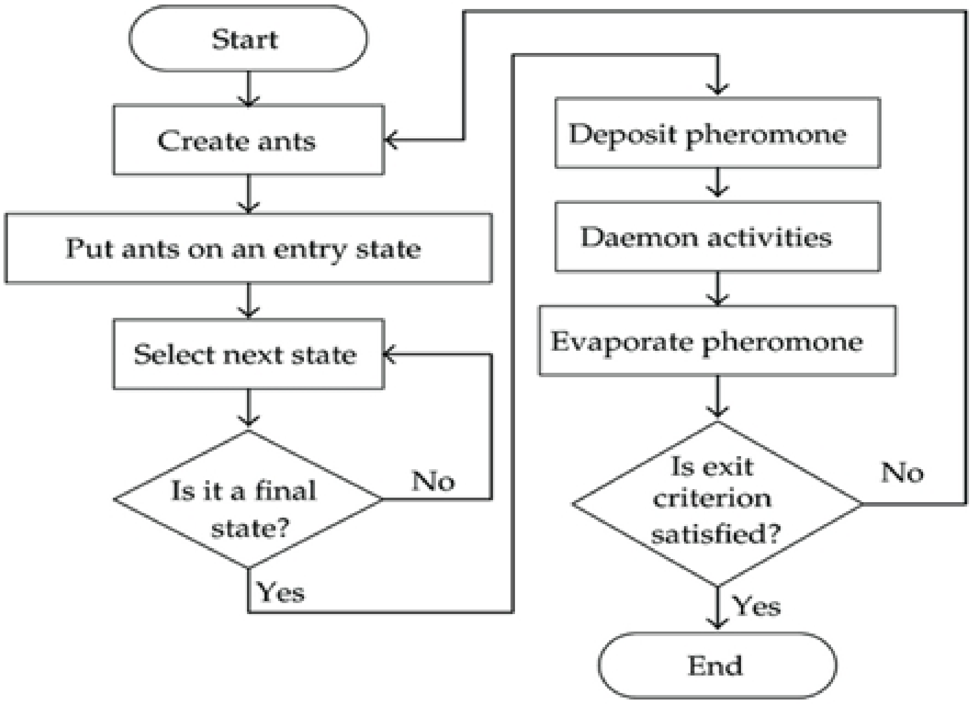

Figure 3. Ant Colony Optimization Flowchart (Xu et al., 2012)

There are numerous ACO-based methods found in the literature that describe different scenarios for feature selecting. (For example see: Kashef & Nezamabadipour, 2015; Xue, et.al, 2016).

Since many of them are limited so that are unable to track any desired sequence of unseen features (i.e. ants travel a constant sequence of features) therefore, we follow an advanced ACO-feature selection algorithm developed by Kashef and Nezamabadipour (2013) to overcome the considered shortcoming.

With applying ACO for variable selection, each feature treats as a graph node wherein all nodes are fully connected to allow any other feature be selected as the next. Then the procedure manages a solution by moving through neighboring nodes under the probabilistic transition rule defined as follows:

$$P_{i,j}^{k}(t) = \begin{cases} \frac{\tau_{ij}^{\alpha}\eta_{ij}^{\beta}}{\sum_i(\tau_{ij}^{\alpha}+\eta_{ij}^{\beta})} & if\ i\ and\ j\ are\ admissible\ nodes \\ 0 & ,\ otherwise \end{cases} \quad (23)$$

Where $P_{i,j}^{k}$ stands for the transition probability from feature (node) i to feature (node) j for the k th ant at time t; τ_{ij} shows the amount of pheromone on a given edge i, j; The attractiveness of the features is represented by Heuristic information parameter (η) depending on the dependency degree which is necessary to be used for finding better solution; otherwise, the algorithm may become too greedy[8] [Ke.L. et. al., 2008].

α and β are two parameters for governing the relative importance of the pheromone value versus the heuristic information. The transition probability $P_{i,j}^{k}$ equilibrates between pheromone intensity τ_{ij} which covers history of previous successful moves and heuristic information η expressing desirability of the move. However, an effective balance can be achieved between them through proper selection of the parameters α and β.

If α=0, no pheromone information is used, and If β=0, the visibility of moves is overlooked. At the end of each iteration, when all ants have completed a solution, the amount of pheromone is updated according to equation 24.

$$\tau_{i,j}(t+1) = (1-\rho)\,\tau_{i,j}(t) + \sum_{k=1}^{m} \Delta\tau_{ij}^{k}(t) + \Delta\tau_{ij}^{g}(t) \quad (24)$$

$$\Delta\tau_{i,j}^{k} = \begin{cases} \frac{Q}{F_k} & if\ ant\ k\ travels\ on\ edge\ ij\ in\ T_k \\ 0, & otherwise \end{cases}$$

[8] Hence, this study adjusts the Heuristic information parameter (η) to 1; in order to prevent a probable poor solution arising from the model.

Where ρ[0, 1] represents the rate of pheromone evaporation; m shows the number of ants; $\Delta\tau_{ij}^{k}(t)$ and $\Delta\tau_{ij}^{g}(t)$ respectively illustrate the amount of pheromone laid on the edge (i, j) by the k^{th} ant, and the amount of pheromone deposited by the global best ant g up to the time t over the edge (i, j); Q is a constant (which adjusted to 1) , and F_k is the cost of the solution found for the k^{th} ant [we considered prediction error as cost function (Leray & Gallinari, 1999)].

The proposed algorithm allows to update the pheromone trails for only the global best ant (max-min ant system), whereas the value of pheromone on each path is limited to [τ_{min} ,τ_{max}]. Hence equation (24) can be modified as follows:

$$\tau_{i,j}(t+1) = \left[(1-\rho)\,\tau_{i,j}(t) + \Delta\tau_{ij}^{g}(t)\right]_{\tau_{min}}^{\tau_{max}} \qquad (25)$$

The process is repeated until the stopping criterion [either a number of iterations or a solution of desired quality (Meena et.al, 2012)] be reached. Learning process in Multi-Layer Perceptron (MLP) neural network with application of ACO tends to minimize the cost of prediction error by considering the importance factors of irrelevant features and maximize the importance factors of relevant features. Finally the features are sorted from more significant to less significant based on their importance factors, and then the optimal feature subset is generated containing the features with more importance factors.

Chapter 4

EMPRICAL FINDINGS

This part presents the application of the techniques discussed in the previous section and provides the findings. Figure 4 plots per capita GDP growth rate for Indonesia over the entire sample period. It shows that GDP growth rate in Indonesia has experienced different regimes over the past decades.

As highlighted in Figure 4, in order to show the structural breaks, we applied Bai and Perron's (2003) tests of multiple structural breaks to the residuals of AR (3)[9] model. The results are shown in appendix A which strongly evidence the existence of structural breaks in individual processes of GDP growth.

Therefore realizing sectoral structural changes for economic development requires analysis of direct economic effects which appear as a change in output of each sector.

At least to some extent, this need can be satisfied by using input-output table. Analyses of the input-output table in Indonesia revealed different classes of sectors as the engine of growth in the economy over the investigated period.

[9] The number of lag (3) is selected based on SIC information criteria.

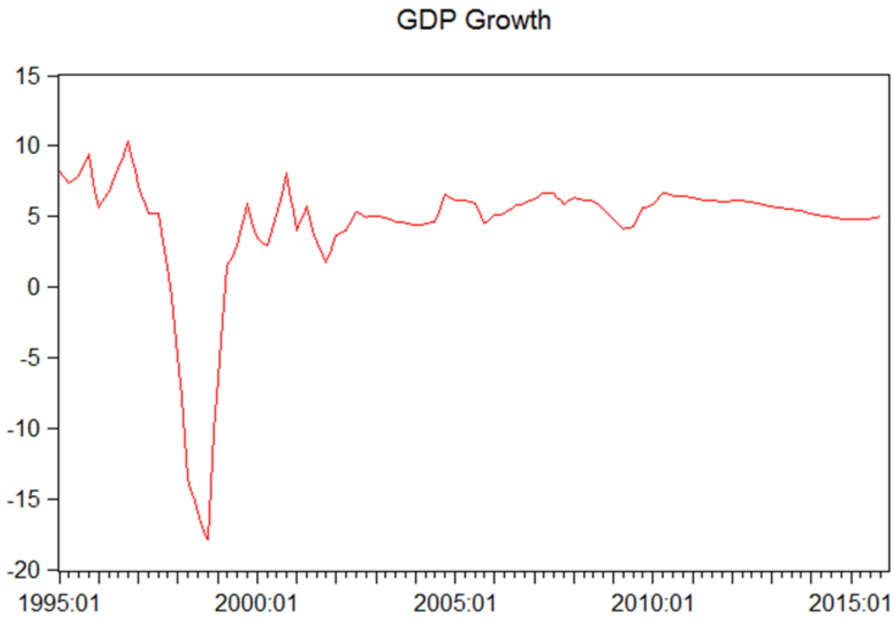

Figure 4. GDP per Capita Growth Rate in Indonesia (1995-2015)

Table 1 presents the key sectors for each year and Table 2 reports the frequency of each sector as well as its corresponding rank. We found that, manufacturing sector ranks first as it has the majority among all the sectors; construction and hotel and restaurants appear second over the sample period and the last ranked sector is agriculture with lowest frequency among the all sectors. The result highlights expansion of manufacturing sector in the period surrounding economic crises and post-crises. The changes in the key sectors showed a trend of transformation in economic activities from agriculture to industry.

Our results are consistent with previous studies that also discovered changes in specialization and industrial concentration patterns during the period 1998 to 2007 (Wahyudi et al., 2010).

Table 1. Input Output Results

Year	Key Sectors
1995	Agriculture
1996	Agriculture
1997	Manufacturing
1998	Manufacturing
1999	Manufacturing
2000	Manufacturing
2001	Construction
2002	Construction
2003	Hotels and Restaurants
2004	Manufacturing
2005	Manufacturing
2006	Hotels and Restaurants
2007	Manufacturing
2008	Manufacturing
2009	Construction
2010	Construction
2011	Construction
2012	Hotels and Restaurants
2013	Hotels and Restaurants
2014	Hotels and Restaurants

Table 2. Frequency of Each Sector over the Entire Sample Period (1995-2014)

Sector	Frequency	Rank
Agriculture	2	3
Hotels and Restaurants	5	2
Construction	5	2
Manufacturing	8	1

As mentioned before, in order to determine the relationship between the extracted annual key sectors and real GDP growth; we adopted a multiple linear regression using the least square estimation method. Before doing this analysis, it is important to clarify whether the series are stationary. In this respect, one can use different methods available for this test. One of the most famous methods is Augmented Dicky Fuller (1979) or ADF approach. Thus, this study utilizes the ADF test with two different scenarios namely Intercept and Trend and intercept. It is noteworthy to mention that, these two scenarios differ regarding the deterministic components included in the autoregressive function. In each test, the first specification was performed with containing only a constant term, while the second exemplar included a time trend in addition to a constant. Table 3 exerts the results of the unit root tests for all variables based on the null hypothesis of unit root tests (The series are non-stationary). In Table 3, after allowing for both intercept and trend one can see that GDP growth appears to be stationary as its null hypothesis is rejected. Whereas when only intercept is allowed, the null hypothesis is not rejected; while the other four variables are stationary at arbitrary critical values.

In the presence of the structural break in the form of a shift in the levels of the data, the conventional unit root tests are not free from spurious rejection. Also because of their break point nuisance invariance, both null and alternative hypotheses are affected by either size or location distortions.

Therefore in the present study, to ensure the appropriate maximal order of integration and get rid of the non-rejection situation of the null hypothesis of unit root in presence of possible structural events which may induce shifts in the regime, the Zivot-Andrews structural break unit root test proposed by Zivot and Andrews (1992) is employed.

Application of unit root hypothesis with structural breaks has at least two benefits. First, it neutralizes a likely biased test outcome in the direction of non-rejection of the null hypothesis. Second, recognizing the probable existence of structural breaks in the model presents crucial witting for testing whether a structural break on certain series is ascribed to particular regime change.

Table 3. Unit Root Test

Variables	**Trend and Intercept**	**Intercept**
Agriculture	-3.46**	-5.01***
Construction	-3.13*	-4.36**
Hotel and Restaurants	-10.25***	-6.03***
Manufacturing	-2.98*	-3.01*
Trade Openness	-4.12**	-3.78**
GDP Growth	-2.79*	-2.12

*, **, and ***, denote significance at 10%, 5%, and 1%, respectively.

Table 4 shows the results of ZA unit root test allowing one structural break with two different specifications. Under different specifications, the null hypothesis in ZA unit root test is that the series has a unit root with a structural break in the intercept for the constant scenario, while the series has a unit root with a structural break in the intercept and trend for the scenario with constant and trend.

As shown in Table 4 the null hypothesis is rejected at 1% level of significance for GDP growth, hence in this manner we conclude that the GDP growth is stationary.

Table 4. ZA Unit Root Test Results

Specification	Constant	Constant and Trend
GDP Growth	-6.129***	-5.787***

***, denotes significance at 1% level.

As we have shown that the all variables in the model are stationary at standard level of significance. Hence we next turn to estimate the model using linear regression which the results are illustrated in Table 5. Given that the existence of the structural break is approved by ZA unit root test as well as the result of Bai and Perron's (2003) test of multiple structural breaks (see Appendix A.), hence in the R.H.S of the estimation model we introduced one dummy variable such that its value is 1 for 1998-Q2, otherwise is 0.

The statistical summary of the linear model yields significant coefficients for all sectors and R^2 of 0.6125, while the variable of trade openness is not significant in all levels of significances. The output of the regression model specified the existence of a relationship between the key sectors and real GDP growth as all variables are significant. This completely makes sense as Hirschman said; growth and development achieved considerable advances through investment in the leading sector(s).

Table 5. Regression Results

Dep. Variable: GDP Growth	
Variables	$\boldsymbol{\beta_j}$
Agriculture	0.3539***
Construction	1.5451***
Hotel and Restaurants	1.4210**
Manufacturing	1.5632**
Trade Openness	1.0003
Intercept	1.2542**
Dummy_98	1.019*
R-squared:	0.6125

** $p < 0.05$,*** $p < 0.01$

The estimated coefficient of the manufacturing sector has a higher weight compared to the other variables, which is somewhat in accordance with the results of correlation analyses (Table 6) between the key sectors and real GDP growth. It is notable to mention that in order to satisfy the assumption of classical regression model, the standard errors are adjusted for the potential heteroscedasticity and serial correlation using the HAC consistent covariance estimators proposed by Newey-West (1987). In many settings for the robust standard errors (i.e. White) the residuals of the estimated equation are assumed to be serially uncorrelated. While HAC is consistent in the presence of both heteroscedasticity and autocorrelation of unknown forms, under the assumption that the autocorrelations between distant observations die out.

In the presence of different leading sectors, the most significant sector can be specified by means of stepwise regression method. This methodology as an appropriate algorithm in the variable ranking area helped us to define the optimal weights of each leading sector. Table 7 presents result of the stepwise regression by forward looking selection and the results of combinatorial approach.

In order to control the stepwise estimation method, we adjusted the stopping criteria using a p-value tolerance for adding or eliminating variables to 5%. Besides, the desired number of variables (Features) has been set to one.

Table 6. Correlation Analyses Table

Variables	***Growth***	***Con.***	***Agri.***	***Manu.***	***Hotel***	***TO***
Growth	1					
Cons.	0.69**	1				
Agri.	0.35**	0.42*	1			
Manu.	0.73**	0.63***	0.78*	1		
Hotel	0.63***	0.51	0.33	0.48**	1	
TO	0.51**	0.61*	0.44	0.51	0.47**	1

Table 7. Stepwise Outputs

Method:	**Stepwise (Forward)**	**Combinatorial**
Dep. Variable:	***RGDP Growth***	***RGDP Growth***
Manufacturing	1.884*	1.889*
Construction	-----	1.089**

Where *p < .01, **p < .05, Stopping Criterion: P-Value Forward/Backward = 0.5/0.5, Number of search repressors is adjusted to one.

The results proved that manufacturing sector is one of the most significant sectors affecting GDP growth among all key sectors throughout the period. Therefore, the highest investment should have been allocated to the manufacturing sector to eliminate the lack of skill-intensive jobs and remove the constraints on growth as pointed out by Oberman et al. (2012) and Harvards report (2010).

According to the results of combinatorial approach, a combination of manufacturing and construction could lead to higher growth. Since the result showed a simultaneous attempt to invest in both sectors, a conflict may emerge with the unbalanced investment strategies leading to a question; which sector should be invested at first.

To answer this question, variables can be ordered based on their magnitudes. However, this approach is not reliable as Courville and Thompson (2001) approved that it is not ideal to rely on β weights since independent variables are sometimes inter-correlated. Therefore, stepwise algorithm (based on the OLS method) does not solve the problem. The incertitude of rank ordering of sectors motivated the authors to apply other valid methods that are discussed later.

In many applications the shape of the functional relationships between the response (dependent) and the explanatory (independent) variables are not predetermined. Therefore, the form of relationships is unknown.

There is a strong belief in many areas, including economics that the correct link between variables should be nonlinear. However, in the presence of weak nonlinearity or strong linearity structure, fitting nonlinear models may well lead to overfitting (Tersvirta et.al., 1993). Hence, in order to show the need for a nonlinear approach, the Teraesvirta's neural network test for neglected nonlinearity (Tersvirta et.al., 1993) has been used.

This test uses a Taylor series expansion of the activation function for improving White's neural network test[10] (Lee et.al., 1993) in order to reach an appropriate test statistic. The test is performed in a Lagrange multiplier (LM) test framework under which ANN model does not need any estimation for the alternative hypothesis. When the null hypothesis is rejected, the model is suffering from neglected nonlinearity which means that better predictions may be attained with the nonlinear model than the linear one.

We performed Teraesvirta's test using a free software environment for statistical computing [R (3.3.3)], while the F-statistic type[11] has been used similar to the classical linear regression. The value of the test statistic ($\mathcal{X}^2$) is 252.23 with 0.001 for the p-value of the test. The small p-value confirms the existence of nonlinear relation goes from independent variables (X_js) to the output (y). Hence, in order to examine the unknown pattern of the data, MLP as a nonparametric nonlinear regression was adopted with the hope of capturing and building a more robust model.

[10] The ANN- based neglected nonlinearity test proposed by White, employs a single hidden layer feedforward neural network with nonlinear activation functions with the capable of approximating an arbitrary nonlinear mapping [For details see: Lee et.al., 1993].

[11]However, the test also has been performed with the other type which is based on the Chi-Squared statistic, whereas, both types rejected the null of linearity.

This thesis used MATLAB 2016a software in neural network analyses having a MLP network that consists of input layer, one hidden layer with five sigmoid neurons and one output layer (Figure 5). All data were divided into three data sets namely training (70% of all data), test (15% of all data) and verification (15% of all data).

We tested for 5, 10 and 15 number of neurons to reach neurons number having the lowest error in the hidden layer. As five neurons had the minimum SIC (Information Criterion) in comparison with the other ones, thereupon we employed five neurons as an optimal number. The activation function is a hyperbolic tangent (tansig) and the linear activation function is adopted for the output layer. In addition, the Levenberg-Marquardt method is applied to define the weighting coefficients for training.

To evaluate the efficiency and predictive accuracy of the proposed model, the Root Mean Squares (RMSE) was computed; it is 0.51. This fact clearly illustrates the power of fitting by suitability of R-squared for MLP (0.79), which clearly outperformed the linear form using multiple regressions.

The MLP results showed that during the fitting process, MLP had adequate ability to model a nonlinear relationship. Therefore, MLP, as an impressive class of nonparametric models, might be suitable to construct a tangible model to examine the complex phenomenon.

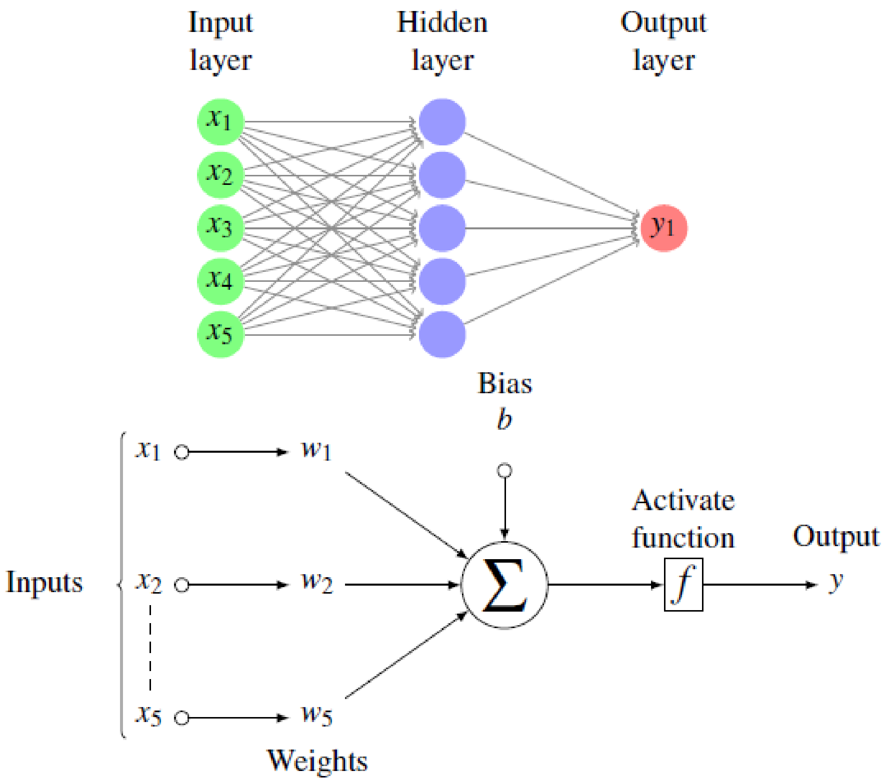

Figure 5. Structure and Block Diagram of the Applied MLP-ANN (Own Source)

To detect the significance of key sectors modelled by MLP, ACO as a population-based metaheuristic approach was adopted for the training of feed-forward artificial neural network (ANN). This technique shortens the nonlinear model provided by MLP with reducing the computation cost of model training, enhancing the generalization abilities of the model and preventing the over-training.

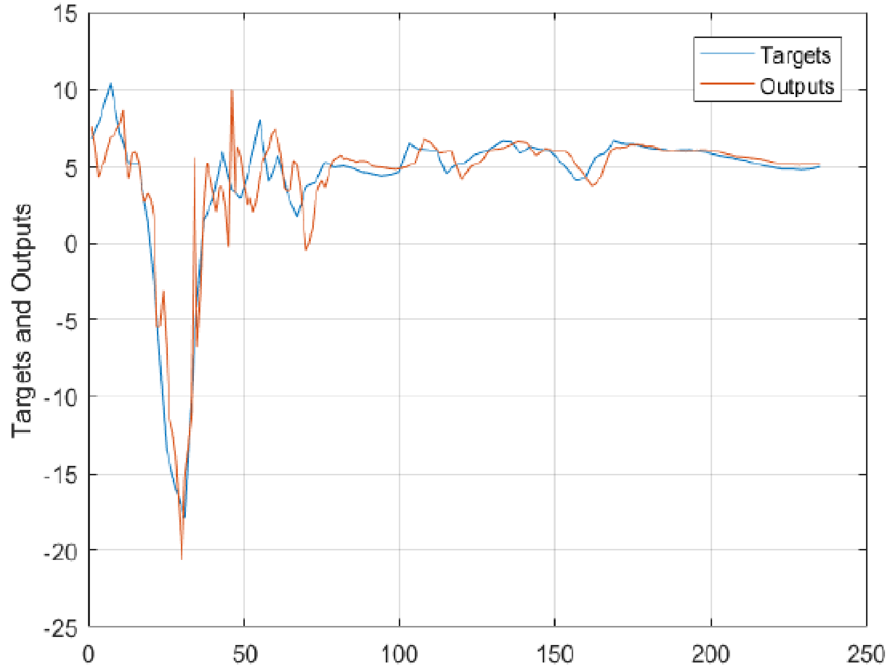

Figure 6. Output by MLP, R-squared 0.79 and RMSE = 0.51

The algorithm was carried out by MATLAB 2016a where the different parameters' values were examined and the best ones were selected for tuning the algorithm. Parameters of algorithm were set according to Kashef and Nezamabadipour (2015). We set population size (number of ants) to 50, while the maximum iterations were adjusted to 100. During each trial 80% of samples were chosen randomly for training and 20% of samples for testing.

The process has been repeated 5 times independently to train MLP-ANN. The evaporation coefficient ρ is 0.05; the initial Pheromone (τ_0) was set to 0.1, while τ_{min} and τ_{max} were set to 0.1 and 5, respectively. The value of β was firstly set to 1 and the value of α varied between [0.2, 1], whereas the best obtained parameters for α and τ_0 were 0.6 and 1, respectively. The acquired results are illustrated in Table 8.

Table 8. ACO Results

Variable	**Rank**
Manufacturing	1
Construction	2
Hotel and restaurants	3
Trade openness	4
Agriculture	5
RMSE (cost minimization)	3.02

The passing trajectory is specified by the rank of each variable. The results indicate an optimal solution for catching higher growth in Indonesia such that the manufacturing sector has been assigned as the starting point with the highest significance degree.

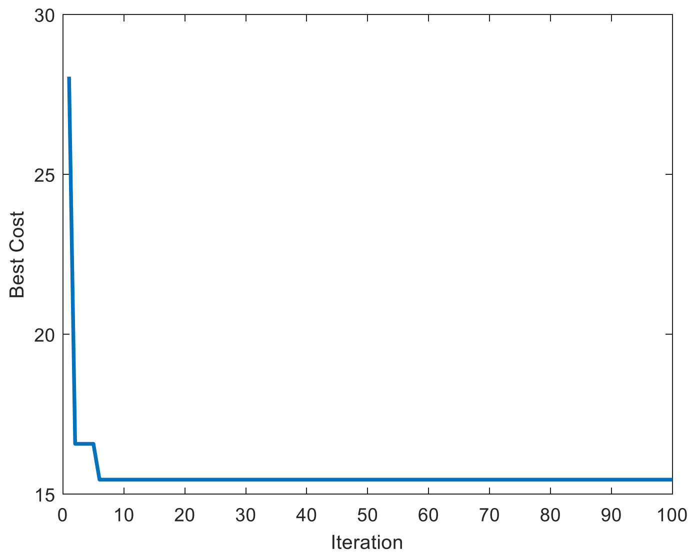

Figure 7. ACO-Cost Minimization Route

However, an optimal route of cost minimization for choosing the optimized solution by ACO is shown in Figure 7. The algorithm achieved its minimum cost around iteration 10.

The result of ACO is comparable with that of stepwise regression (stepwise forward) in terms of identifying the main key sector. Although, both methods found the same sector (manufacturing) as the most significant sector for growth over the sample size but stepwise regression technique has shortcomings, especially in the presence of collinearity (the symptoms are apparent in correlation coefficients in Table 6 [12]) and the narrow confidence intervals which are discussed by many scholars (Hurvich et al., 1989; Tibshirani, 1996). Therefore, relying on the results of the stepwise method is not recommended.

Unlike the stepwise procedure, the problem of rank ordering of sectors could be solved by ACO, since the method determined the optimal path to obtain higher growth. As one can see from Table 8, the best strategy to boost development in Indonesia is investing in manufacturing; this is followed by investment in construction, hotel and restaurants, and finally agriculture. The findings of ACO algorithm are in keeping with the growth strategy published by G20 for Indonesia (2014). It suggests that to improve business environment and create more opportunities for private investors, Indonesia needs to adopt the infrastructure investment policy.

[12] Usually the building of a correlation matrix for the independent variables yields evidences of the possibility that any combination of R.H.S variables creates a multi-collinearity problem. “However, this procedure can make problems and cannot be suggested. Correlation describes a bivariate relationship, whereas collinearity is a multivariate phenomenon (Kumar, 1975:24).” Therefore, the authors applied the variance inflation factor (VIF) method to measure the level of collinearity between the regressors. which, inthe case of tradeopenness, hotels andrestaurants, and agriculture, were 9.82, 10.08, and 10.32, respectively.

Chapter 5

CONCLUSION

Since the unbalanced growth theory has not been modeled in an applicable format for Indonesia to date, this study attempted to develop a practical approach in order to reinforce a sustainable growth path. To this end, this study adopted different methodologies to examine all the sectoral effects and test the corresponding impact of different leading sectors on GDP over the period 1995-2015 for Indonesia. In addition, we employed feature selection methods employed to rank all the leading sectors in accordance with their significance degrees from the highest significant leading sector to the lowest one in the whole period.

Given the fact that, detection of yearly key sector is crucial for the economy, four different leading sectors namely agriculture, manufacturing, hotel and restaurants and construction were obtained through Hirschman-Rasmussen indices over the sample period. Then by using stepwise regression technique, linear association between the extracted sectors was covered. We found that manufacturing and construction sectors have highest impact on GDP rather than the other sectors. In a nonlinear form, MLP was applied to model the complex and nonlinear phenomenon. Shortening the nonlinear model was achieved by ACO as one of the evolutionary feature (variable) selection methods to specify the significance of key sectors.

The findings revealed that the development path could be stabilized by enforcing share of investment in the leading sectors (manufacturing, construction, hotel and restaurants and agriculture respectively), which highlights the core concept of unbalanced growth theory.

Therefore, according to the results of the linear and nonlinear models, we concluded that the effects of different sectors on growth in GDP in Indonesia were completely dependent on the structure of unbalanced growth theory. Manufacturing and construction in Indonesia, as the main key sectors selected both in linear and nonlinear forms, had more impact than the other sectors. Rank ordering of sectors has been performed using ACO, such that the best strategy to boost development in Indonesia, is investing in manufacturing; this is followed by investment in construction, hotels and restaurants, and finally agriculture. Besides, we found that MLP might be an impressive model that yielded suitable results with RMSE equal to 0.51 and R-value of 0.79, which proves the strong nonlinear pattern of development in Indonesia.

It is noteworthy that, according to Asian Development Bank (2015), the recent growth was mainly driven by increases in the services sector. The agricultural sector has continuously grown at a slower rate compared to the average GDP growth. Meanwhile, the manufacturing sector had slower growth, with an average of 4.5% from 2001 to 2010.

The results of this study also proved that Indonesia should have allocated the most investment to the manufacturing sector (the reality of their development program, according to the Asian Development Bank, indicated a lack of adequate investment in this sector) through which the country could have attained higher economic development in the examined period. It should be noted that in case of first investing in direct productive activities like manufacturing, the private investors would face problems in the absence of social overhead capital.

Altogether, our results support the validity of the unbalanced growth theory in Indonesia and provide useful and practical information for policy makers so that they understand the significant effect of manufacturing and construction activities on economic growth and development and so that they seek the expansion of investment in these industries to accelerate the growth in Indonesia.

In further research, the development path can be investigated by the effect of the key sectors exogenously on the growth path (i.e.; nonlinear autoregressive model with exogenous inputs (NARX)). Besides, to generalize the assessment of optimal growth strategy in Indonesia, a comprehensive methodology (such that the number of features is not pre-specified) can be applied by other search algorithms (i.e. Particle Swarm Optimization (PSO) and Genetic Algorithms (GA)).

Using these approaches help to test the resolution of the estimated model based on different scenarios, which depends on the number of desired variables.

REFERENCES

Amores, A. F. 2012. Dynamics in Keysector-Analysis. Sevilla, Spain: Department of Economics, *Quantitative Methods and Economic History*.

Ani, A. A. 2005. “Ant Colony Optimization for Feature Subset Selection.” Transaction on Engineering, *Computing and Technology* 4: 35–38.

Battiti, R. 1994. “Using Mutual Information for Selecting Features in Supervised Neural Net Learning.” *IEEE Transactions on Neural Networks* 5 (4): 537–550.

Bishop, C. M. 1995. *Neural Networks for Pattern Recognition*. New York: Oxford University Press.

Cai, J., and P. Leung. 2004. “Linkage Measures: A Revisit and a Suggested Alternative.” *Economic Systems Research* 16 (1): 63–83.

Chenery, H. B. (1960). Patterns of industrial growth. *The American Economic Review*, 50(4), 624-654.

Cohen, E. (2007). Industrial policies in France: the old and the new. *Journal of Industry, Competition and Trade,* 7(3-4), 213-227.

Copas, J. B. 1983. “Regression, Prediction and Shrinkage.” *Journal of the Royal Statistical Society. Series B (Methodological)* 45 (3): 311–354.

Cornwall, J., & Cornwall, W. (1994). Growth theory and economic structure. *Economica*, 237-251.

Courville, T., and B. Thompson. 2001. “Use of Structure Coefficients in Published Multiple Regression Articles: Beta is not Enough.” *Educational and Psychological Measurement* 61 (2): 229–248.

Cuello, F. A., Mansouri, F., & Hewings, G. J. (1992). The identification of structure at the sectoral level: a reformulation of the Hirschman–Rasmussen key sector indices. *Economic Systems Research, 4(4)*, 285-296.

Dietzenbacher, B. Los, R. Stehrer, M.P. Timmer and G.J. de Vries (2013), "The Construction of World Input-Output Tables in the WIOD Project", *Economic Systems Research, 25*, 71-98.

Dorigo, M. 1992. “*Optimization, Learning and Natural Algorithms*.” PhD thesis, Politecnico di Milano.

Drejer, I. 2003. *Input-Spillover Based Measures of Inter-Industry Linkage Revisited*. Denmark: Department of Business Studies, Aalborg University.

Enu, P., Osei-Gyimah, F., Attah-Obeng, P., & Opoku, C. D. K. (2013). ACHIEVING HIGHER GDP GROWTH RATES IN GHANA: WHICH SECTOR IS TO LEAD?. *Academic Research International*, 4(4), 457.

Fleming, M. (1955). External economies and the doctrine of balanced growth. The *Economic Journal*, 65(258), 241-256.

Gupta, S., Pattillo, C. A., & Wagh, S. (2009). Effect of remittances on poverty and financial development in Sub-Saharan Africa. *World development*, 37(1), 104-115.

Hausmann, R., Rodrik, D., & Sabel, C. (2008). Reconfiguring industrial policy: a framework with an application to South Africa.

Hirschman, A. 1945. National Power and the Structure of International Trade. Berkeley: *University of California Press*.

Hirschman, A. O. 1958. "The Strategy of Economic Development." New Haven, CT: *Yale University Press*.

Holz, C. A. 2011. "The Unbalanced Growth Hypothesis and the Role of the State: The Case of China's State-Owned Enterprises." *Journal of Development Economics* 96 (2): 220–238.

Hurvich, C. M., and C.-L. Tsai. 1989. "Regression and Time Series Model Selection in Small Samples." *Biometrika* 76 (2): 297–307.

Hussain, A., & Khan, A. Q. (2011). Relationship between agriculture and GDP growth rates in Pakistan: An Econometric Analysis (1961-2007). *Academic Research International*, 1(2), 322.

Imbs, J., & Wacziarg, R. (2003). Stages of diversification. *American Economic Review*, 93(1), 63-86.

Kaldor, N. (1986). Limits on growth. *Oxford Economic Papers*, 38(2), 187-198.

Kashef, S., and H. Nezamabadi-pour. 2013. “A New Feature Selection Algorithm Based on Binary Ant Colony Optimization.” In 2013 *5th Conference on Information and Knowledge Technology (IKT), 50–54. Shiraz, Iran: IEEE*.

Kashef, S., and H. Nezamabadi-pour. 2015. “An Advanced ACO Algorithm for Feature Subset Selection.” *Neurocomputing* 147: 271–279.

Ke, L., Z. Feng, and Z. Ren. 2008. “An Efficient Ant Colony Optimization Approach to Attribute Reduction in Rough Set Theory.” *Pattern Recognition Letters* 29 (9): 1351–1357.

Keynes, J. M. (1936). The general theory of money, interest and employment. Reprinted in *The Collected Writings of John Maynard Keynes*, 7.

Khashei, M., and M. Bijari. 2010. “An Artificial Neural Network (p, d, q) Model for Timeseries Forecasting.” *Expert Systems with Applications* 37 (1): 479–489.

Krishna Dutt, A., & Young Lee, K. (1993). The service sector and economic growth: Some cross-section evidence. *International Review of Applied Economics*, 7(3), 311-329.

Krishna, K., & Perez, C. A. (2005). Unbalanced growth. Canadian *Journal of Economics/Revue canadienne d'économique*, 38(3), 832-851.

Kristiansen, S. (2003). Linkages and rural non-farm employment creation: *changing challenges and policies in Indonesia* (No. 03-22).

Kumar, T. K. 1975. "Multicollinearity in Regression Analysis." *The Review of Economics and Statistics* 57 (3): 365–366.

Lee, T.-H., H. White, and C. W. Granger. 1993. "Testing for Neglected Nonlinearity in Time Series Models: A Comparison of Neural Network Methods and Alternative Tests." *Journal of Econometrics* 56 (3): 269–290.

Leontief, W. (1947). Introduction to a theory of the internal structure of functional relationships. Econometrica, *Journal of the Econometric Society*, 361-373.

Leray, P., and P. Gallinari. 1999. "Feature Selection with Neural Networks." *Behaviormetrika* 26 (1): 145–166.

Letter, I. C. 2011. From Reformasi to Institutional Transformation: A strategic Assessment of Indonesias Prospects for Growth, Equity and Democratic Governance. Harvard, MA: *Harvard Kennedy School Indonesia Program*.

Linden, M., and T. Mahmood. 2007. "Long Run Relationships between Sector Shares and Economic Growth-A Panel Data Analysis of the Schengen Region." *Keskustelualoitteita* 50: 1–36.

Liu, H., and L. Yu. 2005. “Toward Integrating Feature Selection Algorithms for Classification and Clustering.” *IEEE Transactions on Knowledge and Data Engineering* 17 (4): 491–502.

Livesey, F. (2012), “Rationales for industrial policy based on industry maturity”, *Journal of Industry Competition and Trade* (2012) 12:349–363.

Loh, W., L. Tim, and Y.-S. Shih. 2000. “A Comparison of Prediction Accuracy, Complexity, and Training Time of Thirty Three Old and New Classification Algorithm.” *Machine Learning* 40 (3): 203–238.

McNelis, P. D. 2005. Neural Networks in Finance: Gaining Predictive Edge in the Market. USA: *Academic Press*.

Meena, M. J., K. Chandran, A. Karthik, and A. V. Samuel. 2012. “An Enhanced ACO Algorithm to Select Features for Text Categorization and its Parallelization.” *Expert Systems with Applications* 39 (5): 5861–5871.

Miller, R. E., and P. D. Blair. 2009. Input-Output Analysis: Foundations and Extensions. New York: *Cambridge University Press*.

Murphy, K. M., Shleifer, A., & Vishny, R. W. (1989). Industrialization and the big push. *Journal of political economy*, 97(5), 1003-1026.

Nathans, L. L., F. L. Oswald, and K. Nimon. 2012. “Interpreting Multiple Linear Regression: A Guidebook of Variable Importance.” *Practical Assessment* 17 (9): 1–19.

Necmi, S. (1999). Kaldor's growth analysis revisited. *Applied Economics*, 31(5), 653-660.

Nurkse, R. 1953. Problems of Capital Formation in Underdeveloped Countries. New York: *Oxford University Press*.

Obasan, K. A., & Adediran, O. A. (2010). The role of industrial sector in the economic development of Nigeria. *Journal of Management and Society*, 1(2), 9-16.

Oberman, R., R. Dobbs, A. Budiman, F. Thompson, and M. Rossé. 2012. The Archipelago Economy: Unleashing Indonesia's Potential. *McKinsey Global Institute*.

Pfajfar, L., & Dolinar, A. L. (2000). Intersectoral linkages in the Slovenian economy in the years 1990, 1992, 1993 and 1995. In Key sectors in the Slovenian economy. Papers of the *13th International Conference on Input-Output Techniques*.

Perroux, F. (1955). A note on the notion of growth pole. *Applied Economy*, 1(2), 307-320.

Pudil, P., J. Novoviov, and J. Kittler. 1994. “Floating Search Methods in Feature Selection.” *Pattern Recognition Letters* 15 (11): 1119–1125.

Rasmussen, P. N. 1956. *Studies in Inter-Sectoral Relations* 15. E. Harck. Amsterdam: North-Holland.

Raymer, M. L., W. F. Punch, E. D. Goodman, L. A. Kuhn, and A. K. Jain. 2000. “Dimensionality Reduction Using Genetic Algorithms.” *IEEE Transactions on Evolutionary Computation* 4 (2): 164–171.

Rodrik, D. (2010). The return of industrial policy. project *Syndicate*, 12.

Rosenstein-Rodan, P. N. (1943). Problems of industrialization of eastern and south-eastern Europe. *The economic journal*, 53(210/211), 202-211.

Say, J. B. (1803). Of the demand or market for products. Critics of Keynesian Economics. New Rochelle (NY): *Arlington House*, 12-22.

Scitovsky, T. (1954). Two concepts of external economies. *Journal of political Economy*, 62(2), 143-151.

Singer, H. 1958. “The Concept of Balanced Growth and Economic Development: Theory and Facts.” Paper presented at the University of Texas *Conference on Economic Development*, Austin, TX.

Singer, H. W. (1964). International development: growth and change. *International development: growth and change*.

Singh, K. P., A. Basant, A. Malik, and G. Jain. 2009. “Artificial Neural Network Modeling of the River Water Quality: A Case Study.” *Ecological Modelling* 220 (6): 888–895.

Singh, A., and S. Dasgupta. 2005. “Will Services be the New Engine of Economic Growth in India?” ESRC Centre for Business Research *Working Paper* no. 310. London: ESRC.

Sonis, M., Hewings, J. D., & Guo, J. (2000). A new image of classical key sector analysis: minimum information decomposition of the Leontief inverse. *Economic Systems Research*, 12(3), 401-423.

Streeten, P. (1963). Balanced versus unbalanced growth. *The Economic Weekly*, 20, 669-671.

Tabor, S. 2015. “Constraints to Indonesia's Economic Growth.” ADB Papers on Indonesia no. 10. Manila, the Philippines: *Asian Development Bank*.

Tanaka, F. J. 2011. “Applications of Leontief's Input-Output Analysis in Our Economy.” Faculty of Economics *Journal University of Nagasaki* 45 (1): 29–96.

Teraesvirta, T., C. F. Lin, and C. W. Granger. 1993. "Power of the Neural Network Linearity Test." *Journal of Time Series Analysis* 14 (2): 209–220.

Tibshirani, R. 1996. "Regression Shrinkage and Selection Via the Lasso." *Journal of the Royal Statistical Society. Series B (Methodological)* 58 (1):267–288.

Wahyudi, S. T., and M. D. Jantan. 2010. "Regional Patterns of Manufacturing Industries: A Study of Manufacturing Industries in Java Region, Indonesia." *Philippine Journal of Development* 37: 95–115.

Wei, S. J. 2015. Asian Development Outlook 2015. Mandaluyong, Philippines: *Asian Development Bank.*

Widodo, W., R. Salim, and H. Bloch. 2014. "Agglomeration Economies and Productivity Growth in Manufacturing Industry: Empirical Evidence from Indonesia." *Economic Record* 90 (s1): 41–58.

Wilber, S. 2002. The Service Sector and Long-Run Economic Growth. Washington DC: *Georgetown University Department of Economics*.

Xue, B., M. Zhang, W. N. Browne, and X. Yao. 2016. "A Survey on Evolutionary Computation Approaches to Feature Selection." *IEEE Transactions on Evolutionary Computation* 20 (4): 606–626.

Yotopoulos, P. A., and J. B. Nugent. 1973. “A Balanced-Growth Version of the Linkage Hypothesis: A Test.” *The Quarterly Journal of Economics* 87 (2): 157–171.

Zhang, G., B. E. Patuwo, and M. Y. Hu. 1998. “Forecasting With Artificial Neural Networks: The State of the Art.” *International Journal of Forecasting* 14 (1): 35–62.

Zhang, H., and G. Sun. 2002. “Feature Selection Using Tabu Search Method.” *Pattern Recognition* 35 (3): 701–711.

APPENDICES

Appendix A. Bai and Perron (2003) Test

Bai and Perron (2003) test of multiple structural breaks with the break specification of sequential determination by 15% trimming, applied to the AR(3) models of GDP Growth of Indonesia (For details see: Bai and Perron (2003)). The identified break date (1998-Q2) is an evidence of regime changing of GDP growth in Indonesia over the period of Asian financial crisis.

Table 9. Bai and Perron's (2003) test of multiple structural breaks

Model	**Break Date**
AR(3)	1998-Q2

Appendix B. MLP Error Histogram

The corresponding plots of applied ANN are shown as follows:

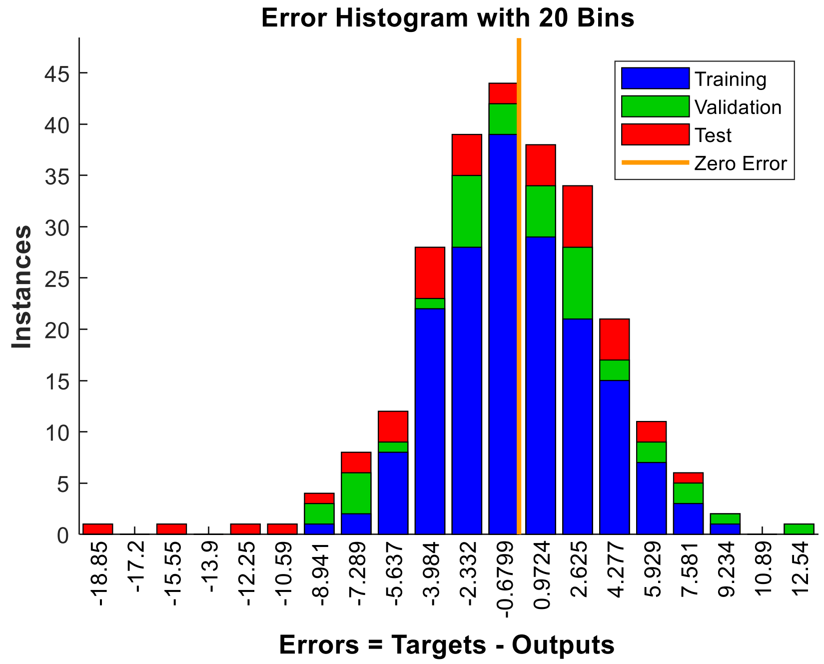

Figure 8. Error Histogram of ANN

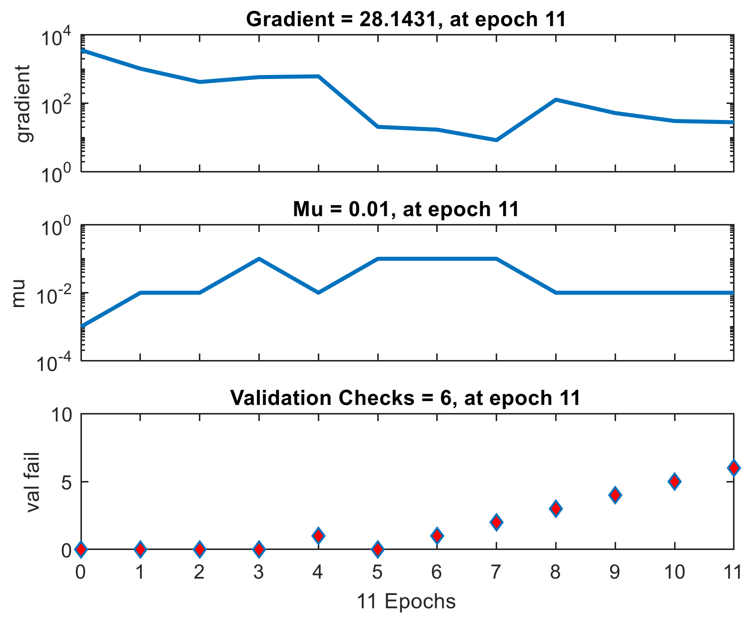

Figure9. Training State of ANN

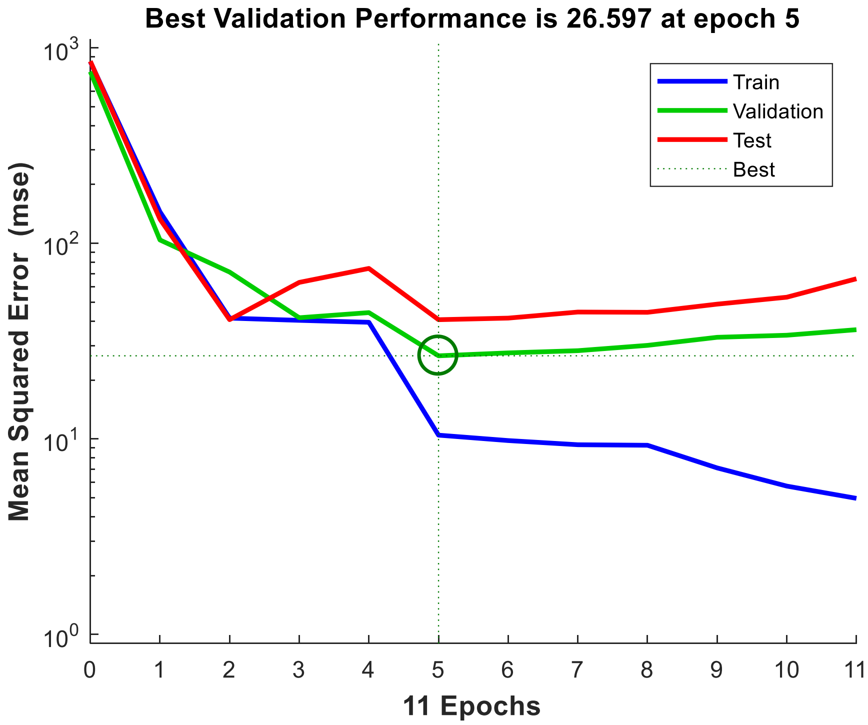

Figure 10. ANN Performance

Appendix C. Economic Structure of Indonesia

Indonesia maintained high economic growth during 1970s and 1980s due to rapid boom in oil production and microeconomic reforms implemented by the government. Indonesia joined the ranks of lower middle-income countries in 1997 by moving from a low-income country in 1965 with a 7% average annual economic growth rate. Late in the 1990s, the Asian financial crisis originated in Thailand in 1997 had a negative impact which led to a 13.6% decrease in the Gross Domestic Product (GDP) in 1998 and only 0.3% economic growth maintained in 1999. With an average annual GDP growth rate of 4.6% between 2000 and 2004, the country gradually recovered from the financial crisis.

The average annual economic growth increased to 6% between 2005 and 2011 after an impressive recovery in global commodity prices that caused a commodity boom. After the peak of the country's economic growth in the second quarter of 2010, the successive years of 2011, 2012, 2013, and 2014 witnessed a steady decline to 6.2%, 6.0%, 5.6%, and 5.0%, respectively.

However, during this depression period, the economy generally grew to 888 $ billion in 2014, with a steady increase in per capita income from 1,643$ in 2006 to 3,523$ in 2014. The average economic growth of 5.7% between 2009 and 2014 was lower than the all-time high of 7% experienced before the financial crisis. While there is an increase in the GDP share of private consumption, the GDP of total investment had decreased (ADB, 2015). The service sector largely drove this growth, whereas the manufacturing sector has been inconspicuous. Sluggish economic growth has contributed to decelerating poverty reduction. Indonesia's overall poverty between 2006 and 2010 fell by 4.5% (from 17.8% to 13.3%) , but this rate declined by just 1.5% between 2011 and 2014 (from 12.5% to 11%) (ADB, 2015). The table below presents the GDP growth of Indonesia from 2009 to 2014.

Table 10. Indonesia GDP 2009-2014

	2009	2010	2011	2012	2013	2014
GDP (in billion dollar)	578.1	755.6	845.5	878.7	871.1	888.4
GDP(annual percentage)	4.7	6.4	6.2	6.0	5.6	5.0
GDP per capita	2498	3250	3667	3761	3663	3523

Source: IMF, BPS-Statistics Indonesia, and ADB estimate (Letter, 2011)

Appendix D. Input Output Table (List of Sectors)

Industry name
Agriculture, hunting, forestry and fishing
Mining and quarrying
Food, beverages and tobacco
Textiles and textile products
Leather, leather products and footwear
Wood and products of wood and cork
Pulp, paper, printing and publishing
Coke, refined petroleum and nuclear fuel
Chemicals and chemical products
Rubber and plastics
Other non-metallic mineral
Basic metals and fabricated metal
Machinery
Electrical and optical equipment
Transport equipment
Manufacturing
Electricity, gas and water supply
Construction
Sale and repair of motor vehicles and motorcycles; retail sale of fuel
Wholesale trade, except of motor vehicles and motorcycles
Retail trade and repair, except of motor vehicles and motorcycles;
Hotels and restaurants
Inland transport
Water transport
Air transport
Other supporting transport activities
Post and telecommunications
Financial intermediation
Real estate activities
Renting of machinery & equipment and other business activities
Public administration and defense; compulsory social security
Education
Health and social work
Other community, social and personal services
Private households with employed persons

Note: The data in the WIOTs are expressed in millions of dollars. Exchange rates used to convert national values into US$. Source (Dietzenbacher et al. 2013)