ESTIMATING LONG-RUN ELASTICITIES OF RURAL WAGE RATE DETERMINANTS IN INDONESIA: THE JOHANSEN COINTEGRATION METHOD

by

FAISAL R. HARAHAP

B.S., Arizona State University, 1991
M.A., The University of Oklahoma, 1993

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES
(Agricultural Economics)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

August 2012

© Faisal R. Harahap, 2012
Abstract

This thesis examines the rural-urban linkage in Indonesia by investigating the degree of integration between the rural labor market and the urban labor market in Indonesia. It focuses on finding the determinants of rural wage rates in Indonesia, which from policy perspective is especially important since most Indonesian poor work as agricultural laborers in the rural area and are net consumers of rice; and hence any policy intervention directly or indirectly affecting rural wage rates is likely to affect the poor. Long-run relationships among real agricultural wage rates and real urban wage, agriculture prices, and real urban GDP are empirically studied using the Johansen cointegration framework. The estimates of the long-run elasticities are obtained based on 1983-2009 quarterly data compiled from government official statistics.

Results for three most populated Indonesian provinces (West Java, Central Java and East Java) show that urban variables (ie. real urban wage rates and urban GDP) are much more important than rural variables (ie. rice price) in determining the long-run rural wage rates. The long-run elasticities for real urban wage rates are between 0.14 and 0.53, with an average elasticity of 0.31, while the agriculture (rice) price has elasticities merely between zero and 0.17. Furthermore, the long-run elasticities of real urban GDP for West Java, Central Java, and East Java are 0.14, 0.19, and 0.19, respectively.

The policy implications are significant. For instance, the long-run impact of government protectionist trade policy by raising rice prices to help the poor is likely to be very limited. A
25% import tariffs, for example, would have raised rural wage rates in Java on average by merely 2.6%, a negligible increase, given it’s only a once-and-for-all increase. Conversely, using a similar argument, if the government decided to liberalize trade policy in rice by completely dismantling tariff barriers, there would be virtually no effect on rural poverty, as measured by the rural wage rate. In fact, lower rice prices would translate into increased food security for poor consumers in urban areas, and even in rural areas where almost all the truly poor are net consumers of rice.
Table of Contents

Abstract ........................................................................................................................................... ii

Table of Contents .............................................................................................................................. iv

List of Tables ...................................................................................................................................... vi

List of Figures .................................................................................................................................... vii

List of Abbreviations ....................................................................................................................... viii

Acknowledgements .......................................................................................................................... ix

Dedication .......................................................................................................................................... xi

Chapter 1: Introduction .................................................................................................................... 1

1.1 Background and rationale ......................................................................................................... 1

1.2 Problem statement and research question ............................................................................. 6

1.3 Thesis organization ................................................................................................................... 9

Chapter 2: Survey of literature ..................................................................................................... 10

2.1 Relevant empirical studies on Indonesia .............................................................................. 10

2.2 Relevant empirical studies on other countries ................................................................... 13

Chapter 3: Theoretical framework ............................................................................................... 16

3.1 Harris-Todaro migration model ............................................................................................. 17

3.2 Rural wage determination model ......................................................................................... 21

Chapter 4: Methodology ............................................................................................................... 27

4.1 Unit root test ............................................................................................................................. 32

4.2 Unit root test in the presence of structural break ............................................................... 36

4.3 Cointegration test .................................................................................................................... 37
4.4 Cointegration vector identification ................................................................. 41

Chapter 5: Data ............................................................................................................. 44

Chapter 6: Empirical results and discussions .......................................................... 54

6.1 Unit root test results: ADF, PP and DF-GLS tests ............................................. 54
6.2 Confirmatory unit root test results: KPSS test ................................................. 57
6.3 Unit root with structural break tests results ...................................................... 59
6.4 Cointegration test results .................................................................................... 61
6.5 Cointegration vector identification .................................................................... 65
   6.5.1 Cointegration vector identification results for West Java .......................... 66
   6.5.2 Cointegration vector identification results for Central Java ....................... 67
   6.5.3 Cointegration vector identification results for East Java ............................ 68

Chapter 7: Conclusion ................................................................................................. 71

7.1 Summary ............................................................................................................... 71
7.2 Policy implications ............................................................................................... 74
7.3 Thesis limitations and recommendations for future studies ............................. 77

References .................................................................................................................. 80

Appendix ..................................................................................................................... 95

Mathematical derivation of Harris-Todaro model .................................................... 95
List of Tables

Table 6.1: Results of non-stationarity tests for the time-series variables..........................55
Table 6.2: KPSS confirmatory unit root test (observation period: 1983:Q1 - 2009:Q1)........58
Table 6.3: Unit root tests in the presence of structural shift using LLS method and critical
values ......................................................................................................................................60
Table 6.4: Johansen cointegration test results for West Java..............................................63
Table 6.5: Johansen cointegration test results for Central Java...........................................64
Table 6.6: Johansen cointegration test results for East Java................................................65
List of Figures

Figure 4.1: Framework of analysis for assessing long-run rural wage rates ..................31

Figure 5.1: Trends in Indonesian real manufacturing wage rates and real agricultural wage rates in West Java, Central Java, and East Java (2007 = 100) ..............................48

Figure 5.2: Trends in rice price indices in West Java, Central Java, and East Java (2007 = 100) .................................................................................................................................49

Figure 5.3: Trends in real manufacturing GDP (2000 = 100), urban employment, and urban employment rates in Indonesia ...............................................................................................50

Figure 5.4: Trends in Indonesian Real Manufacturing GDP (2007 = 100) ....................51

Figure 5.5: Real rural wage rates, real urban wage rates, agricultural prices, and real urban GDP in West Java (1983:Q1 – 2009:Q1) (2007 = 100) ..........................................................52

Figure 5.6: Real rural wage rates, real urban wage rates, agricultural prices, and real urban GDP in Central Java (1983:Q1 – 2009:Q1) (2007 = 100) .........................................................53

Figure 5.7: Real rural wage rates, real urban wage rates, agricultural prices, and real urban GDP in East Java (1983:Q1 – 2009:Q1) (2007 = 100) .........................................................53

Figure 6.1: Logarithm of agriculture price in West Java (2007=100) ..........................57
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF</td>
<td>Augmented Dickey-Fuller</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike Information Criterion</td>
</tr>
<tr>
<td>AR</td>
<td>Autoregressive</td>
</tr>
<tr>
<td>ARIMA</td>
<td>Autoregressive Integrated Moving-Average</td>
</tr>
<tr>
<td>BPS</td>
<td>Central Bureau of Statistics</td>
</tr>
<tr>
<td>BULOG</td>
<td>Badan Usaha Logistic (National Logistic Agency)</td>
</tr>
<tr>
<td>CPI</td>
<td>Consumer Price Index</td>
</tr>
<tr>
<td>DF-GLS</td>
<td>Dickey-Fuller Generalized Least-Square</td>
</tr>
<tr>
<td>ECM</td>
<td>Error-Correction Model</td>
</tr>
<tr>
<td>GDP</td>
<td>Gross Domestic Product</td>
</tr>
<tr>
<td>KPSS</td>
<td>Kwiatowski, Phillips, Schmidt &amp; Shin</td>
</tr>
<tr>
<td>LLS</td>
<td>Lanne Lutkepohl and Saikkonen</td>
</tr>
<tr>
<td>LM</td>
<td>Langrange Multiplier</td>
</tr>
<tr>
<td>LR</td>
<td>Likelihood Ratio</td>
</tr>
<tr>
<td>MDG</td>
<td>Millennium Development Goals</td>
</tr>
<tr>
<td>OLS</td>
<td>Ordinary Least-Square</td>
</tr>
<tr>
<td>PP</td>
<td>Phillips and Perron</td>
</tr>
<tr>
<td>SAKERNAS</td>
<td>Survey Tenaga Kerja Nasional (National Labor Force Survey)</td>
</tr>
<tr>
<td>SIC</td>
<td>Schwarz Information Criterion</td>
</tr>
<tr>
<td>SUSENAS</td>
<td>Survey Sosial Ekonomi Nasional (National Socio-Economic Survey)</td>
</tr>
<tr>
<td>VAR</td>
<td>Vector Autoregressive</td>
</tr>
<tr>
<td>VECM</td>
<td>Vector Error-Correction Model</td>
</tr>
</tbody>
</table>
Acknowledgements

This thesis would not have been possible without the collaborative support of many individuals during my study at UBC. I offer my enduring gratitude to the faculty, staff and my fellow students at UBC, who have inspired me to continue my work in this field.

I owe particular thanks to my thesis supervisor and personal mentor, Professor Richard Barichello, for inspiring me to question more deeply on the topics of rural-urban migration and agricultural wage rates determination in Indonesia, and for enlarging my general vision in applied economics and providing coherent answers to my endless questions. Similarly, I am very grateful to all other members of my thesis and examining committees: Professor Daniel Gordon of the Department of Economics at the University of Calgary, and Dr. Sumeet Gulati of the Faculty of Land and Food Systems at the University of British Columbia. In particular, Professor Gordon has generously provided valuable review and suggestions on my use of applied time-series econometrics methods. I also thank Professor James Vercammen for chairing my thesis defense committee, and Dr. Florencia Jaureguiberry for being an external examiner in my thesis defense.

I am also highly indebted to the extensive and able research assistance from Andre Tampubolon, especially during the critical stage of data collection and compilation process, which involved his frequent visits to the office of Central Bureau of Statistics (BPS) in Jakarta, Indonesia. In the end, I believe both Andre and I have learned much and rewardingly
harnessed new sets of technical as well as people skills from this experience. Generous help from the BPS staff, particularly Mr. Hadi and Ms. Dewi, is also highly appreciated.

I would like also to express my gratitude to the Faculty of Land and Food Systems and the Institute of Asian Research at UBC for providing financial support during my research through various teaching assistantships, research assistantships, scholarships as well as research funding.

Last but not least, special thanks go to my family, who has supported me throughout my years of education, both morally and financially.
To Josephine
Chapter 1: Introduction

1.1 Background and rationale

One of the most universal human aspirations is to achieve freedom from poverty. Despite the dramatic increase in the global living standards the word economy has experienced since the end of World War Two, about 1.2 billion people, or about a quarter of the world’s population, still live under the poverty line, with the majority living in the rural area by the turn of the last Millennium. The Millennium Development Goals (MDGs), signed by heads of states in September 2000, constitutes the boldest commitment by the international community to alleviate the global poverty problem.

The poverty in Indonesia deserves special attention. Part of the dynamic region of East Asia, the Indonesian economy has come a long way from a near-collapse in the late 1960s to becoming one of the world’s fastest growing economies. Between 1976 and 2009, the Indonesian economy grew at an annual average of 5.7 percent, primarily fueled by rapid growth in the industrial manufacturing sector. Indonesia’s remarkable economic development in the past four decades has increased living standard, as reflected by a recent move from World Bank’s lower-income country status to middle-income country status. Nonetheless, Indonesia remains a relatively poor country. In 2009, there were 32.5 million Indonesians living below the national poverty line of $1.55-a-day, or about 14.2 percent of the population. While the 2009 figure should be considered a substantial improvement compared to 54.2 million people in 1976 (about 40.1 percent of the total population), it is still high especially if
we take into account that in 2009 there were as many as 50.6 percent of all Indonesian still living under $2-a-day, a profile that resembles more that of lower-income countries than middle-income countries.

While the more than three decades of rapid economic growth has significantly reduced the number of poor people in Indonesia, poverty remains persistent not only among migrant workers in the urban areas who fail to enter the formal sectors, but also among many agricultural laborers who remain in rural areas. In fact, poverty in Indonesia is still heavily concentrated in the rural area. In 2009, 20.6 million poor people reside in the rural area, which is about 63.4 percent of the total number of poor people in the country. Most of the poor people in rural areas are agricultural laborers without, or with very limited, landholdings and access to capital. The agricultural laborers are net consumers of agriculture crops, and they earn their living primarily from providing labor services through the rural labor market to survive. While the rural poor are likely to be engaged in one or more low productivity farming activities, many also generate and/or combine their household income from non-farming activities in the villages as well as by migrating to nearby urban city centers. Many of these migrant workers are able to secure jobs in the better-paying formal sectors. However, many more are only able to find seasonal or semi-permanent jobs in the informal sectors.

_______________________________

1 Source: BPS and World Development Indicators.
2 BPS
3 For example, “Bank Dunia tidak pahami soal beras [The World Bank misunderstands the rice problem],” Kompas, 16 November 2006, and “Kemiskinan dan harga beras [Poverty and rice prices],” Bisnis Indonesia,
Clearly any effective general poverty reduction policy should be based on a thorough understanding on the dynamic interplay among various variables in the rural labor markets and their linkages to the urban labor markets. Traditional empirical literature in urban planning, demography and environmental studies has viewed strong rural-urban linkage rather negatively as an undesirable “side effect” of development. Higher population density in urban areas is expected to create social and physical pressures that can lead to conflicts, pollutions and health problems. Rapid urbanization can also lead to detrimental environmental consequences (Brockerhoff and Brennan, 1998; Brennan, 1999; and Goudie, 2006). In developing countries, the typical policy solution has been to put strict controls on migration rather than to provide better infrastructure and facilitate the rural-urban migration flow. For example, the Hukou system in China institutionalized formal control on urban citizenship (Chan and Zhang, 1999; MacKenzie, 2002; and Zhang and Wang, 2010). As a result, recent migrants in Chinese cities are frequently deprived of basic rights to social services such as health care, education, etc., which would ensure stable living conditions for themselves and their families in the urban areas.

More recently, however, a strong rural-urban linkage has been thought to have an important role in poverty reduction. Aggregate-level empirical studies in economics on Ghana, Bangladesh, and the Philippines (Palmer-Jones and Parikh, 1998; Abdulai and Delgado, 2000; Lasco et al., 2008) suggest a strong rural-urban linkage can have significant positive effect on agricultural wages in rural areas, where most of the poor people reside. The fact that most of the poverty in developing countries is in rural areas, where there are abundant supplies of labor with relatively low agricultural wages, suggests a strong rural-urban linkage
can actually provide a breathing space for the depressed rural labor market. In theory, individual's decision to migrate to the urban areas is based on the positive difference in the expected earnings in the formal urban sector and expected earnings in his or her village (Todaro, 1969, 1976, and 1986; and Harris and Todaro, 1970). Hence, the higher the real wages migrant workers earn in the urban areas, the more positive are the poverty reduction impacts for the rural areas. Particularly, rural labor can be pulled into urban areas by strong growth in manufacturing sector (Deshingkar, 2006). Furthermore, as suggested in the Vietnamese experience, even access to the less formal sectors such as in the case of seasonal migration can be welfare enhancing (de Brauw and Harigaya, 2007).

Recent data for Indonesia suggest employment opportunities in the urban areas have provided an important path for the rural poor to get out of poverty. In fact, Mason and Baptist (1996) show the labor market has had an increasing role in reducing poverty in Indonesia since the mid-1980s. The wage labor markets are expected to have larger poverty reduction impacts on rural areas (where majority of the poor reside), as the economy continues to undergo structural change, and as the workforce continues to move out of the agriculture sector into the manufacturing or services sector. Indeed, Indonesia provides a useful case study due to high levels of internal migration since the early Suharto years, especially from the rural areas to the urban areas. The rural-urban migration is aided by relatively dense population; moderately rapid urban economic growth for more than three decades; slower growth in the agricultural sector; moderately skilled rural labor force; and very few migration restrictions. These factors lower costs and raise benefits from migration activities, which is somewhat unique among low-middle income developing countries.
In the period between 1976 and 2009, the Indonesian industrial manufacturing sector grew at an annual average of 9.2 percent, increasing its share of national income from merely 10.4 percent in 1976 by almost threefold to 27 percent in 2009. Similarly, in 2009, about 49 percent of the total employment in Indonesia is in the manufacturing and other industrial sectors, as compared to merely 34 percent in 1976. In contrast, the agricultural sector grew only by an annual average of 3.3 percent during the same period, while its share of national income declined almost by half from 29.7 percent in 1976 to 15.8 percent in 2009 (World Bank, 2010). The rapid industrialization during this period was accompanied by substantial rural-urban migration. Between 1976 and 2009, urban population grew at an average of 4.6 percent per year, while rural population rose by a barely positive average of 0.1 percent per year. Moreover, 53 percent of Indonesians lived in urban areas in 2009, as compared to 20 percent in 1976 (World Bank, 2010). As in many newly industrializing countries in Asia during that period, a dramatic rise in living standards in Indonesia happened with substantial rural-urban migration (World Bank, 2009).

If the poor people in rural areas are indeed rural laborers and net consumers of agricultural crops, then the rural wage rate is expected to be the most important variable that measures their well-being. Improvement in the rural wage rate is expected to reduce rural poverty. As a consequence, an empirical investigation on the determinants of the rural wage rate rates is necessary. In fact, such empirical examination not only will provide us with a better understanding about the rural-urban labor market linkage in Indonesia, but also help us resolve the on-going debate on whether the government should intervene or liberalize the rice
market. Specifically, this thesis is aimed at estimating long-run elasticities of major variables that determine the rural wage rates in Indonesia. These variables are expected to include rural variables as well as urban variables.

1.2 Problem statement and research question

The main objective of the thesis is to provide a detailed examination on the Indonesian rural-urban labor linkage. Specifically, the interest is in measuring the degree of integration between the rural labor market and the urban labor market in Indonesia. One important way to measure the degree of rural-urban labor market integration is by examining the process of wage transmission between the labor markets. Wage changes in one market are transmitted to the other market at certain speed of adjustment and degree of completeness, which altogether determine the degree of market integration. Theoretically, the wage transmission concept here is grounded on the Law of One Price and the standard spatial price determination models (Rapsomanikis et al., 2006; Enke, 1951, Samuelson, 1952; Takayama and Judge, 1971). Applied to labor markets, this concept implies in an integrated rural-urban labor market, a co-movement of agricultural wage rates and urban wage rates is expected. Clearly the degree of rural-urban labor market integration, as that of other market integrations, is a long-run issue. While in the short-run rural wage rates and urban wage rates may diverge significantly, in an integrated rural-urban labor market, we should expect that in the long run both wages will tend to move together despite any possible short-run distortions.
The thesis is aimed at investigating the long-run relationship between rural wages rates and selected major urban and rural variables in Indonesia. The following are the main research questions the thesis intends to address:

(1) How do equilibrium rural wage rates respond to changes in rural variables and urban variables in the long-run?

The focus here is on to what degree the rural labor market is integrated with the urban labor market and whether there is a significant co-movement among rural variables (e.g. agricultural wage rates, agricultural prices, etc.) and urban variables (e.g. urban wages, urban employment, etc.).

(2) Do changes in the rural variables have more significant impact on the equilibrium rural wage rates than the change in the urban variables, or it is the other way around?

The key issue here is the on-going debate on rice price policy in Indonesia. Despite negative effect of higher rice prices on consumers, successive Indonesian administrations have used various trade policy measures to raise rice prices. Rhetoric used by policy makers as well as local commentators to rationalize such a policy is to help reduce rural poverty, based on the assumption that most poverty rests with rice farmers (with landholdings) who are net producers of rice.
and that poverty among net consumers of rice is not so important. Therefore, it is assumed that raising rice prices through trade tariffs will improve welfare of the rural poor. By contrast, McCulloch (2008) found a large majority of the population in rural areas consume more rice than they produce and most are therefore harmed by higher rice prices.

Do equilibrium rural wage rates respond differently to changes in rural variables and urban variables across provinces in Indonesia?

Clearly the above research questions focus on the equilibrium or long-run rural wage rates, which is a key indicator in measuring rural poverty, given that poor people own little capital, and rely heavily on their only asset, their own labor. What determines the rural wage rates in the long-run? This is the key question for poverty reduction.

The answers to the above questions should have important policy implications on efforts to reduce poverty in Indonesia. It should also provide valuable quantitative information that may enlighten our understanding of the linkage between the rural labor market and the urban labor market at the provincial-level in Indonesia. The results of the study should also be a useful guide in assessing the effectiveness of the already on-going various poverty reduction programs in Indonesia, by which an evaluation can be made as to whether each of such programs provides a sustainable path out of poverty.

---

1.3 Thesis organization

Following the Introduction, Chapter Two provides a survey of the literature on a number of relevant empirical works on Indonesia as well as other developing countries. Chapter Three discusses theoretical issues and analytical framework used in this study. Chapter Four provides a detailed description of the methodology including the applied econometric techniques employed in the empirical study. Chapter Five briefly describes the data used in this study along with some descriptive statistics. Chapter Six discusses the empirical results from the econometric estimates. Finally, a full summary, policy implications, and recommendations for future studies are given in the last chapter.
Chapter 2: Survey of literature

2.1 Relevant empirical studies on Indonesia

Robust answers to the research questions in Section 1.2. require formal econometric estimation on a theoretically grounded model. While there are many empirical works examining the determinants of rural wage rates in a number of developing countries, not many of such studies have been done on Indonesia. The quite extensive literature on employment in Indonesia mostly does not address the determinants of wages through formal model econometric estimations (Lim, 1997; Islam and Nazara, 2000; and Suryadarma, Suryahadi and Sumarto, 2007). Some other empirical works make attempts to analyze the long-run trend of wages in Indonesia, but lack serious application of econometric techniques. For example, Manning (2004) reviews real wage trends between 1971 and 1993 in the context of various episodes of economic growth and labor market developments. While the study provides enlightening descriptive results, it lacks the econometric estimates necessary for making solid policy inference. Other similar descriptive studies on Indonesian wage trends include those of Naylor (1990) and Mazumdar and Sawitt (1986).

The few available econometric wage studies on Indonesia have exclusively focused on urban or migrant wages and used primarily cross-sectional survey data and an econometric approach with little attempt to uncover the long run aspects of the wage trends using time-series data and appropriate time-series econometric methods. These studies have primarily relied on the traditional, static human capital investment model (Mincer, 1974) for estimating
the determinants of various wage variables, including government and private sector salaries in urban Java (Byron and Takahashi, 1989); government wages (Filmer and Lindauer, 2001); urban wages (Comola and De Mello, 2009; Pirmana, 2006); manufacturing wages (Sjöholm and Lipsey, 2006); and urban migrant wages (Alisjahbana and Manning, 2010).

Byron and Takahashi (1989), using the National Socio-Economic Survey data (SUSENAS), analyze the earnings determinants of government and private sectors’ employees in Javanese cities, and find experience, years of schooling, and gender difference are significant determinants of earnings. They also conclude that the private sector has a higher rate of return to schooling, while the public sector rewards experience more generously.

In a similar line of study, Filmer and Lindauer (2001) use the 1998 National Labor Force Survey (SAKERNAS) and 1999 SUSENAS data to obtain estimates of earnings determinants for workers in public and private sectors. On the contrary to the common belief that the relatively low pay among Indonesian civil servants has been the primary factor behind rampant corruption among government officials, they find government workers with secondary education or less, representing three-quarters of the civil service, earn a pay premium over private sector workers with similar educational qualifications. On the other hand, those government officials with more than a high school education indeed earn less than their private sector counterparts, but the private sector premium is far smaller than previously thought.
Utilizing four annual SAKERNAS data to estimate earning differential among groups of workers, Pirmana (2006) concludes that education, experience, along with socio-demographic factors and place of residence are powerful determinants of individual earnings, and that only 42 percent of the earnings differential between males and females is caused by differences in individual characteristics. Similarly, Comola and De Melo (2009) use 1996-2004 annual SAKERNAS data to estimate the determinants of earnings in Indonesia. One of their distinctive findings is that workers with higher levels of educational attainment are most likely to find a job in the formal sector, and that the informal sector is perceived by those workers who cannot obtain a job in the formal sector as an alternative to inactivity.

Furthermore, Sjoholm and Lipsey (2006) use the longitudinal Manufacturing Survey data to study the relationship between manufacturing wages in domestically-owned plants taken over by foreign firms. They find wages in such plants increase sharply between the year before takeover and two years after take over, relative to plants remaining in domestic ownership. An econometric analysis of the whole panel suggests that both foreign ownership throughout the period and foreign takeover resulted in higher wages relative to domestically-owned plants.

Most recent work by Alisjahbana and Manning (2010) focuses on finding the wage determinants of migrant workers in four major Indonesian cities, namely Medan, Tangerang, Samarinda, and Makasar. Using the 2008 data from the on-going, 5-year longitudinal survey of Rural-Urban Migration in Indonesia, the study provides earnings determinants estimates for three cohorts of urban workers, namely recent migrants, long-term migrants, and urban non-migrants. Although the sample of the survey does not have information that would
enable the study to examine the cohort of rural non-migrant workers, it concludes that the recent migrants earn as well as the urban non-migrants. Moreover, the long-term migrants do significantly better than the other two cohorts of urban workers. It also argues that rural-urban migration is an inevitable outcome of a rapid industrialization that not only provides a path to poverty reduction, but also a further momentum for the next stage of economic development.

2.2 Relevant empirical studies on other countries

Some notable applied econometric works that address the agricultural wage determination have attempted to directly find an explicit long-run relationship between agricultural wage rates and its determinants by employing the “new dynamic model” suggested by the cointegration concept and error-correction model (ECM). While there has been no study applying the cointegration and/or the ECM approach on Indonesian wage time-series data, there are several of such studies on other developing countries.

Palmer-Jones and Parikh (1998) investigate the existence in Bangladesh of a long-run relationship between agricultural wage rates, rice, jute, cotton prices, manufacturing wage rates and agricultural production over the period 1949-1991. Using the cointegration and ECM approach, the study finds that agricultural wages have strong positive long run relationships with rice prices, manufacturing wages and agricultural productivity.
Abdulai and Delgado (2000) examines the variables that influence real agricultural wage rates in Ghana based on 1957-1991 annual data. The Johansen cointegration approach is used to reveal and quantify the long-run relationships among agricultural and urban wage rates, the domestic terms of trade between agriculture and non-agriculture, urban unemployment, capital stock in agriculture and the size of rural population. One of the main results from cointegration test and identification shows a 1 percent change in the domestic terms of trade between agriculture and non-agriculture leads to a 0.83 percent change in the real agricultural wage rate in the long-run. Furthermore, a 1 percent change in the real urban wage rates is likely to change the real agricultural wage rates by 0.56 percent in the long-run. In addition, if the urban unemployment rate rises by 1 percent, real agricultural wage rates are expected to decline by 0.29 percent in the long-run. These findings imply the significance of domestic farmer’s of terms of trade and rural-urban migration on increasing wages in the rural areas, where most of the poor reside.

Lasco et al. (2008) empirically explore the short-run and long-run relationship between rice prices and agricultural wages in the Philippines, using the neoclassical wage determination model. Employing three different empirical frameworks, namely the cointegration/ECM approach, the first-differenced approach, and the traditional OLS approach, they found that agricultural wages adjusts positively to 1 percent change in rice price with a short-run elasticity of 0.29 to 0.57 and a long-run elasticity of 0.70 and 1.0 in preferred model specification. Moreover, without the inclusion of any other variables such as the urban variables in the estimated model, the results suggest that decrease in rice price, for example
as a result of trade liberalization, is likely to lead to a net welfare loss for the households that are heavily reliant on agricultural wages for income.
Chapter 3: Theoretical framework

The research questions in Section 1.2 leads to the following hypothesis: If migration in Indonesia occurs widely and if agricultural employment is sufficiently small share of total employment, the agricultural sector may be a price-taker in rural labor markets. In such a case, rice prices and other agricultural variables would then have little effect on the formation of rural wage rates in Indonesia. Rather, rural wage rates would be determined primarily by urban variables, such as unskilled wage rate in service and manufacturing sectors in local towns and larger cities; and urban unemployment or urban economic growth.

From policy perspective, given that the majority of the rural poor are net consumers of rice, this hypothesis (if proven true) implies that any intervention to raise rice prices (e.g. through open market operations) will largely benefit land owning farmers, particularly the minority of farmland owners with large land holdings. On the other hand, a trade liberalization program that lowers rice price will not necessarily have a significant negative effect on the rural landless poor, aside from broader effect to help other net rice consumers in the urban areas.

A more formal model is necessary to make the above hypothesis testable. Before deriving a testable reduced form model for equilibrium rural wage rate determination, a brief discussion on some migration theories is useful. Arthur Lewis (1954) proposed a model in which the underdeveloped economy consists of two sectors: a traditional, overpopulated rural subsistence sector characterized by zero marginal labour productivity (“surplus labor” condition) and a high-productivity modern urban industrial sector, characterized by full
employment condition. Such less connected, dual rural and urban labor market hypothesis, however, contradicts our above hypothesis.

Von Braun (2005) asserts that, more broadly, migration can be determined by push and pull factors. Push factors may include droughts, land scarcity, low wages and absence wage labor in rural areas. Pull factors may include better job opportunities and possibility of higher wages in urban areas. Deshingkar (2006) also shows that rural labor can be pulled into urban areas by strong growth in manufacturing. Furthermore, using Vietnamese data, de Breuw and Harigaya (2007) contends that even seasonal migration can be welfare enhancing. One most notable relevant migration theory is the Harris-Todaro (1970) migration model, which allows a degree of integration between the rural and urban labor markets in its assumptions. This assumption is in fact in line with our casual observation in the Indonesian case. A further exploration on the Harris-Model migration model is necessary, as it provides a useful analytical framework for individual’s migration decision, which is consistent to our above hypothesis.

3.1 Harris-Todaro migration model

The Harris-Todaro migration model (Todaro, 1969, 1976, and 1986; and Harris and Todaro, 1970) was inspired by the attempt to explain the following phenomena: an increasing of internal migration from the rural areas to the urban areas; large rural-urban wage differential; high urban unemployment rate, and rapidly growing urban informal sector; which are commonly observed in many developing countries in the 1960-1970s (Wang et al., 2000).
The Harris-Todaro migration model is based on the neo-classical migration theory, which views migrants as individual, rational actors, who decide to move on the basis of a cost-benefit calculation. In doing so, free choice and full access to information for every migrant individual are assumed, so that these migrants are expected to go where they can be the most productive and earn the highest wages (de Haas, 2010). To elaborate, the Harris-Todaro model has four basic characteristics (Todaro and Smith, 2009):

1. Migration is motivated primarily by rational economic decision based on relative benefits and costs, predominantly financial and/or psychological.

2. The migration decision depends on expected rather than actual urban-rural real-wage differentials where the expected differential is a function of two variables: the actual urban-rural wage differential and the probability of successfully securing job(s) in the urban area.

3. The probability of securing job(s) in the urban area is directly related to the urban employment rate and thus inversely related to the urban unemployment rate.

4. Rural-to-urban migration growth rates in excess of urban job opportunity growth rates are not only possible but also rational, given the wide urban-rural expected wage differentials. Hence, high rates of urban unemployment are inevitable outcomes of the chronic imbalance of economic opportunities between urban and rural areas observed in most developing economies.
A mathematical presentation of the Harris-Todaro model is given in the Appendix.

The intuition behind the Harris-Todaro model (1970) can be summarized as follows. The minimum wage in the urban sector is set to be higher than the wage in the rural sector, which is valued at its marginal product. This results in a wage differential between the two sectors. Rural workers have an incentive to migrate to the urban areas despite urban unemployment, because of the prospect of higher earnings in the urban sector. Such migration will continue as long as there is a possibility for migrants to increase their income by moving to the urban areas. Clearly some migrants will have arranged employment before leaving the countryside. Others will begin searching for employment only after they have arrived at their destination. Some will necessarily join the pool of urban unemployment. However, even in the latter case, the presence of the migrant in the urban area may increase his/her likelihood of securing urban employment at a later date. This explains why there is a continuous flow of migrants observed in developing countries despite of the high urban unemployment rates. Thus, the Harris-Todaro migration model appear to reflect characteristics of a developing economy (Banhs, 2005).

The main point of Harris-Todaro migration model is that if the expected urban wage

(1) equals the current rural wage, there is no incentive to migrate from the rural area to the urban area.
(2) is greater than the current rural wage, there is an incentive to migrate from the rural area to the urban area.

(3) is less than the current rural wage, there is an incentive to reverse-migrate from the urban area to the rural area, given adjustments for differences in the cost of living.

The Harris-Todaro model has been criticized for being Western-centric due to the assumption that a large flow of internal migration from the rural areas to the urban areas fulfils the same facilitating role in the modernization of the current developing economies as it did in the western market economies (Skeldon, 1997). The model also assumes that migrants have perfect knowledge of the costs and benefits of migration (McDowell & de Haan 1997). It has also been criticized for largely ignoring the presence of market imperfections and other structural and institutional constraints in developing countries. For example, the model does not account constraining factors such as government restriction on migration (de Haas, 2008). However, historical examination on the characteristics of the Indonesian economy and the rural-urban migration in Indonesia seems to make these criticisms less valid, particularly on the issue of government migration restriction, which is interestingly not so significant in the Indonesian case.

Nevertheless, despite the appeal of the Harris-Todaro model’s realistic assumptions for the Indonesian case, without the ability to put the model into a more testable hypothesis, there is
no way to reject or accept the model’s hypothesis. A translation in the spirit of the original Harris-Todaro migration model is broadly pursued below.

### 3.2 Rural wage determination model

The rural wage determination model in this section is derived by following the approach described by Abdulai and Delgado (2000). The model assumes rural wage rate is determined by the interaction of the supply and demand of the rural labor within a competitive rural labor market. This assumption is justified by the observed rural labor market conditions in Indonesia as well as many other developing countries, where minimum wage legislation is largely toothless, and labor unions are practically non-existent. The aggregate demand for rural labor is specified as follows:

\[
L^d = L^d(RUW, AGP),
\]  

(3.1)

where \( RUW \) and \( AGP \) denote real rural wage rate and agriculture price, respectively. The specification of demand side of the rural labor market hypothesizes that real rural wage rate and agriculture price influence the amount of labor services producers are willing to hire in agricultural production in the rural areas. In other words, the quantity of labor services demanded in the rural labor market is a function of agriculture producers’ output and labor input price.
The Harris-Todaro migration model suggests internal migration from the rural areas to urban areas will occur if the expected urban real wage rate is greater than the rural wage rate. While this may sound as an oversimplification, Larson and Mundlak (1997) have shown that wage differentials between the sectors alone can have a large effect on intersectoral-sectoral migration. In our attempt to incorporate this idea in the formulation of the the aggregate labor supply function, we define

\[ \Delta W_{ru} = E(URW) - RUW, \]  

(3.2)

where \( \Delta W_{ru} \) indicates rural-urban wage differentials, and \( E(URW) \) and \( RUW \) denote expected urban wage rate and prevailing rural wage rate, respectively. We then specify the following equation:

\[ \Delta L' = g(\Delta W_{ru}), \]  

(3.3)

here \( \Delta L' \) represents a change in the rural labor force supply through out-migration due to short-run positive changes in the rural-urban wage differentials \( \Delta W_{ru} \). In other words, assuming migration cost (including all kinds of transaction costs) is not prohibitive, a change in the expected wage rate in favor of the urban sector is likely to increase the flow of migration from the rural areas to the urban areas in the long-run. Furthermore, this out-migration will in turn affect the aggregate supply for rural labor, which functionally can be described as:
\[ L' = h(\Delta L') \]  

(3.4)

Given equation (3.3), the aggregate supply of rural labor can also be expressed as:

\[ L' = j(\Delta W_{ru}) \]  

(3.5)

and given equation (3.2), we have

\[ L' = k(URW, E(URW)) \]  

(3.6)

Let us also define the following:

\[ E(URW) = l(URW, UREM) \]  

(3.7)

where \( URW \) and \( UREM \) denote the prevailing urban wage rate and the rate of urban employment (or unemployment), respectively. The variable \( UREM \) here is used to reflect the probability of obtaining employment in the urban sector. Higher employment rate (or lower unemployment rate) likely leads to a higher expected real urban wage rate, or vice versa. Hence, the expected urban wage rate can be defined as the prevailing real urban wage \( URW \) corrected by the urban employment rate (or unemployment rate), \( UREM \). Urban employment (or unemployment) rate now becomes the objective probability that serves as a proxy variable for the individual’s subjective probability of successfully securing job(s) in the urban area, making the Harris-Todaro model more directly testable.
Given equation (3.7), the aggregate supply for rural labor in equation (3.6) can be expressed as follows:

\[ L^* = L^* \left( RUW, URW, UREM \right), \]  

(3.8)

in which the aggregate supply of agricultural labor is a function of the urban wage rate, the rural wage rate, and the rate of urban unemployment.

Assuming there is a competitive equilibrium in the aggregate rural labor market, market equilibrium condition stipulates:

\[ L^* = L^d \]  

(3.9)

Substituting equations (3.1) and (3.8) into equation (3.9), and solving for \( RUW \), we have the following reduced-form real rural wage function:

\[ RUW_t = f \left( URW_t, AGP_t, UREM_t, \epsilon_t \right), \]  

(3.10)

where \( \epsilon_t \) is an error term capturing the influence of all other omitted variables in the model.

From econometric estimation perspective, the error term \( \epsilon_t \) is assumed to be stationary, or I(0). If the omitted variables were stationary as well, or I(0), we would not expect any
problem in the estimation results. However, if the omitted variables were non-stationary, say I(1), this would imply the error term \( \varepsilon \), to be I(1) as well. Specifically, from time-series econometric perspective, such underspecification of the statistical model may lead to either failure in detecting cointegration or underestimation of the cointegrating rank (Pashourtidou, 2003). For that reason, we shall pay careful attention to the behavior of the residual upon obtaining econometric estimation results.

Equation (3.10) implies that an increase in the real urban wage rate and an increase the urban employment rate (or a decrease in the urban unemployment rate) will stimulate workers in the rural areas to migrate to the urban areas, reducing the labor supply in the rural labor marker, causing a positive change in the real rural wage rate.

In addition, an increase in agricultural prices might reduce the real rural wage rate, especially since food expenditures comprises a major portion of individual total expenditures in the rural households, but it might also favor workers in rural areas, as higher agricultural prices will induce farmers to increase production, which requires more labor inputs. Moreover, with the assumption of flexible rural labor market and non-backward-bending rural labor supply curve, the higher demand for labor in the rural areas leads to higher real rural wage rates (Khan, 1984; Sah and Stiglitz, 1987).

Certainly causality directions among the variables in equation (3.10) can be an issue. Nonetheless, given the actual condition in the Indonesian rural and urban labor markets, and the results from the Harris-Todaro migration model, it is safe to deduce that the direction of
the effect goes from the real urban wage rates to the real rural wage rates; and from the urban employment rate (or the urban unemployment rate) to the rural wage rates. Furthermore, one might argue that agricultural prices may be thought as endogenous, with higher rural wage rates driving up agricultural prices. However, agricultural prices in Indonesia (particularly rice prices and food crops prices) are exogenously determined by the world market, with BULOG regular intervention primarily only for pure price stabilization (Timmer, 1996; Barichello et al., 1998; and Dawe, 2001). Therefore, overall, the right-hand variables in equation (3.10) can be restricted as to possess a weak exogeneity property.

The functional form in equation (3.10) allows us to employ econometric estimation in a more straightforward manner, and perform tests on a number of hypotheses, which would help us answer the research questions in Section 1.2. For example, with the formal estimation and test, we will be able to test our hypothesis stated in the beginning paragraph of this section, which suggests that urban variables (e.g. real urban wage rates and urban employment (or unemployment) rates are more important determinants of the real rural wage rate than the rural variables (e.g. agricultural prices), given the price-taker nature of the rural labor market. This is achieved by comparing the magnitude of the coefficient estimate for every independent variable in equation (3.10) and examining its statistical significance.
Chapter 4: Methodology

While it is tempting to apply the ordinary least-square (OLS) regression method on a linear specification of equation (3.10) and obtain the parameter estimates for every independent variable, technical advances in time-series econometrics in the last twenty-five years have suggested doing so can potentially produce inconsistent parameter estimates. This is true if any of the variable time series, such as real rural wage, real urban wage, agricultural price, and urban employment, contains a unit root, and therefore is non-stationary.

As Rao (2008) intuitively explains, the classical OLS econometric estimation method is based on the assumption, among others, that the means and variances of the variables are finite constants and independent of time. Yet recent discoveries based on formal unit root test methods have shown that many time-series variables in economics do not satisfy these assumptions. In other words, many of these economic variables are non-stationary variables. Furthermore, it has also been shown that the use of the OLS estimation method on non-stationary (or unit root) variables gives misleading inferences, known as the spurious regression problem (Granger and Newbold, 1974). From statistical viewpoint, if the means and variances of the unit root variables are not constant over time, all the computed summary statistics, in which these means and variances are used, are time dependent and fail to converge to their true values as the sample size increases, which is an important theorem used as the basis for inference calculations. Likewise, conventional tests of hypothesis with standard distributions will be seriously biased towards rejecting the null hypothesis of no
relationship between the dependent and independent variables. This is could lead to wrong inferences with serious policy implications if the null hypothesis is actually true.

Accordingly, formally pre-testing each of the involved variables for unit roots has become an important necessity in any serious applied econometric works involving time-series economic variables, as it is important to confirm or refute the existence of unit root within each of the variable time-series before any appropriate estimation strategies and procedures are adopted.

When unit roots are statistically significant in one or more time-series variables specified in the model, one way to proceed is by taking the first difference of the non-stationary time-series variables to make them stationary. Once the time-series variables are stationary, the classical OLS method can be employed as usual within the Box and Jenkins’ autoregressive integrated moving-average (ARIMA), or Vector Autoregressive (VAR), model (Greene, 2003). The first-difference approach may be acceptable, but one serious drawback of this approach is differencing may result in a loss of important information about the long-run relationship that may exist among the concerned variables (Sargan, 1964; and Hendry and Mizon, 1978). Since economics is particularly interested in the long-run equilibrium, such loss, if preventable, is not desirable.

To overcome this limitation, Engle and Granger (1987) offer the cointegration concept. They show that it is possible for a linear combination of two or more non-stationary (or unit-root) time-series variables to be stationary. If such a stationary linear combination exists, the non-
stationary time-series variables are said to be cointegrated. Hence, cointegration is the appropriate technique to estimate the equilibrium or long-run parameters in a relationship with unit root variables. For our specific case, it is useful to view cointegration as the long-run equilibrium between non-stationary urban and rural time-series variables that are integrated of the same order.

Murray (1994) provides a very intuitive explanation of the cointegration concept via a humorous example of a drunk and a dog. Each of them wanders aimlessly. The drunkard’s path illustrates a random walk process, as the drunk loses alertness due to alcohol consumption. So is the path of the unleashed dog, as each new scent that crosses the dog’s nose dictates a direction for the dog’s next step, with the last scent forgotten as soon as the new one arrives. This situation is comparable to the processes of two non-stationary time-series variables. In the illustration for the cointegration idea, the dog is assumed to be the drunk’s. Despite reduced alertness, the drunk hears her dog, and the dog hears the drunken master. As each wanders around, each also assesses how far way the other is and moves to partially close that gap. Despite the random walk the drunk’s or the dog’s path follows, if the drunk is found, the dog is unlikely to be very far away from the drunk. They both belong to each other.

It is important to emphasize that cointegration is not the same as correlation. When a set of time-series variables are correlated, they will co-move in synchrony. In the correlation case, no widening or narrowing of spreads between the two wage series is expected. In other words, they will rise and fall in tandem. By contrast, when a set of variable series is co-
integrated, on a daily basis the co-integrated series do not need to move in synchrony at all. Yet, they cannot wander off in opposite directions for very long without coming back to a mean distance eventually. Indeed, putting it more technically, in our study we are not claiming that real rural wage rates, real urban wage rates, agriculture price, and urban employment are correlated, but we are hypothesizing that the four time-series variables are cointegrated, which means there is a statistically significant long-run relationship among the four time-series variables.

Provided each of the time-series variables in equation (3.10) contains unit root, our model naturally lends itself to the cointegration concept, in which the equilibrium wage rate of the rural labor market is achieved in the long run. If a long-run equilibrium can be found in the relationship among the urban and rural variables, long-run elasticities are then estimated directly and explicitly through the estimation of cointegration equation(s), as opposed to being derived from short-run estimates as in the traditional ARIMA or VAR model. We proceed to apply the time-series econometric techniques in a sequence depicted by the diagram in Figure 4.1:
Figure 4.1: Framework of analysis for assessing long-run rural wage rates

**Unit Root Test:** test for the order of integration of real rural wage rate (RUW), real urban wage rate (URW), agriculture price (AGP), and urban employment (UREM) series (all in log forms), using ADF, PP, and DF-GLS (ERS)

- If order not the same
  - Conclude absence of labor market integration. Perform Granger causality tests
  - If 0
    - Confirmatory Unit Root Test
      - KPSS
      - Structural Break Unit Root Test: Lanne, Ludkephol & Saikkonen (2002)

- If 1
  - Cointegration Test:
    A. Determine optimal range of lag-length to be used in cointegration test with VAR model involving RUW, URW, AGP, UREM (in levels), based on:
      1. LR (likelihood ratio) test and Akaike Information Criterion (AIC)
      2. Residual Diagnostics
    B. Test the null of no cointegration among RUW, URW, AGP, and UREM (Johansen procedures).
      Use intercept and/or linear trend specification. Vary lag length based on the optimal range result from the above step.
    - If 0
      - accept "No Cointegration" null
    - If 1
      - reject "No Cointegration" null

**Cointegration Vector Identification:**
Use theory for restrictions:
- Normalization of RUW
- Weak Exogeneity of URW, AGP, and UREM
- Hypothesis testing: exclusion of URW, AGP and/or UREM across/within equation(s) (if warranted)

Assess long-run relationships (coefficients) in the rural-urban labor market interaction

No long-run relationship
Once a cointegration can be established among the included variables, by Granger representation theorem (Engle and Granger, 1987), a short-run rural wage adjustment model can be developed by means of the Vector Error-Correction Model (VECM). While from technical perspective it is interesting to further investigate the short-run adjustment of the rural wage rates and the speed of such adjustment, in this study we don’t attempt to build the short-run wage adjustment model, as the scope of the task is beyond what is necessary to answer the research questions in this thesis.

4.1 Unit root test

The very first step that needs to be done is applying unit root tests on each of the time-series variables to determine if the variable is stationary or non-stationary. The presence of unit root in a price series indicates that it is a non-stationary process. Engle and Granger’s definition for cointegration features an OLS regression involving non-stationary time-series variables that are integrated with the order one, or I(1), in which the regressions’s residual is stationary, or I(0), (1987). As a result, before any cointegration test is perfomed on our urban and rural time-series variables, we must first be certain that all of these time-series variables are indeed non-stationary and integrated of the same order. A unit root test employed on each of the concerned time-series variables provides the necessary information. Furthermore, if the unit root can be removed after a time-series variable is first-differenced, then it can be confirmed that the time-series variable is an I(1) variable (Davidson and MacKinnon, 1993; and Hamilton, 1994).
A number of unit root tests are available, each of which has its own strengths and weaknesses. We include in our empirical investigation three of the most widely used unit root tests, namely, the Augmented Dickey-Fuller (ADF) test (Dickey and Fuller, 1979), the Phillips and Perron (PP) test (Phillips and Perron, 1988), and the Dickey-Fuller Test with GLS Detrending (DF-GLS) (Elliott et al., 1996).

The ADF test provides appropriate test statistics which can be used to determine whether or not a time-series variable has a unit root, a unit root plus drift, and/or a unit root plus drift plus a time trend (Enders, 2004). As an example, a regression model without drift and without a time trend, is specified for a time-series variable $y_t$, as follows:

$$\Delta y_t = \rho y_{t-1} + \sum_{i=1}^{m} \lambda_i \Delta y_{t-i} + e_t$$ (4.1)

The $m$ lagged $\Delta y_{t-i}$ terms in the model take care of the possibility of serial correlation and heteroskedasticity in the error terms $e_t$. The number of lagged difference terms to be included is determined by applying information criterion method such as the Akaike Information criterion (AIC), Schwarz Information Criterion (SIC), etc. The system of hypotheses is $\rho = 0$ (the time-series variable $y_t$ is non-stationary) under the null hypothesis, and $\rho < 0$ (the time-series variable $y_t$ is stationary) under the alternative hypothesis. The null hypothesis is rejected if the t-statistic is smaller than the relevant critical value. If $\rho = 0$, the series $y_t$ has a unit root and is non-stationary; otherwise, $y_t$ is stationary.
The system of hypotheses of the PP test is similar to that of the ADF test. They have the same procedures for computing the t-statistic, and both share the same asymptotic distribution, and hence, critical values. The difference lies in the lack of lagged $\Delta y_{t-1}$ terms in the regression model of the PP test. As discussed before, the terms are used in the ADF regression model to account for possible serial correlation and heteroskedasticity in the error terms. In contrast, the PP test uses nonparametric statistical methods to take care of this problem, by direct modifying the test statistic. Additionally, as in the case of the ADF test, a drift (intercept) and/or a time trend can also be appended to the regression model of the PP test. One advantage of the PP tests over the ADF tests is that the PP tests are robust to general forms of heteroskedasticity in the error term. Another advantage is that there is no need to specify a lag length for the test regression.

As noted above, a constant, or a constant and a linear time trend can be appended to the basic ADF test regression in equation (4.1). For these two cases, the DF-GLS test offers a simple modification of the ADF tests, in which the data are detrended so that explanatory variables are “taken-out” of the data prior to running the test regression. This modification provides substantial improvement to the ADF test (Maddala and Kim, 1998).

A practical issue in the ADF and DF-GLS tests is in determining the length of the lagged $\Delta y_{t-1}$ terms. In this study, the lag length is specified by employing the Schwarz Information Criterion (SIC) automatic selection method with a maximum lag length of 14 (fourteen) (Schwarz, 1978). For the PP test, a Bartlett kernel spectral estimation method is specified with Newey-West bandwidth selection (Newey and West, 1994).
Furthermore, the hypothesis testing in unit root tests require critical values which are non-standard. MacKinnon (1991, 1996) provides the critical values used in the ADF and PP tests, which implement a much larger set of simulations than those tabulated by Dickey and Fuller, allowing the calculation of Dickey-Fuller critical values for arbitrary sample sizes. As for the DF-GLS test, it follows a Dickey-Fuller distribution in the constant only case, but the asymptotic distribution differs in the case when both a constant and trend are included. Simulation results from Elliott et al. (1996) provide the critical values for the latter case. In all three unit root tests, the null hypothesis is rejected for values that fall below these critical values.

Kwiatkowski, Phillips, Schmidt and Shin (1992) derive an alternative unit root test, in which the time-series variable is assumed to be stationary in the null, as opposed to non-stationary null hypothesis in ADF, PP or DF-GLS test. In other words, under the so called KPSS unit root test, the pair of hypotheses are \( H_0 : y_t \sim I(0) \) against \( H_1 : y_t \sim I(1) \). It has been suggested that tests using stationarity as null can be used for confirmatory analysis, especially when the presence of unit root(s) is inconclusive under the ADF, PP or DF-GLS test. However, Maddala and Kim (1998) do not believe in the merit of using the KPSS test as confirmatory test, since it has the same poor power properties as the ADF or PP tests. Instead, they recommend the use of the DF-GLS test, which is thought to be superior to the ADF or the PP test. In this study, we will perform the KPSS test anyway for the sake of completeness.
4.2 Unit root test in the presence of structural break

Cook and Manning (2004) show the disappointing performance of the ADF and DF-GLS tests in the presence of breaks in the time-series variables. Indeed, any shift in the time series should be taken into account in testing for a unit root because unit root tests may have very low power if the shift is simply ignored (Perron, 1989). Even the KPSS test in the above would not help as its confirmation may suffer the same low power. Visual inspection may provide a significant hint whether the time series features a shift or break. Nonetheless, a formal unit root test that allows for a structural break is necessary.

To take into account a shift in the time series, one possible approach is to assume that the shift is deterministic.

\[
y_t = \mu_0 + \mu_t + f_t(\theta)\gamma_x_t,
\]

(4.2)

where \(\theta\) and \(\gamma\) are unknown parameters or parameter of vectors and the errors \(x_t\) are generated by an AR\((p)\) process with possible unit root.

Following Lutkepohl (2004), we will consider 3 (three) possible shift functions for \(f_t(\theta)\).

First, a simple shift dummy variable for a shift date will be considered. The second shift function is based on an exponential distribution function, which allows for a non-linear gradual shift to a new level starting at a shit date. The third shift function is based on a rational function in the lag operator applied to a shift dummy for a shift date.
In our study, we follow Saikkonen and Lutkepohl (2002) and Lanne, Lutkepohl and Saikkonen (2002) to perform unit root tests for the model in Equation 4.2, which is based on estimating the deterministic term first by a generalized least-squares (GLS) procedure and subtracting it from the original series, before an ADF-type test is performed on the adjusted time-series. Furthermore, if the break date is unknown, Lanne, Lutkepohl and Saikkonen (2003) have provided a procedure for picking the most appropriate break date.

4.3 Cointegration test

In the event that the tests indicate that the series are integrated of the same order of one, or I(1)), we proceed by testing the null of non cointegration against the alternative hypothesis of one cointegrating vector by using the Johansen Maximum Likelihood procedure (Johansen 1988, 1991). A relatively simpler method for testing the presence of cointegration is the Engle-Granger Two-Step procedure. Nonetheless, it is commonly acknowledged that the statistical properties of the Johansen Maximum Likelihood procedure are generally better and the cointegration test is of higher power compared to the Engle-Granger Two-Step procedure. Dickey et al. (1991) show that the Johansen cointegration test yields more satisfactory results compared to the Engle-Granger Two-Step procedures or the Stock-Watson procedures. In fact, Maddala and Kim (1998) discuss several Monte Carlo studies that have been conducted to compare different cointegration estimation methods and provide robust evidence for avoiding the Engle-Granger Two-Step procedure.
Unlike the Engle-Granger procedure which is OLS-based, the Johansen procedure is a multivariate test based on a Vector Autoregressive (VAR) model, where all the variables are assumed to be endogenous. It allows for testing and imposing restrictions not available in the Engle-Granger Two-Step procedure. The test relies on the relationship between the rank of a matrix and its eigenvalues (or characteristic roots). As adapted in Eviews 7 User’s Guide (2009), Johansen (1991) considers a VAR of order $p$:

$$y_t = A_1y_{t-1} + \ldots + A_p y_{t-p} + Bx_t + \varepsilon_t$$  \hspace{1cm} (4.3)$$

where $y_t$ is a $k$-vector of non-stationary I(1) variables, $x_t$ is a $d$-vector of deterministic variables, $\varepsilon_t$ is a vector of innovations. Equation (4.4) may be rewritten as,

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + Bx_t + \varepsilon_i$$  \hspace{1cm} (4.4)$$

where:

$$\Pi = \sum_{i=1}^{p} A_i - I,$$  \hspace{1cm} (4.5)$$

and

$$\Gamma_i = -\sum_{j=i+1}^{p} A_j$$  \hspace{1cm} (4.6)$$

By the Granger’s representation theorem (Engle and Granger, 1987), if the coefficient matrix $\Pi$ has reduced rank $r < k$, then there exist $k \times r$ matrices $\alpha$ and $\beta$ each with rank $r$ such
that $\Pi = \alpha \beta'$ and $\beta' y_i$ is I(0). $r$ is the number of cointegrating relations (or the cointegrating rank) and each column of $\beta$ is the cointegrating vector.

The Johansen cointegration test works by estimating the $\Pi$ matrix from an unrestricted VAR and testing whether the restrictions can be rejected, as implied by the reduced rank of $\Pi$. It can be shown that for a given $r$, the maximum likelihood estimator of $\beta$ defines the combination of $y_{t-1}$ that yields the $r$ largest canonical correlations of $\Delta y_i$ with $y_{t-1}$ after correcting for lagged differences and deterministic variables when present (Johansen, 1995).

There are two test statistics in the Johansen procedure, namely, the trace test and the maximum eigenvalue test. Each of the two tests are based on likelihood ratio tests of the significance of these canonical correlations and thereby reduced the rank of the $\Pi$ matrix. Both tests can be used to determine the number of cointegrating vectors present, although they do not necessarily always indicate the same number of cointegrating relations. The trace test and the maximum eigenvalue test are described in equation (4.8) and equation (4.9), respectively, below:

$$J_{trace} = -T \sum_{i=r+1}^{n} \ln \left( 1 - \hat{\lambda}_i \right)$$  \hspace{1cm} (4.7)

$$J_{max} = -T \ln \left( 1 - \hat{\lambda}_{r+1} \right)$$  \hspace{1cm} (4.8)

Here $T$ is the sample size and $\hat{\lambda}_i$ is the $i$-th largest canonical correlation. The trace test features the null hypothesis of $r$ cointegrating vectors against the alternative hypothesis of $n$ cointegrating vectors. The maximum eigenvalue test, on the other hand, features the null
hypothesis of $r$ cointegrating vectors against the alternative hypothesis of $r + 1$ cointegrating vector vectors. Neither of these test statistics follows the standard distribution. Osterwald-Lenum (1992) provides critical values for both tests obtained by using simulation studies. In addition, Johansen and Julius (1990) assert that the maximum eigenvalue test has more power than the trace test, and therefore produce more clear-cut results.

Furthermore, despite the stationary nature of the cointegrating relations, the cointegration equations may still have intercepts and deterministic trends. In carrying out the cointegration test, an assumption regarding the intercept and the deterministic trends needs to be made. Three specifications of cointegrating equations from Johansen (1995) relevant to our study and data are considered, as follows:

1. Model 1: The level data $y_t$ have no deterministic trends and the cointegrating equations have intercepts:

   $\Pi y_{t-1} + B x_t = \alpha (\beta'y_{t-1} + \rho_0)$  \hspace{1cm} (4.9)

2. Model 2: The level data $y_t$ have linear trends but the cointegrating equations have only intercepts:

   $\Pi y_{t-1} + B x_t = \alpha (\beta'y_{t-1} + \rho_0) + \alpha' y_0$  \hspace{1cm} (4.10)

3. Model 3: The level data $y_t$ and the cointegrating equations have linear trends:
\[ \Pi y_{t-1} + B x_t = \alpha (\beta' y_{t-1} + \rho_0 + \rho_t) + \alpha_\perp y_0 \]  

(4.11)

The terms associated with \( \alpha_\perp \) are deterministic terms “outside” the cointegrating relations.

Finally, the results of the Johansen cointegration test can be quite sensitive to the choice of lag length (Enders, 2004). In this study, following the most common procedure, we first estimate a VAR model using the undifferenced (or level) data. We pick the optimal range of lag length from information provided by the likelihood ratio (LR) test and the Akaike Information Criterion (AIC) as a starting point. We then apply residual diagnostics to every lag length within the optimal range. We begin with the longest lag length and test whether it can be shortened. We choose the most optimal lag length based on the results of the residual diagnostics. Finally, we apply the optimal lag length in the proposed Johansen cointegration test.

### 4.4 Cointegration vector identification

It is possible to have more than one cointegrating vector in the system given by equation (4.4), when there are more than two time –series variables specified in the cointegrating system. Cointegrating vectors can be thought of as representing constraints that an economic system imposes on the movement of the variables in the system in the long-run (Dickey at al., 1991). Hence, from this perspective, *ceteris paribus*, the presence of multiple cointegrating vectors is desirable, since stationarity in as many directions as possible means
the economic system is more stable. On the other hand, the fewer the number of cointegrating vectors, the less constrained is the long-run relationship.

While it seems desirable to have many cointegrating vectors, obtaining precise estimates for the cointegrating vectors can be rather challenging if we do not have a unique cointegrating vector. Such problem of multiple long-run relationship is known as an identification problem (Granger, 1986). From matrix algebra perspective, the identification requires restrictions being imposed on matrix $\Pi$ in equation (4.4). Pesaran and Shin (2002) demonstrate that identification of the cointegrating vector $\beta'$ requires knowledge of $r$ and that the number of restrictions necessary to identify the cointegrating vector(s) is a direct function of the number of cointegrating vectors. In general, if the number of available restrictions $k < r^2$, the system is underidentified; if $k = r^2$, the system is exactly identified; and when $k > r^2$, the system is overidentified.

In general the complete exact (or over) identification of the system will involve a combination of four restrictions (Greenslade et al., 2002):

1. Restrictions on the cointegrating rank of $\Pi$; $r < N$;
2. Restrictions on the dynamic path of adjustment (the $\Gamma_i$);
3. Restrictions on the cointegrating vectors, $\beta$, where “$\Pi = c\beta$”; and
4. Restrictions on the exogeneity or long-run causality of the system, which will imply restrictions on $\alpha$. 

42
The actual choice of imposed restrictions is determined by the researcher’s *a priori* information, in the form predicted by economic theory, market arbitrage conditions or institutional characteristics. Furthermore, Greenslade et al. (2002) also argue that imposing exogeneity restrictions at the earliest possible stage of the model reduction process and then restricting the dynamic adjustment of the model hugely increases the power of tests of overidentifying restrictions on the long-run cointegrating vectors.
Chapter 5: Data

Obtaining long time-series wage data for Indonesia has always been challenging regardless of the sectors. Studies involving analysis of wage trends in Indonesia are therefore rare despite their importance for policy-makings. Comparing the availability and the reliability of various labor market data in Indonesia, Manning (1994) correctly calls wage data as the “anak tiri” (step child) of the Indonesian labor force statistics. Because of this constraint, it is hardly surprising that labor market studies in Indonesia have primarily concentrated on quantity and quality, but not the wages.

Nonetheless, our empirical study on the equilibrium rural wage rate determination requires time series data of some length. The availability of aggregate time-series wage, employment, and price data with a relatively long span of time for both the rural sector and the urban sector is crucial for the successful application of the Johansen cointegration framework to answer the research questions. Indeed, the greatest challenge and the most time-consuming activity in this empirical research are in securing lengthy time-series variables (ie. rural agricultural wage rates, urban wage rates, agricultural prices, and urban employment rates) required by the model specification. The sole source of data used in this study is BPS (the Indonesian Central Bureau of Statistics). Very few of the data used are available in existing digital sources at BPS. Often data from historical hard copy sources had to be compiled manually and, subsequently, digitized for statistical software database inputs.
The model is estimated using secondary data from the period 1983-2009. It is the period when data have been relatively available and reliable, and formal surveys have been regularly conducted by BPS on a monthly or quarterly basis. While it seems reasonable to use annual data as in many other similar studies on other countries, the twenty-six annual observations in our case were not long enough and would likely pose the problem of low degree of freedom. Therefore, quarterly time-series data for the period 1983-2009 are used instead.

Furthermore, the BPS data also allows us to do analysis at the provincial level rather than the country level. The availability provincial-level data is especially useful because Indonesia is a large archipelago with diverse characteristics across its 33 provinces. However, wage data for non-Java provinces have only been collected more recently (since 1991), thus relatively a much shorter span of time-series observations, which would also significantly reduce the degrees of freedom critically needed in making statistical inference. Moreover, there is data quality problem for the non-Java data, as they show less consistency compared to those for Javanese provinces. For that reason, our empirical analysis focuses on West Java, Central Java, and East Java, which are the three major provinces in the island of Java, where roughly 70% of the 240 million Indonesians reside. The rural labor market and the urban labor market in each of these three provinces is expected to be relatively more integrated, compared to other provinces in Indonesia.

In this study, real agricultural wage rates are used as a proxy for the real rural wage rate variable specified in the model. The real agricultural wage rate is obtained by deflating the nominal agricultural wage rate by the consumer price index (CPI) for rural areas, with 2007
as the base year. The nominal agricultural wage rate is the simple average of three wage rates for three different activities in the rice sector, namely hoeing, planting, and weeding. Note that in the Indonesian context, hoeing is a traditional male occupation, while planting and weeding are female occupations (Godfrey, 1993). BPS has consistently collected nominal wage rates for hoeing, planting, and weeding, as well as rural CPI since 1976 from 180 farmers in West Java, 261 in Central Java, and 261 in East Java on half-day rates through its monthly surveys of farmers’ cost structure and, later succeeded by, survey of rural producers’ and consumers’ prices. Moreover, the quarterly figure for the nominal agricultural wage rate or the rural CPI is derived from taking the simple average of three successive monthly figures within every quarter. BPS have converted any in-kind wage payments into monetary terms. In Indonesia, these in-kind payments are typically a small portion of the total wages.

Real wage rates in the manufacturing sector are used in this study as a proxy for the urban wage rate variable specified in the model. Using 2007 as the base year, the nominal manufacturing wage rate is deflated by the urban CPI to obtain the real manufacturing wage rate. The monthly urban CPI data are widely available and constantly monitored by the central bank or various other agencies in Indonesia, including BPS. We obtain the quarterly urban CPI figure by averaging three monthly figures within each quarter. Meanwhile, the quarterly nominal manufacturing wage rates are the average wage costs in various manufacturing industries collected by BPS wage surveys that have been conducted on a quarterly basis since 1981. The wage survey covers several thousands firms, mainly in the manufacturing sector. The quarterly frequency of manufacturing wage rates dictates the frequency of the dataset.
It is important to note that our manufacturing wage dataset has eight missing quarterly observations between 1992 and 1993. Due to changes in internal BPS administration and survey re-designing, no wage survey was held during this period. We performed linear interpolation to replace the missing observations. The results of the cointegration procedure can be sensitive towards interpolated data. Such linear interpolation makes the time-series more systematic and reduce the probability of finding unit roots in the series. However, we balk up from performing a formal bootstrap technique to fill in the missing observations, as we believe such linear interpolation will only have limited effect on the result of the unit root test on the series, especially given the relatively small number of missing observation (ie. 8 out of total 105 observations).

Furthermore, we were confronted by the choices to use either the manufacturing wage rates for Indonesia, the manufacturing wage rates for Jakarta, or the weighted average of manufacturing wage for Jakarta and Surabaya (Indonesia’s first and second largest cities, respectively). We have decided to choose the general manufacturing wage for Indonesia as the proxy for urban wage rate over the other choices for two reasons. First, the locations of the manufacturing firms are not just limited to the Jakarta and/or Surabaya greater urban area(s), and yet regardless wherever they are located in Indonesia, they still represent alternative employment opportunity for the workers in agricultural sector. Second, the manufacturing wage rate for Indonesia seems to be the most complete series available from BPS with a minimum number of missing observations.
Figure 5.1 shows trends in real manufacturing wage rates, and real agricultural wage rates for West Java, Central Java, and East Java between 1983 and 2009. Generally, similar movements in real wage rates are observed in the two sectors, except for few years after the 1998 Asian financial crisis. It is consistent with our main hypothesis that agricultural variables play a secondary role in determining agricultural variables.

The rice price index is used as a proxy variable for the agricultural price in the model. The rice price index data are obtained from the monthly farmers’ terms trade of trade statistics published by BPS at the provincial-level. The rice price index is based on the prices received by farmers at the farm gates. Similar to agricultural wage data, these producers’ rice prices have been collected by BPS since 1976 on a monthly basis. For comparability purpose, we
have used 2007 as the base year for the quarterly rice price index time-series variable. Furthermore, the quarterly rice price index figures are obtained by taking the simple average of three successive monthly rice price index figures within every quarter.

Figure 5.2 shows trends in rice price indices for West Java, Central Java, and East Java between 1983 and 2009. Despite differences in price index levels in each of the three provinces, the movements of the three series have generally shown similar patterns, indicating a high degree of integration in the Javanese rice markets.

Figure 5.2: Trends in rice price indices in West Java, Central Java, and East Java (2007 = 100)

For urban employment (or unemployment) rates, no data is available beyond the bi-annual frequency data. In fact, for pre-2005 period, only annual-frequency data are available. For that reason, real quarterly manufacturing GDP is used as a proxy variable. This is reasonably justified, as most manufacturing industry locations in Indonesia are in the urban or suburban
areas to which rural workers are typically attracted. Figure 5.3 shows the trends in the real manufacturing GDP, the urban employment, and the urban employment rate time-series variables, utilizing the available annual-frequency data. Furthermore, because the time-series variables are in different units, they have been converted to their natural logarithmic forms for a more convenient visual comparison. For the sake of consistency, all urban employment data have been obtained from a single source, the SAKERNAS (National Labor Survey). Note that urban employment and urban employment rate data for 1995 are missing because no annual national labor survey was held in 1995. As the graph shows, real manufacturing GDP, urban employment and urban employment rate time-series follow roughly the same long-run path, providing some further degree of visual justification for using the real manufacturing GDP time-series as a proxy for the urban employment or the urban employment rates time-series.

Figure 5.3: Trends in real manufacturing GDP (2000 = 100), urban employment, and urban employment rates in Indonesia
Every quarter BPS has published real quarterly manufacturing GDP figures as part of their Quarterly Indonesian National Income or Quarterly Indonesian Gross GDP publication since 1983 with consistent and standardized methods. In our study, the base year of the real quarterly manufacturing GDP is adjusted to 2007 for comparability purposes. Trends in real manufacturing GDP between 1983 and 2009 are presented in Figure 5.4.

Figure 5.4: Trends in Indonesian Real Manufacturing GDP (2007 = 100)

Because the data figures in each of the four time-series variables in our cointegration analysis is stated in different scale and/or units, it is helpful to convert them into natural logarithmic forms. The advantage of doing this is such transformation simplifies the interpretability, as it allows estimated the model coefficients to be viewed as long-run elasticities expressed in terms of percent change, as opposed to unit change.
Figures 5.5, 5.6, and 5.7 show three sets of 105 quarterly time-series observations of provincial-level real agricultural rural wage rates (RUW), Indonesian real manufacturing wage rate (URW), provincial-level real agricultural price (AGP), and Indonesian manufacturing GDP (UGDP) for West Java, Central Java, and East Java, respectively, covering the period from 1983:Q1 to 2009:Q1. We intend to analyze the series in each of these three graph sets with the Johansen Maximum Likelihood cointegration approach that have been discussed previously in this chapter. All the series in the graphs have been converted into natural logarithmic form.

**Figure 5.5: Real rural wage rates, real urban wage rates, agricultural prices, and real urban GDP in West Java (1983:Q1 – 2009:Q1) (2007 = 100)**
Figure 5.6: Real rural wage rates, real urban wage rates, agricultural prices, and real urban GDP in Central Java (1983:Q1 – 2009:Q1) (2007 = 100)

![Graph of real rural wage rates, real urban wage rates, agricultural prices, and real urban GDP in Central Java.](image)

Figure 5.7: Real rural wage rates, real urban wage rates, agricultural prices, and real urban GDP in East Java (1983:Q1 – 2009:Q1) (2007 = 100)

![Graph of real rural wage rates, real urban wage rates, agricultural prices, and real urban GDP in East Java.](image)
Chapter 6: Empirical results and discussions

Results of the unit root tests, cointegration tests, and cointegration vector identifications for the West Java, Central Java, and East Java provinces are discussed below. All results in this thesis are estimated with Eviews statistical software, based on the theoretical framework presented in Chapter 3, and the methods and procedures discussed in Chapter 4. Note that for a more intuitively understandable symbol, the UREM variable name has been replaced by UGDP to represent the urban GDP in the entire tables, figures and equations presented in this chapter and the next chapter. As discussed in Chapter 5, the urban GDP variable is used as a proxy variable for the urban employment rate variable in the model.

6.1 Unit root test results: ADF, PP and DF-GLS tests

The condition for cointegration stipulates all involved variables must be of the same order of integration of I(1) (Engle and Granger, 1987). Therefore, as the first step, the order of integration or stationarity of the individual time-series variable must be examined and formally tested. The ADF, PP and DF-GLS tests, discussed in Chapter 4, are employed on each time-series variable both with and without deterministic trend. All ADF, PP, and DF-GLS tests have the same null hypothesis that the time-series variable under investigation has a unit root (ie. non-stationarity), against the alternative that it does not. In general, the substantially negative values of the reported test statistic under each test lead to rejection of the null hypothesis. In addition, Schwarz’s Information Criterion (SIC) is used to automatically determine the appropriate lag-length truncation with the number of maximum lagged terms set to 14 (fourteen) in each case. This method works by trading off parsimony.
against reduction in the sum of squares. The results of the unit root tests are reported in Table 6.1:

Table 6.1: Results of non-stationarity tests for the time-series variables

<table>
<thead>
<tr>
<th>(all series are in natural logarithms)</th>
<th>ADF</th>
<th>PP</th>
<th>DF-GLS</th>
<th>ADF</th>
<th>PP</th>
<th>DF-GLS</th>
<th>ADF</th>
<th>PP</th>
<th>DF-GLS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Real rural real wages (RUW)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agricultural wage - West Java</td>
<td>-1.50</td>
<td>-1.31</td>
<td>-0.66</td>
<td>-3.06</td>
<td>-2.54</td>
<td>-3.08**</td>
<td>-6.67***</td>
<td>-6.61***</td>
<td>-6.09***</td>
</tr>
<tr>
<td>Agricultural wage - Central Java</td>
<td>-1.68</td>
<td>-1.29</td>
<td>-1.30</td>
<td>-2.98</td>
<td>-2.39</td>
<td>-2.86*</td>
<td>-5.37***</td>
<td>-6.64***</td>
<td>-4.84***</td>
</tr>
<tr>
<td>Agricultural wage - East Java</td>
<td>-1.33</td>
<td>-1.46</td>
<td>-0.60</td>
<td>-1.86</td>
<td>-2.04</td>
<td>-1.90</td>
<td>-8.03***</td>
<td>-7.97***</td>
<td>-7.45***</td>
</tr>
<tr>
<td><strong>Real urban real wages (URW)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manufacturing wages (all sectors)</td>
<td>-1.19</td>
<td>-1.14</td>
<td>-0.43</td>
<td>-2.24</td>
<td>-2.21</td>
<td>-2.21</td>
<td>-8.07***</td>
<td>-8.05***</td>
<td>-8.02***</td>
</tr>
<tr>
<td><strong>Agriculture price (AGP)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice price index - West Java</td>
<td>-0.54</td>
<td>-0.29</td>
<td>1.52</td>
<td>-2.10</td>
<td>-2.95</td>
<td>-1.84</td>
<td>-5.68***</td>
<td>-9.07***</td>
<td>-0.62</td>
</tr>
<tr>
<td>Rice price index - Central Java</td>
<td>-0.27</td>
<td>-0.26</td>
<td>1.62</td>
<td>-3.90**</td>
<td>-3.03</td>
<td>-1.84</td>
<td>-5.58***</td>
<td>-8.70***</td>
<td>-1.10</td>
</tr>
<tr>
<td>Rice price index - East Java</td>
<td>-0.20</td>
<td>-0.27</td>
<td>1.73</td>
<td>-2.15</td>
<td>-2.84</td>
<td>-1.73</td>
<td>-5.82***</td>
<td>-8.47***</td>
<td>-0.62</td>
</tr>
<tr>
<td><strong>Real Urban GDP (UGDP)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real manufacturing GDP</td>
<td>-1.46</td>
<td>-1.83</td>
<td>1.21</td>
<td>-1.86</td>
<td>-2.09</td>
<td>-1.36</td>
<td>-4.37***</td>
<td>-11.06***</td>
<td>-1.79*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Critical value</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF or PP</td>
<td>-3.49</td>
<td>-2.89</td>
<td>-2.58</td>
<td>-4.05</td>
<td>-3.45</td>
<td>-3.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DF-GLS</td>
<td>-2.59</td>
<td>-1.94</td>
<td>-1.61</td>
<td>-3.58</td>
<td>-3.03</td>
<td>-2.74</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
- ADF, PP, and DF-GLS denote the Augmented Dickey-Fuller, Philips Perron, and Elliott-Rothenberg-Stock Dickey-Fuller-Generalized Least Square unit root tests, respectively.
- For ADF, PP, and DF-GLS tests, one (*), two (**), and three (***)) asterisks indicate rejection of unit root at the 10%, 5%, and 1% level of significance, respectively.

Results from the three unit root tests are broadly consistent with the hypothesis that all the time-series variables are individually I(1) (i.e., they have the same integration order of one).
Stationary test results using ADF and PP tests convincingly show that at 1 percent level of statistical significance (or 5 percent in only 1 case), all provincial-level real rural wage rates (RUW), country-level real urban wage rates (URW), all provincial-level agriculture price (AGP), and country-level urban GDP (UGDP) variables are non-stationary (with or without trend) at the level, and stationary at the first-difference. This implies each of these rural or urban wage series is indeed an I(1) time-series variable.

Similarly, at 1 percent or 5 percent level of significance, the results of the DF-GLS unit root test reveal the levels of each of the time series are non-stationary. The first differences of all the rural wage rate series and urban wage rate series are stationary at 1 percent level of significance. However, the first difference of real urban GDP is stationary only at 10 percent level of significance. Furthermore, it is important to note that even at 10 percent significant level, the DF-GLS test on the three first-differenced agriculture price (AGP) time-series variables fail to reject the null hypothesis of non-stationarity, indicating the possibility for the rice price index time-series being an I(2) variable, instead of I(1). Nonetheless, visual inspection suggests that while the agriculture price time-series in West Java, Central Java, and East Java, are non-stationary, the behavior of their first-differenced time-series appears to be stationary (see Figure 6.1 below). Similar behavior can also be observed in the agriculture price time-series for Central Java and West Java. Accordingly, no further unit root tests are performed on the agriculture price time-series. We proceed by allowing a wider band of statistical significance for the agriculture price time-series and treat them as I(1) variables, instead of pursuing the multicointegration method suggested by Granger and Lee.
(1990) for handling a cointegration of two or more time-series variables with different orders of integration.

Figure 6.1: Logarithm of agriculture price in West Java (2007=100)

All in all, based on visual inspections and performed unit root tests, we comfortably conclude there is sufficient evidence that all of our time-series variables are integrated in the same order of one. In other words, they are all I(1) variables.

6.2 Confirmatory unit root test results: KPSS test

All ADF, PP and DF-GLS tests have non-stationarity as its null hypothesis. The KPSS test provides an alternative test to confirm the above findings by specifying stationarity as its null
Table 6.2 below provides results of the KPSS tests on same set of time-series variables.

Table 6.2: KPSS confirmatory unit root test (observation period: 1983:Q1 - 2009:Q1)

<table>
<thead>
<tr>
<th>(all series are in natural logarithms)</th>
<th>Level Without Trend</th>
<th>Level With Trend</th>
<th>First Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real rural real wages (RUW)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agricultural wage - West Java</td>
<td>0.96***</td>
<td>0.11</td>
<td>0.04</td>
</tr>
<tr>
<td>Agricultural wage - Central Java</td>
<td>0.78***</td>
<td>0.13*</td>
<td>0.08</td>
</tr>
<tr>
<td>Agricultural wage - East Java</td>
<td>0.53**</td>
<td>0.11</td>
<td>0.07</td>
</tr>
<tr>
<td>Real urban real wages (URW)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manufacturing wages (all sectors)</td>
<td>0.99***</td>
<td>0.17**</td>
<td>0.06</td>
</tr>
<tr>
<td>Agriculture price (AGP)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice price index - West Java</td>
<td>1.14***</td>
<td>0.10</td>
<td>0.05</td>
</tr>
<tr>
<td>Rice price index - Central Java</td>
<td>1.14***</td>
<td>0.11</td>
<td>0.07</td>
</tr>
<tr>
<td>Rice price index - East Java</td>
<td>1.14***</td>
<td>0.10</td>
<td>0.06</td>
</tr>
<tr>
<td>Real Urban GDP (UHDR)</td>
<td>1.12***</td>
<td>0.24***</td>
<td>0.23</td>
</tr>
</tbody>
</table>

KPSS Critical value

<table>
<thead>
<tr>
<th></th>
<th>1 percent</th>
<th>5 percent</th>
<th>10 percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 percent</td>
<td>0.74</td>
<td>0.22</td>
<td>0.35</td>
</tr>
<tr>
<td>5 percent</td>
<td>0.46</td>
<td>0.15</td>
<td>0.35</td>
</tr>
<tr>
<td>10 percent</td>
<td>0.35</td>
<td>0.12</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Notes:
-KPSS denotes Kwiatowski, Phillips, Schmidt & Shin the unit root test.
-Null hypothesis: series is stationary
-One (*), two (**), and three (*** ) asterisks indicate rejection of stationarity (no unit root)
at the 10%, 5%, and 1% level of significance, respectively.

At 1 percent level of significance, each individual time-series variable without trend
specification is indeed confirmed as an I(1) variable. These results are consistent with the
findings under the ADF, PP and DF-GLS unit root tests.
6.3 Unit root with structural break tests results

Visual inspection on each of the time-series variable in Chapter 5 gave us the impression that there might be a break or a shift in each of the time-series variables. If such break does exist, the results of of the ADF, PP or DF-GLS unit root tests may be distorted, as in such case it is less likely for the test to project the null hypothesis of non-stationarity. The KPSS test with its stationarity null hypothesis indeed provides a way to confirm the results of the ADF, PP or DF-GLSS tests. However, if both tests fail to reject the respective null hypotheses or both reject the respective null hypotheses, we still do not have confirmation. A unit root test that considers structural break in the time-series in a more formal way is given by Perron (1989) who proposes a unit root test that allows a single break in the time-series.

Recently, Lanne Lutkepohl and Saikkonen (2002) argue that the results such single break unit root test are not valid if the structural change in time series occurs in a number of periods. Besides, it would be challenging to arbitrarily choose the break date to be specified in the test. They develop a test for the unit root in time series that allows a shift or break to be spread out over a number of periods. In other words, the proposed model allows the shift or break to feature a smooth transition to a new level. Such provision for a smooth transition to a new level is highly useful for our analysis, as visual inspections on the each of time-series variables indicate no abrupt shift or break took place. Thus, if there were a shit or break at all, it must have been one with a smooth transition to a new level. In the proposed model, a more generalized shift function is specified to capture the smooth transition is considered by adding the general nonlinear form $f_{t}(\theta)$ to the deterministic term of the tested model specification, as described in Chapter 4 Section 2. Lutkepohl (2004) proposes 3 (three)
different functional forms for $f_t(\theta)$ to estimate the model and test for unit root. The test hereinafter is called the LLS unit root test. Results of the LLS unit root test is presented in Table 6.3.

Table 6.3: Unit root tests in the presence of structural shift using LLS method and critical values

<table>
<thead>
<tr>
<th></th>
<th>Shift dummy</th>
<th>Exponential shift</th>
<th>Rational shift</th>
<th># of Lag(s)</th>
<th>Suggested break date</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Real rural real wages (RUW)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agricultural wage - West Java</td>
<td>-2.35</td>
<td>-3.04**</td>
<td>-3.06*</td>
<td>1</td>
<td>2004 Q2</td>
</tr>
<tr>
<td>Agricultural wage - Central Java</td>
<td>-0.92</td>
<td>-0.87</td>
<td>-1.12</td>
<td>1</td>
<td>2008 Q1</td>
</tr>
<tr>
<td>Agricultural wage - East Java</td>
<td>-2.08</td>
<td>-1.99</td>
<td>-2.24</td>
<td>2</td>
<td>2004 Q2</td>
</tr>
<tr>
<td><strong>Real urban real wages (URW)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manufacturing wages (all sectors)</td>
<td>-0.57</td>
<td>-0.32</td>
<td>-0.03</td>
<td>1</td>
<td>1998 Q1</td>
</tr>
<tr>
<td><strong>Agriculture price (AGP)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice price index - West Java</td>
<td>-1.04</td>
<td>-1.05</td>
<td>-1.08</td>
<td>6</td>
<td>1998 Q3</td>
</tr>
<tr>
<td>Rice price index - Central Java</td>
<td>-0.65</td>
<td>-0.67</td>
<td>-0.90</td>
<td>6</td>
<td>1998 Q3</td>
</tr>
<tr>
<td>Rice price index - East Java</td>
<td>-0.49</td>
<td>-0.96</td>
<td>-1.42</td>
<td>6</td>
<td>1998 Q2</td>
</tr>
<tr>
<td><strong>Real Urban GDP (UGDP)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real manufacturing GDP</td>
<td>-2.26</td>
<td>-2.29</td>
<td>-2.24</td>
<td>4</td>
<td>2005 Q1</td>
</tr>
<tr>
<td><strong>Critical Values</strong></td>
<td>1 percent</td>
<td>5 percent</td>
<td>10 percent</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-3.48</td>
<td>-2.88</td>
<td>-2.58</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
- No linear trend is specified in the data generating process.
- One (*), two (**), and three (***) asterisks indicate rejection of unit root at the 10%, 5%, and 1% level of significance, respectively.

Break dates for each time-series variable is assumed to be unknown. The suggested break date is obtained by following the procedure in Lanne, Lutkepohl, and Saikkonen (2003). The number of lagged own terms included in the tested model is obtained using the Schwarz Information Criterion (SIC). The results of LLS unit root test presented in Table 6.3 provide further evidence for the existence of a unit root when breaks are allowed. The test values for all three test statistics for each time-series variable are all quite similar across functional
specifications as well as across time-series variables, and they do not provide evidence against unit root. The estimated results of unit root tests indicate that all variables are I(1), as the null hypothesis of non-stationary can not be rejected at conventional significance levels.

Of course, there could still be other forms of structural breaks than those considered here. For example, if the series has a deterministic linear trend function, there may be also a break in the trend slope. However, in our cases, results from various unit roots tests indicate robustly that we are not to choose a deterministic linear trend functional specification for all time-series variables.

6.4 Cointegration test results

Having established that all time-series variables are I(1), the Johansen procedure is performed to examine the possible presence of cointegration among the variables, and to estimate the associated cointegrating vector(s) if such cointegration does, indeed, exist.

Because results of cointegration test can be quite sensitive to the chosen lag length, we must first determine the lag length and justify our selection. To begin with, we sequentially estimate a VAR model involving the undifferenced (or level) data of the four variables (ie. URW, RUW, AGP, and UGDP) for West Java, Central Java, and East Java. Based on the VAR estimates, Eviews is employed to generate choices of optimal lag length based on several different criteria. An optimal range of lag length is obtained for each province is obtained based on the lag lengths provided by the Likelihood Ratio (LR) test and the Akaike Information Criterion (AIC). Employing residual diagnostics on the VAR model for every
lag length within the optimal range, we choose 10 (ten), 6 (six), and 10 (ten) as the most optimal lag lengths in the models for West Java, Central Java, and East Java, respectively. The selection seems reasonable given the quarterly frequency of the data. The performed residual diagnostics include Langrange Multiplier (LM) autocorrelation test, normality test, and White heteroskedasticity test. While there is no evidence of statistically significant autocorrelation or heteroskedasticity, the results of the normality test on skewness and kurtosis are less convincing for each provincial case. The implication is, however, less serious. Given the large sample size of our quarterly time-series data, normality can still be statistically rejected even though the residuals are actually not far from normality.

In the next step, we apply the optimal lag length from the VAR model estimates into the proposed Johansen cointegration test for West Java, Central Java, and East Java. For each case, three models with different assumptions on the intercept and deterministic trends, as specified in equations (4.1), (4.2) and (4.3), are tested.

Table 6.4 displays the results of the cointegration test for West Java. To determine the presence of cointegration among all the variables included in the model, the calculated values of maximum eigenvalue statistic and trace statistic are compared with the Osterwald-Lenum’s critical values of these test statistics. For West Java, at 5 percent level of statistical significance, both maximum eigenvalue and trace tests indicate the presence of at least 1 (one) cointegrating relation in every model specification. As discussed previously in Chapters 3 and 4, the cointegrating relation(s) can be interpreted as the long-run relationship
or equilibrium of the variables concerned. It provides a basis for identifying the cointegrating vector, which would give us the empirical estimates of the long-run relationship.

Table 6.4: Johansen cointegration test results for West Java

<table>
<thead>
<tr>
<th>Model</th>
<th>$r = 0$</th>
<th>$r \leq 1$</th>
<th>$r \leq 2$</th>
<th>$r \leq 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0: \lambda = 0$</td>
<td>37.12</td>
<td>79.17</td>
<td>53.12</td>
<td>60.16</td>
</tr>
<tr>
<td>$H_0: \lambda \leq 1$</td>
<td>23.10</td>
<td>42.05</td>
<td>34.91</td>
<td>41.07</td>
</tr>
<tr>
<td>$H_0: \lambda \leq 2$</td>
<td>13.29</td>
<td>18.95</td>
<td>19.96*</td>
<td>24.6**</td>
</tr>
<tr>
<td>$H_0: \lambda \leq 3$</td>
<td>5.66</td>
<td>5.66</td>
<td>9.24</td>
<td>12.97</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>$r = 0$</th>
<th>$r \leq 1$</th>
<th>$r \leq 2$</th>
<th>$r \leq 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0: \lambda = 0$</td>
<td>32.97</td>
<td>65.18</td>
<td>47.21</td>
<td>54.46</td>
</tr>
<tr>
<td>$H_0: \lambda \leq 1$</td>
<td>19.03</td>
<td>32.21</td>
<td>29.68</td>
<td>35.65*</td>
</tr>
<tr>
<td>$H_0: \lambda \leq 2$</td>
<td>9.29</td>
<td>13.18</td>
<td>15.41**</td>
<td>20.04</td>
</tr>
<tr>
<td>$H_0: \lambda \leq 3$</td>
<td>3.89</td>
<td>3.89</td>
<td>3.76</td>
<td>6.65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>$r = 0$</th>
<th>$r \leq 1$</th>
<th>$r \leq 2$</th>
<th>$r \leq 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0: \lambda = 0$</td>
<td>43.80</td>
<td>77.47</td>
<td>62.99</td>
<td>70.05</td>
</tr>
<tr>
<td>$H_0: \lambda \leq 1$</td>
<td>19.19</td>
<td>33.68</td>
<td>29.68</td>
<td>35.65*</td>
</tr>
<tr>
<td>$H_0: \lambda \leq 2$</td>
<td>10.20</td>
<td>14.48</td>
<td>25.32</td>
<td>30.45</td>
</tr>
<tr>
<td>$H_0: \lambda \leq 3$</td>
<td>4.29</td>
<td>4.29</td>
<td>12.25</td>
<td>16.26</td>
</tr>
</tbody>
</table>

Notes:
- Model 1: no trend in data; intercept (no trend) in cointegrating equation; no intercept in VAR
- Model 2: linear deterministic trend in data; intercept (no trend) in cointegrating equation; intercept (no trend) in VAR
- Model 3: linear deterministic trend in data; intercept and trend in cointegrating equation; no intercept in VAR
- Critical values are from Osterwald-Lenum (1992)
- *, **, and *** indicate 1, 2, and 3 cointegrating equations, respectively
- Lags interval; (in first differences): 10

Table 6.5 shows the cointegration analysis for Central Java. On the basis of the test statistics in the table, we conclude of there is at least 1 (one) cointegrating relation in every specified model at 5 percent level of statistical significance.
Table 6.5: Johansen cointegration test results for Central Java

<table>
<thead>
<tr>
<th>Model</th>
<th>( \lambda - \text{max} ) statistic</th>
<th>( \lambda - \text{max} ) statistic</th>
<th>Trace statistic</th>
<th>Trace statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>1%</td>
<td>5%</td>
<td>1%</td>
</tr>
<tr>
<td>Model 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H_0: r = 0 )</td>
<td>44.05</td>
<td>32.24*</td>
<td>92.40</td>
<td>53.12</td>
</tr>
<tr>
<td>( H_0: r &lt;= 1 )</td>
<td>22.31</td>
<td>26.81</td>
<td>48.36</td>
<td>34.91</td>
</tr>
<tr>
<td>( H_0: r &lt;= 2 )</td>
<td>20.50</td>
<td>20.20</td>
<td>26.05</td>
<td>19.96</td>
</tr>
<tr>
<td>( H_0: r &lt;= 3 )</td>
<td>5.55</td>
<td>9.24***</td>
<td>5.55</td>
<td>9.24***</td>
</tr>
<tr>
<td>Model 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H_0: r = 0 )</td>
<td>38.83</td>
<td>32.24</td>
<td>69.24</td>
<td>47.21</td>
</tr>
<tr>
<td>( H_0: r &lt;= 1 )</td>
<td>21.00</td>
<td>25.52*</td>
<td>30.41</td>
<td>29.68</td>
</tr>
<tr>
<td>( H_0: r &lt;= 2 )</td>
<td>7.34</td>
<td>18.63</td>
<td>9.41</td>
<td>15.41**</td>
</tr>
<tr>
<td>( H_0: r &lt;= 3 )</td>
<td>2.07</td>
<td>6.65</td>
<td>2.07</td>
<td>6.65</td>
</tr>
<tr>
<td>Model 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H_0: r = 0 )</td>
<td>38.88</td>
<td>36.65</td>
<td>87.47</td>
<td>62.99</td>
</tr>
<tr>
<td>( H_0: r &lt;= 1 )</td>
<td>28.68</td>
<td>30.34*</td>
<td>48.60</td>
<td>42.44</td>
</tr>
<tr>
<td>( H_0: r &lt;= 2 )</td>
<td>13.57</td>
<td>23.65</td>
<td>19.92</td>
<td>25.32**</td>
</tr>
<tr>
<td>( H_0: r &lt;= 3 )</td>
<td>6.35</td>
<td>16.26</td>
<td>6.35</td>
<td>16.26</td>
</tr>
</tbody>
</table>

Notes:
- Model 1: no trend in data; intercept (no trend) in cointegrating equation; no intercept in VAR
- Model 2: linear deterministic trend in data; intercept (no trend) in cointegrating equation; intercept (no trend) in VAR
- Model 3: linear deterministic trend in data; intercept and trend in cointegrating equation; no intercept in VAR
- Critical values are from Osterwald-Lenum (1992)
- *, **, and *** indicate 1, 2, and 3 cointegrating equations, respectively
- Lags interval; (in first differences): 6

The cointegration test results for East Java are presented in Table 6.6 below. For model 1, at 5 percent level of statistical significance, results from both maximum eigenvalue and trace tests indicate the presence of at least 1 (one cointegrating relation). Test on model 2 however gives mixed results. At 5 percent level of statistical significance only the trace test found the presence of at least 1 (one) cointegrating relation, while the maximum eigenvalue test shows none. Similar test results are obtained for model 3. Given the technical superiority of the maximum eigenvalue test over the trace test (Enders, 2004), The mixed results from models 2 and 3 should put us into caution on the quality of the cointegration estimates under these model specifications.
Table 6.6: Johansen cointegration test results for East Java

<table>
<thead>
<tr>
<th>Model</th>
<th>$H_0: r = 0$</th>
<th>$H_0: r \leq 1$</th>
<th>$H_0: r \leq 2$</th>
<th>$H_0: r \leq 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda_{\text{max}}$</td>
<td>$\lambda_{\text{max}}$</td>
<td>Trace statistic</td>
<td>Trace statistic</td>
</tr>
<tr>
<td></td>
<td>statistic</td>
<td>5%</td>
<td>1%</td>
<td>5%</td>
</tr>
<tr>
<td>Model 1</td>
<td>36.19</td>
<td>28.14</td>
<td>33.24</td>
<td>82.30</td>
</tr>
<tr>
<td>Model 2</td>
<td>24.13</td>
<td>22.00</td>
<td>26.81*</td>
<td>46.11</td>
</tr>
<tr>
<td>Notes:</td>
<td>- Model 1: no trend in data; intercept (no trend) in cointegrating equation; no intercept in VAR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Model 2: linear deterministic trend in data; intercept (no trend) in cointegrating equation; intercept (no trend) in VAR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Model 3: linear deterministic trend in data; intercept and trend in cointegrating equation; no intercept in VAR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Critical values are from Osterwald-Lenum (1992)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- $*$, **, and *** indicate 1, 2, and 3 cointegrating equations, respectively</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Lags interval; (in first differences): 10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6.5 Cointegration vector identification

As discussed in Section 4.3, while the existence of more than one cointegrating vector implies a higher level of stability in the long-run relationship of the concerned variables, it could pose an identification challenge. We proceed to the cointegration vector identification by following for the most part the Greenslade et al.’s approach (2002). Specifically, we first recall the theoretical discussion on Section 3.2, in which we safely deduce the exogeneity of real urban wage rates, urban employment rate, and agricultural price variables in the model based on the main results of the Harris-Todaro migration model. The exogeneity along with the knowledge of the actual conditions and the institutional characteristics of the Indonesian
rural and urban labor markets are then used for imposing on the estimated system three restrictions, namely, the unique cointegrating vector restriction, the exogeneity restriction, and the optimal lag length restriction obtained from the pre-testing residual diagnostics.

6.5.1 Cointegration vector identification results for West Java

For West Java, model 2 from the cointegration test is selected, and after imposing the exogeneity restriction, the unique long-run relationship normalized by real rural wage rate is represented in equation (6.1).

\[
RUW = 0.26URW + 0.14AGP + 0.14UGDP + 4.54
\]  \quad (6.1)

This is the cointegrating vector obtained from the largest eigenvalue in the cointegration test results for model 2. In other words, this is a stable equilibrium relationship to which all the variables in the model have a tendency to return in the long run (Engle & Granger, 1987). All restrictions on normalization and exogeneity are binding. The \( \chi^2 \) (Chi-Square) statistic is 2.30 with the \( p-value \) equals to 0.51, which suggests the restrictions cannot be rejected even at 10 percent level of statistical significance, implying all the imposed restrictions are statistically binding.

In West Java, in the equilibrium real rural wage rate increases by 0.26 percent if real urban wage increases by 1 percent; by merely 0.14 percent if agriculture price increases by 1 percent; and by 0.14 percent if real urban GDP increases by 1 percent.
Residual diagnostics on equation (6.1) is performed to ensure that the behavior of the residual does not violate the independence and stationarity assumption. Specifically, we check for autocorrelation, heteroscedasticity and normality of the residuals. Results of the autocorrelation (LM) test show that there is no significant evidence of autocorrelation left behind. Furthermore, the White Heteroskedasticity (with no cross term) test provides significant evidence of homoscedasticity in the residuals. In addition, while the results of VAR Residual Normality test employing Jarque-Bera test indicate significant evidence for non-normal skewness and kurtosis of the residuals, we believe this non-normality is not worrisome, as in a large sample case it is possible to get a highly significant Jarque-Bera statistic (indicating rejection of normality null hypothesis) even though the residuals are actually not far from normality. This merely indicates that a lot more of lags are necessary in the model specification in order to erase the non-normality.

6.5.2 Cointegration vector identification results for Central Java

After imposing exogeneity, the unique long-run estimates for Central Java normalized by real rural wage rate are obtained and shown in equation (6.2).

\[ RUW = 0.53URW + 0.17AGP + 0.19UGDP + 0.88 \]  

(6.2)

Model 2 specification from the cointegration tests on Central Java data has been chosen, as visual inspection of the data series indicate that each series has some linear deterministic
trend, and adding intercept to the cointegration specification should sufficiently capture it. \( \chi^2 \) statistic is 1.87 with \( p \text{ value} \) equals to 0.60, implying that all imposed restrictions are statistically binding at a high level of significance.

In Central Java, the equilibrium real rural wage rate is expected to increase by 0.53 percent when real urban wage rate increases by 1 percent; by 0.17 percent if agriculture price goes up by 1 percent; and by 0.19 percent when real urban GDP rises by 1 percent.

Residual diagnostics employing autocorrelation (LM) test, White heteroskedasticity test, and VAR residual normality test is run on equation (6.2). Similar results as in equation (6.1) are obtained, in which there is no significant evidence for autocorrelation and heteroskedasticity in the residuals. Again, rejection on the normality hypothesis is likely due to a large high frequency sample, which requires more lags to be included in the model specification.

### 6.5.3 Cointegration vector identification results for East Java

For East Java, model 3 from the cointegration test is selected, and after imposing the exogeneity restriction, the unique long-run relationship normalized by real rural wage rate is represented in equation (6.3).

\[
RUW = 0.14URW - 0.70AGP + 019UGDP + 0.02t + 6.83
\]  

(6.3)
\( \chi^2 \) statistic is 7.77 with 0.05 \( p - value \), implying that all imposed restrictions are binding at 5 percent level of statistical significance.

However, the negative sign of the long-run coefficient for agriculture price is counter-intuitive for an interpretation. So, an additional restriction on the coefficient equal to zero is further imposed, and tested whether the restriction is binding at a higher level of statistical significance. The \( \chi^2 \) statistic of the more restricted model is rises to 9.03 while the \( p - value \) goes up to 0.06, indicating that all imposed restrictions are binding at a slightly higher level of statistical significance. The estimates suggest that in the long-run rural wage rate in East Java will increase by 0.59 percent if urban wage rate increases by 1 percent, and by 0.86 percent if urban GDP improve by 1 percent.

To sum up, in all provinces included in this thesis, the urban wage rate is a much more important determinant of rural wage rate than agricultural price or urban GDP, except in East Java where the long-run impact of real urban GDP on equilibrium real rural wage rate is much higher than that of real urban wage rate. The long-run elasticities for real urban wage rate are between 0.14 and 0.53, with an average elasticity of 0.31. By comparison, the agriculture (rice) price has elasticities between zero and 0.17. On average the real urban wage rate elasticity is slightly less than twice as large as the agricultural price elasticity. Furthermore, the elasticities of real urban GDP for West Java, Central Java, and East Java are roughly similar at 0.14, 0.19 and 0.19, respectively. It is also noteworthy that in East Java the long-run impact of real urban GDP on equilibrium real rural wage rate is slightly higher than that of real urban wage rate.
Residual diagnostics on equation (6.3) yield similar results as when the same set of tests are applied to the West Javanese and Central Javanese residual data. There is no significant evidence of autocorrelation and heteroskedasticity in the residuals, while non-normality persist due to large high-frequency sample, indicating more lag need to be included in the tested model specification in order to erase non-normality. Note, however, adding more lags to rectify this problem would eventually lead to another problem, namely, multicollinearity.

It is also important to note that unlike in the case of West Java and Central Java, the cointegration relation model specification for East Java includes a linear time trend (in addition to an intercept). The residual diagnostic results suggest that this specification is the most suitable one for East Java. A time trend in the cointegration equation could make sense, for example, if the equilibrium for a variable follows a trend, and there are no other trending variables in the cointegrating relationship to cancel it out. In such case, a trend belongs to one of the cointegrating variables. As Kennedy (2003) asserts, the ratio of wage to prices, for example, may follow a trend. Since our wage time-series are in real terms and expressed as ratio of nominal wages to prices, the equilibrium level of the real (urban or rural) wages rates may certainly follow a trend, because their price component typically follows a trend. In addition, from an economic interpretation view, it is plausible in the case of East Java that the real agricultural wage rates grow more rapidly than the other variables in the cointegrating vector, especially in the long-run.
Chapter 7: Conclusion

7.1 Summary

The thesis examines the long-run determinants of equilibrium real rural wage rates in Indonesia over the period 1983-2009, using quarterly time-series data published by the Indonesian government. Based on the Harris-Todaro migration theory and the equilibrium rural wage determination model, the Johansen Maximum Likelihood cointegration framework is used to determine the long-run relationship between real rural wage rate, real urban wage rate, agriculture price, and real urban GDP.

The first step in the Johansen cointegration approach is performing unit root tests on all the four time-series variables. The unit root test results indicate that the level of all the tested time-series variables are non-stationary, while their first-differences are stationary, implying that they are all $I(1)$ time-series variables. These results are consistent with our prior expectations, except for results of the agriculture (rice) price time-series. There was a good reason to believe that the rice price index time-series are stationary due to active intervention in the rice market by BULOG (the Government Logistics Agency). In particular, the agency is known to have engaged in open market operations to stabilize prices received by rice farmers. That kind of targeted market intervention is likely to make the rice price time-series more systematic, making them stationary time-series. Hence, the presence of unit root in the rice price index time series is a little surprising. Nonetheless, the results are consistent with findings in some recent works (Timmer, 1996; Barichello et al., 1998; and Dawe, 2001),
which suggest that actual prices paid to rice farmers by BULOG closely follow the world market prices. Apparently, instead of following certain welfare-driven target prices, BULOG’s intervention seems aimed at smoothing out local prices received by rice farmers, so that they would mimic world market prices.

The second step in the Johansen cointegration approach is performing the cointegration tests on the suspected cointegrating variables, whose results suggest the statistically significant presence of at least one cointegrating relationship among all the time-series variables for West Java, Central Java, and East Java. This implies the presence of at least one stable equilibrium relationship to which the concerned variables have a tendency to return in the long-run. Once the existence of cointegration is established, the third step is identifying the cointegrating vectors. Accordingly, cointegrating vector(s) for each province are identified using matrix restriction techniques, taking advantage of the exogeneity property derived of the rural wage determination theory and the residual diagnostics results from the cointegration model at various lag lengths.

In our first research question, we are interested in to what degree the rural labor market is integrated with the rural labor market. The identified long-run relationship estimates provide quantitative information about the level of rural-urban labor market linkage in West Java, Central Java, and East Java. The estimated long-run elasticities provide a way to quantitatively measure how rural variables (ie. agricultural wage rates and rice prices) respond to changes in urban variables (ie. urban wage rates and manufacturing GDP).
In our second research question, we would like to know whether the rural variables play a more or a less important role than the urban variables in determining the equilibrium rural wage rates. The estimation results in each of the three identified cointegrating vectors allow us to directly compare the magnitudes of the long-run elasticities estimates for the rural variable determinant (ie. agricultural price) and the urban variables (ie. urban wage rates and manufacturing GDP). As we have seen in the estimation results, in each of the three cases (West Java, Central Java, and East Java), we have the urban variables playing a more important role than the rural variable in determining the long-run agricultural wage rates.

In our last research question, we ask whether rural wage rates respond differently to changes in rural variables and urban variables across provinces in Indonesia. It is possible to directly compare the long-run coefficient estimates of every rural wage rate determinants for the three provinces; however, while is convenient to do such direct comparison of the long-run elasticities for each determinant of the agricultural wage rates across provinces, it has to be done cautiously by bearing in mind that the estimates come from three different data generating processes, and that there is no formal statistical method to test whether there is a difference (or not) between them as in the OLS short-run structural estimation procedure.

Similarly, it would be tempting to compare our long-run elasticities estimates with those from similar studies that have been done for other countries. For example, comparing our result with those of Ghana (Abdulai and Delgado, 2000), Bangladesh (Palmer-Jones and Parikh, 1998), and the Philippines (Lasco et al., 2008), we can clearly conclude that the long-run elasticities for urban variable determinants in Indonesia are much larger than in those
countries, while the long-run elasticities for rural variable determinants are lower in Indonesia than in those countries, implying that there is higher degree of rural-urban labor market integration in Indonesia, compared to Ghana, Bangladesh or the Philippines. However, for the same reason as in the above paragraph, such inter-country comparison should be approached with caution.

7.2 Policy implications

The identification of the cointegrating vectors reveals estimates for the long-run elasticities of agricultural wage rate determinants in each province. These results have several policy implications worth discussing.

First, the long-run impact of Government protectionist trade policy by raising rice prices to help the poor are likely to be very limited in Indonesia. For instance, in 2000, the government imposed import tariffs on rice, which raised prices by about 25 percent. This would have raised rural wage rates by merely 3.5 percent in West Java, 4.3 percent in Central Java, and practically nothing in East Java. The impact of the imposed tariff on rural wage rate on Java on average is about 2.6 percent, which is a negligible increase, given that the imposed tariff only provides a once-and-for-all increase.

By comparison, annual urban GDP growth has always been between 6 percent and 8 percent per year in the last decade. A minimum annual urban GDP growth rate of 6 percent would have raised rural wages by 0.84 percent, 1.14 percent, and 1.14 percent per year in West
Java, Central Java, and East Java, respectively. The impact of such urban GDP growth rate growth rate on rural wage rate in Java as a whole is about 1.04 percent per year on average. Unlike the tariff-imposed rice price rise, urban GDP growth generates more than one-time effects. Over the decade since 2000 the compounded impact of a 6 percent annual urban GDP growth would have amounted to a 10.4 percent increase rural wage rates.

In addition, urban wage rates in Indonesia also keep rising each year at faster rate than the urban GDP growth rate. However, even if they rose only as fast as the minimum urban GDP growth rate of 6 percent per year, this would be raising rural wage rates by 1.6 percent in West Java, 3.2 percent in Central Java, and 0.8 percent in East Java. Over the decade this would have raised rural wage rates for the whole of Java by 18.6 percent on average. Hence, within the ten-year period, the one-time 25 percent tariff increase would have boosted rural wage rates by only 1/10-th as much of the combined increase from the urban wage rate growth and the urban GDP growth.

Second, using a similar argument as in the above, if the government decided to liberalize trade policy in rice by completely dismantling tariff barriers, it would lower rural wage rates by merely 2.6 percent, which means it would have virtually no effect on rural poverty as measured by the rural wage rates. At the same time, free trade in rice would mean lower rice prices which would translate into increased food security for poor consumers in urban areas, and even in rural areas where almost all the truly poor are net consumers of rice. Clearly for the case of Indonesia, or at least Java, unlike what many suspect, there is no substantial trade-off between globalization and poverty reduction. In fact, as Ananta and Barichello (2012)
describe, many, if not most, departures from free trade in the agricultural sector are the result of lobby efforts by larger farmers who may not be “rich” but are also far from poor. Judging by the political economy of agricultural policy in Southeast Asia, these lobby groups feature a high representation of agricultural land owners. By contrast the consumers of those food commodities who are hurt by those higher prices include large numbers of the poor. So agricultural trade protection often, perhaps typically in lower to middle income countries, is regressive, and worsens poverty.

One might argue if a higher rice price harmed rural poor laborers, why the positive long-run elasticity estimates for the rice price (in the case of West Java and Central Java) seemed to suggest that the agricultural wage rates would improve as a trade tariff increased the rice price. After all, the agricultural wage rates are expressed in real terms and, hence, have taken into account any rice price increase. However, note here that we are not claiming that increasing rice prices by the trade tariff will reduce the real agricultural wage rates. We only show that such tariff-induced increase in rice prices will bring only limited and one-time positive impact on the agricultural wage rates, as opposed to a persistent increase in the urban wage rates or the manufacturing GDP whose positive impacts are much more significant and long lasting. Moreover, the consumption bundle of the rural poor are likely to have more weights on rice than the consumption bundle used in the calculation of the rural CPI, and therefore the already limited positive impact brought by the tariff-induced rice price increase on the rural poor by an increase in the agricultural wage rates will further be reduced. It is also worth noting that we have not included here the calculation for the welfare loss from imposing the trade tariff on rice prices.
On the other hand, using the above same notion that the consumption bundle of the rural poor is likely to carry more weights on rice than the consumption bundle used in computing the rural CPI, we can also infer that any trade liberalization program that reduces rice prices will only harm the rural poor even less.

Finally, the estimation results also provide a valuable guide in assessing the effectiveness of on-going poverty reduction programs in Indonesia. An evaluation can be made as to whether a program provides a sustainable path out of poverty. For example, between 2005 and 2010, the Indonesian government implemented an unconditional cash transfer program, which provides about $11 cash per month (or 36-cents-a-day) to each of the targeted 19.2 million poor and near-poor households. The world’s largest ever cash transfer program was initially introduced to serve as a cushion to shocks that had come from the drastic fuel subsidy reduction, which had severe impacts on the living costs of the poor and near-poor families. The unconditional cash transfer program may probably be inevitable as a short-term solution with little market distortions, but it clearly does not provide a sustainable path out of poverty given the empirical estimates of the determinants of equilibrium real rural wage rates.

7.3 Thesis limitations and recommendations for future studies

The main contribution of the thesis is an examination of the long-run relationship between rural variables and urban variables in Indonesia. From technical perspective, the strengths of the thesis are the inclusion of the effects of rural-urban labor mobility in the determination of
the equilibrium rural wage rates, and the application of the Johansen Maximum Likelihood cointegration method in the model estimation. Nevertheless, several limitations remain in this thesis, as follows:

First, it would be interesting if we could obtain time-series datasets for provinces outside of Java and perform the same empirical analysis that we did on the three Javanese provinces. The availability of such data would allow a direct comparison on the degree of rural-urban labor market integration in the Java region and the non-Java region. Our preliminary hypothesis suggests the rural-urban labor market integration is higher in Java than outside Java, implying urban variables are more important in determining equilibrium rural wage rates in Java.

Second, it would be useful if quarterly employment rate or unemployment rate data were available. The accuracy of the long-run estimates would be enhanced when these data were used instead of the proxy real manufacturing GDP variable. Of course, we could choose to change from the quarterly to annual data frequency and use the available annual unemployment rate data, but the estimation will suffer the loss of many degrees of freedom due to the much shorter span of time, which will ultimately affect the ability to make robust statistical inference.

Third, while the cointegration estimates gives important information about the long-run relationship among the concerned variables, it is not clear how each variable adjust itself in the short-run. It will be further enlightening if in the future short-run wage adjustment models can be developed. As suggested by the Granger Representation theorem (Engle and
Granger, 1987), the presence of cointegration relationship(s) allows us to use the empirical estimates from the cointegration vector identification in the second-step Vector Error-Correction Model (VECM), which will give empirical estimates for both long-run and short-run responses of the changes in each of variables in the model. In fact, the VECM estimates will allow us to assess not only the long-run coefficients and short-run coefficients, but also the speed of adjustment for any exogenous shock(s).

Fourth, it would be important to go beyond the aggregate data and examine the effect of changes in rural and urban variables on rural wage rates at the individual or household level. In particular, effects on different gender, ethnicities, and religion are relevant to Indonesia. This however will require longitudinal surveys of rural employment and labor markets, and social and demographic characteristics of rural labor. In addition, different model and econometric techniques are also needed for such study.
References


Deshingkar, P. 2006. Internal migration, poverty and development in Asia: Including the excluded. *IDS BULLETIN-INSTITUTE OF DEVELOPMENT STUDIES* 37 (3): 88-.


Rapsomanikis, George, David Hallam, and Piero Conforti. 2006. Market integration and price transmission in selected food and cash crop markets of developing countries.


Appendix

Mathematical derivation of Harris-Todaro model

Our theoretical model of migration is based on the H-T model of rural-urban migration. The future expected income from migration is given by

\[ \int_0^\infty \left[ PW_u + (1 - P)W_b \right] e^{-rt} dt - C = \frac{1}{r} \left[ PW_u + (1 - P)W_b \right] - C \]  

(A.1)

where \( C \) is the direct cost of migration, \( r \) is the migrants’ discount rate, \( P \) is the probability of employment at real wage \( W_u \) and \( W_b \) is the real income received if unemployed or employed in the informal sector. The would-be migrants compare (2.1) with the future income from remaining in the rural sector.

\[ \int_0^\infty W_r e^{-rt} dt = \frac{1}{r} W_r \]  

(A.2)

If employment is a certain prospect (i.e. \( P=1 \)) then migration takes place only if there are gains from moving, i.e., only if

\[ \frac{1}{r} W_u - C > \frac{1}{r} W_r \quad \text{or} \quad W_u - W_r > rC \]  

(A.3)

Under conditions of uncertainty, the probability of obtaining employment is given by

\[ P = \frac{\bar{L}_u}{N_u} = \frac{\bar{L}_u}{\bar{L}_u + M \bar{N}_r} \]  

(A.4)

where \( L \) is population employed, \( N \) is total population and \( M \) is the rate of migrants coming from the rural region and the subscript \( u \) refers to urban regions while \( r \) refers to rural areas. Equation (A.4) assumes that migrants compete on equal terms with the incumbent urban
employed population. Thus as $M$ rises, $P$ falls and migration continues only until the returns from (A.1) and (A.2) are exactly equal. Hence, the equilibrium migration rate $M$ is given by

$$PW_u + (1 - P)W_b - W_r = rC$$  \hspace{1cm} (A.5)

with $P$ given by (A.4). Substituting (A.4) into A.5) and solving for $M$ gives the equilibrium migration rate. Equation (A.5) is derived assuming equality holds in (A.3).

$$M = \frac{W_u - W_r - rC}{rC - W_b + W_r} \frac{\bar{L}_u}{\bar{N}_u}$$  \hspace{1cm} (A.6)

We require that $W_b - W_r < rC$ for $M > 0$ which implies there is no incentive to leave rural areas for urban unemployment.

From (A.6), we get the familiar results

$$\frac{\partial M}{\partial W_u} > 0; \frac{\partial M}{\partial W_r} < 0; \frac{\partial M}{\partial L_u} > 0; \frac{\partial M}{\partial C} < 0$$  \hspace{1cm} (A.7)

Inequalities (A.7) state that any marginal increase in urban wage ($W_u$) or decrease in the rural wage ($W_r$) will increase migration. Paradoxically, any policy to increase employment in the advanced urban sector will raise the migration rate and may increase urban unemployment. Hence, as predicted in the H-T models, a policy of creating more employment opportunities in the advanced regions may only enlarge the migration from the backward region. Also, any decrease in the cost of migration will increase $M$. Clearly, the H-T model in the above form, still ignores the impact of human capital, availability of public goods like health care, housing stock and road infrastructure in migration decisions.