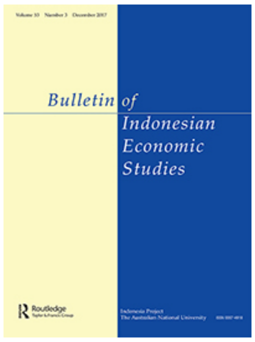
Gravity models of interregional migration in Indonesia

Wajdi, Nashrul; Adioetomo, Sri Moertiningsih; Mulder, Clara

Published in:
Bulletin of Indonesian Economic Studies

DOI:
10.1080/00074918.2017.1298719

Publication date:
2018



Routledge
Taylor & Francis Group

Bulletin of Indonesian Economic Studies

ISSN: 0007-4918 (Print) 1472-7234 (Online) Journal homepage: http://www.tandfonline.com/loi/cbie20

Gravity models of interregional migration in Indonesia

Nashrul Wajdi, Sri Moertiningsih Adioetomo & Clara H. Mulder

To cite this article: Nashrul Wajdi, Sri Moertiningsih Adioetomo & Clara H. Mulder (2017) Gravity models of interregional migration in Indonesia, Bulletin of Indonesian Economic Studies, 53:3, 309-332, DOI: 10.1080/00074918.2017.1298719

To link to this article: https://doi.org/10.1080/00074918.2017.1298719

Accepted author version posted online: 21 Mar 2017.
Published online: 12 Feb 2018.

Routledge
Taylor & Francis Group

Bulletin of Indonesian Economic Studies, Vol. 53, No. 3, 2017: 309–32

GRAVITY MODELS OF INTERREGIONAL MIGRATION IN INDONESIA

Nashrul Wajdi*
University of Groningen

Sri Moertiningsih Adioetomo*
Universitas Indonesia

Clara H. Mulder*
University of Groningen

This article explores the determinants of interregional migration in Indonesia. Employing basic and modified (extended) gravity models, and using data from the 2000 and 2010 Population Censuses and the 2005 Intercensal Population Survey, we test Long's (1985) hypothesis that in the early stages of population redistribution, economic development is positively related to a concentration of the population. Using per-capita GDP as a proxy for income and as an indicator of economic development, we find that migration in Indonesia is indeed directed towards more developed regions. This finding supports the notion that regional disparities in development are an important factor in interregional migration in Indonesia. In line with classical gravity models, our findings show that distance is negatively related to the size of migration flows. However, unlike previous studies of interprovincial migration in Indonesia, we find that the effect of distance has weakened over time.

Keywords: *migration, gravity models, Poisson pseudo-maximum likelihood, PPML*

JEL classification: O15, R23

INTRODUCTION

The strong concentration of Indonesia's population on the island of Java has long been a major concern among policymakers and researchers. Previous studies on interregional migration in Indonesia show that Java—particularly its large metropolitan areas—continues to be the main destination for migrants (Alatas 1993; Chotib 1998; Darmawan and Chotib 2007; Firman 1994; Rogers et al. 2004; Wajdi 2010; Wajdi, Van Wissen, and Mulder 2015). Regardless of the formation of new metropolitan areas on other islands, the attractiveness of Java's metropolitan areas for migrants remains high. These areas, especially the country's two largest metropolises, Jakarta and Surabaya, have high economic densities (as measured by regional GDP per square kilometre of urban land area) and high population

* This research was made possible by NFP/Nuffic Scholarship grant number NFP PhD CF 8776, granted to Nashrul Wajdi. The authors would like to thank Professor Leo J. G. van Wissen for his valuable input on the first draft of this manuscript.

ISSN 0007-4918 print/ISSN 1472-7234 online/17/000309-24
http://dx.doi.org/10.1080/00074918.2017.1298719

concentrations (World Bank 2012). In contrast, the regions outside Java have had low economic densities for many decades.

According to Long (1985), population concentrates in urban centres during the early stages of development and deconcentrates during the later stages of development. A study by Wajdi, Van Wissen, and Mulder (2015) indicates that the migration pattern in Indonesia is in line with Long's thesis, which posits that economic development has a strong relationship with migration. However, very few studies have investigated the association between economic development and migration flows within the local context, and only a few have used an explanatory modelling approach to explain migration flows.

Darmawan and Chotib (2007) used per-capita GDP, minimum regional wages, and unemployment rates to model interprovincial migration flows in Indonesia using hybrid gravity models. Wajdi (2010) modelled migration as a function of wage differentials, unemployment rates, and economic structure. Van Lottum and Marks (2012) modelled interprovincial migration using a gravity model framework and showed that gravity models are very suitable for analysing internal migration flows in a large country such as Indonesia. The authors modelled migration as a function of population size, per-capita income, distance, contiguity between regions, and two control variables, namely transmigration and urban primacy. They found that wage differentials between regions were relatively unimportant, but that the existence of Jakarta as a primate city was a very important determinant of migration.

These three studies found that internal migration in Indonesia was mainly directed towards more developed regions. However, all three employed rather large regions—islands in the case of Wajdi (2010) and provinces in the case of Darmawan and Chotib (2007) and Van Lottum and Marks (2012)—many of which were quite heterogeneous with regard to economic development and degree of urbanisation. As a consequence, they failed to take into account differences between metropolitan and non-metropolitan areas in Indonesia—except for Van Lottum and Marks, who considered the existence of Jakarta as a primate city.

In order to address these limitations of the previous research, we focus on two research questions. The first is an existing question to which we attempt to provide a new answer: to what extent are migration flows in Indonesia directed towards the more developed regions? We address this question in a considerably more detailed and comprehensive way than has been done before, in particular by distinguishing between metropolitan and non-metropolitan areas. We also explore the impact of determinants of migration that have rarely been considered for the case of Indonesia, namely the percentage of agricultural workers, the percentage of highly educated workers, contiguity between regions, and the migrant stock. We also use a different statistical estimation method, the Poisson pseudo-maximum likelihood (PPML) estimator, which is more suitable for count data. The main aim is to test Long's (1985) hypothesis that during the early stages of development, economic development is positively related to a concentration of the population. Since our theoretical explanations of migration are adopted from studies in developed countries, we aim to investigate to what extent these theories are applicable to the case of Indonesia. Therefore, our second research question is: to what extent do common migration determinants explain interregional migration flows in Indonesia? To help us explore these questions, we use data from the Population Censuses of 2000 and 2010 and the Intercensal Population Survey (Supas) of 2005, employed in a gravity model framework.

THEORETICAL BACKGROUND

Long's Thesis and the Basic Gravity Model

According to Long (1985), population concentrates in urban centres during the early stages of development because these centres fulfil the need for social and economic interaction; and deconcentrates during the later stages of development because improvements in transportation and communication permit easier interaction over longer distances. A study by Wajdi, Van Wissen, and Mulder (2015) finds that Indonesia is currently in the early stages of population redistribution, but is moving towards the later stages. The authors find some indications of urbanisation and overurbanisation in Indonesia (indicating concentration), but only weak signs of suburbanisation and metropolitan to non-metropolitan migration (indicating deconcentration).

We argue that the current pattern of population redistribution in Indonesia is in line with Long's thesis, that is, that during the early stages of development, people migrate from less developed regions to more developed regions. This thesis can be examined using one of the most popular models to predict migration flows, the spatial interaction model, in particular the gravity model of migration.

According to Öberg (1997), the spatial gravity model is one of the strongest theories in applied geography. The model is based on the works of Ravenstein (1885), who stated that the volume of migration is inversely related to distance. This so-called social physics theory (analogical to the physical laws of Newtonian physics) was introduced into geography by Zipf (1946), whose P1P2/D hypothesis postulates that migration is directly proportional to the origin's population (P1) and the destination's population (P2), and inversely proportional to the distance between the origin and the destination (D) (see also Anderson 1979; Niedercorn and Bechdolt 1969).

The basic formulation of the gravity model of migration is as follows:

$$M_{ij} = g \cdot \frac{P_i^{\alpha} P_j^{\beta}}{D_{ij}^{\gamma}}, \qquad (1)$$

where M_{ij} is the migration from region i to region j, P_i and P_j are the population sizes of the two regions i and j respectively, D_{ij} is the distance between i and j, and g is a constant (Bunea 2012).

When applying Newton's law in the gravity model of migration, the total population is the most representative variable representing the mass of the two objects i and j. The total population represents the capacity of a region to send migrants; the more populated a region is, the bigger the volume of migration from that area (Flowerdew and Aitkin 1982; Kim and Cohen 2010). For the case of Indonesia, Van Lottum and Marks (2012) found that origin and destination regions with larger total populations attracted more migrants, where the coefficient for the total population at the origin was slightly larger than the coefficient for the total population at the destination.

The distance decay in the gravity model of migration can be used as a representation of the physical costs of migration, and also to some extent of the non-physical costs, such as language and cultural barriers. The actual costs of migration are not usually measured, although they do affect migration flows. When physical distance increases, the costs of moving will also increase, and migration will therefore diminish. Over time, however, improvements in technology, communication, information, and transportation will reduce the costs of migration. Thus, the effect of distance on migration is negative, but the magnitude of the effect diminishes over time (Bodvarsson and Van den Berg 2013; Bunea 2012; Etzo 2008; Fan 2005; Greenwood 1997; Greenwood and McDowell 1991; Zipf 1946). Therefore, it is necessary to assess the effect of distance over time. We expected that the effect of distance would progressively decline over the periods we studied: 1995–2000, 2000–2005, and 2005–10. It should be noted, however, that Van Lottum and Marks (2012) found that the effect of distance on interprovincial migration in Indonesia actually increased over time.

The Modified Gravity Model: Push and Pull Factors

Because there are so many potential determinants of migration flows, estimating the basic formulation of the gravity model will almost always suffer from omitted variable bias. To overcome this problem, researchers have introduced other variables into the basic gravity model (Bodvarsson and Van den Berg 2013; Greenwood 1997). The extended form of the gravity model is known as the extended or modified gravity model. The general representation of the modified gravity model as proposed by Greenwood (1997) contains per-capita real income or GDP in source region *i*, per-capita real income or GDP in destination region *j*, a vector of explanatory variables describing different characteristics of the origin (push factors), and a vector of explanatory variables describing different characteristics of the destination (pull factors). Push factors are characteristics of the origin that may encourage out-migration or inhibit in-migration, while pull factors are characteristics of the destination that may encourage in-migration or discourage out-migration (Bodvarsson and Van den Berg 2013; Bunea 2012; Greenwood 1997). One of the major push/pull factors of migration is the attractiveness of a region.

A key determinant of the attractiveness of an area is the expected earnings of an individual, indicated by income per capita (Beine, Bertoli, and Moraga 2014; Fan 2005). Because potential migrants will evaluate the real value of their expected net gains from migration by considering the present discounted value of their expected future stream of net gains, current earnings can be considered a good proxy for expected future earnings (Borjas 2001, 2013; Bunea 2012; Greenwood 1975; Sjaastad 1962; Todaro 1980). As Beine, Bertoli, and Moraga (2014) have stated, GDP per capita at the destination is a good measure of the income prospects of potential migrants from all regions of origin. Besides representing the income difference between two areas, GDP per capita can be used as an indicator of the difference in level of economic development (Bodvarsson and Van den Berg 2013; Fan 2005). In the Indonesian context, GDP including oil and gas (hereafter 'GDP') has been widely used as a tool to assess regional development performance (Bappenas 2015). Using GDP divided by urban land area to measure economic density, the World Bank (2012) found that the metropolitan areas in Java had some of the highest economic densities in the country. This would provide further evidence of the gap in economic development between Java and the rest of the country.

The effects of income on migration can be viewed from two different perspectives: micro and macro. From a micro perspective, migration generally occurs because a migrant gains income benefits from moving (Greenwood 1975). From a macro perspective, migration occurs from low-income to high-income regions, or from less developed to more developed regions. Therefore, the higher the GDP at the destination, the greater the attractiveness of the destination. Put in economic terms, migration will occur if the income elasticity is negative at the origin and positive at the destination.

However, migration may also be positively related to a higher level of economic development in the region of origin, for two reasons. First, as Massey (1988) has argued, the development process may produce a group of workers who start looking for greater rewards elsewhere. Second, the higher the level of economic development in the origin area, the more resources and opportunities potential migrants have, and therefore the higher the migration propensity will be.

Similarly, a large income differential between origin and potential destination does not necessarily induce migration, for two reasons. First, there is a high probability that a migrant will not fulfil the requirements for quick re-employment at the destination (Fan 2005; Greenwood 1975; Todaro 1969). Second, migrants may want to improve their incomes relative to their local communities, rather than improving their absolute incomes. This type of migration is known as 'migration as a response to relative deprivation', an idea introduced into migration studies by, among others, Stark and Yitzhaki (1988). The relative deprivation concept, which was developed in the field of psychology, implies that a person's happiness is derived not only from how many goods they can afford from their own income, but also from the relative ranking of their income compared with the income of their community. When potential migrants expect to experience an increase in their relative income at the destination, even though their absolute income stays the same, then migration occurs, because they will experience a higher level of well-being or satisfaction (Bodvarsson and Van den Berg 2013; Stark and Yitzhaki 1988). For the case of Indonesia, Van Lottum and Marks (2012) found a negative effect of the log of the ratio of per-capita income in the source region to the per-capita income in the destination region. However, because the effect of income on migration can be different at the origin and the destination, it is necessary to assess the income variable at both origin and destination.

Another feature of economic development and modernisation is the migration of labour out of agriculture, which occurs in both developed and developing nations (Rozelle, Taylor, and DeBrauw 1999). Minami (1967) stated for the case of Japan that migration from agricultural to non-agricultural areas was caused by the rise of non-agricultural wages relative to agricultural wages, as a result of economic development. However, Adams (1969) argued that it was not necessarily the income differential between agricultural and non-agricultural areas that induced migration; he found that people were simply attracted to more industrialised areas. This can be regarded as a sociological phenomenon because the economic motives behind the movement are minor. A study by Butzer, Mundlak, and Larson (2003) on intersectoral migration in Indonesia, Thailand, and the Philippines revealed that labour surpluses had not been redistributed from agriculture to other sectors, and that migration rates from agricultural to non-agricultural areas in these three countries were low compared with those of other countries. Moreover, the low migration rates out of agriculture had caused a persistence of intersectoral income differentials. Although migration had been responsive to income differences in each country, it had also been affected by the absorptive capacity of the non-agricultural sectors of the economy.

The level of educational attainment in a region can be expected to have a substantial effect on migration. The effect of educational level is expected to be positive for both destination and origin. A region that has good higher-education facilities (senior high schools or universities) will attract people who are seeking higher education. A high level of educational attainment is associated with the occupational structure of the region and with greater demand for educated persons. Moreover, regions with highly educated inhabitants are more likely to have good social and cultural amenities that will attract better-educated persons. Highly educated potential migrants generally have a higher propensity to migrate from origin regions and are better equipped to adapt to the new situation in destination regions (Beals, Levy, and Moses 1967; Dahl 2002; Girsberger 2015; Greenwood 1969a, 1969b; Greenwood and McDowell 1991; Lessem 2009; Sahota 1968).

However, the estimated effect of educational attainment may be counterintuitive or may not be found in macro analyses. Greenwood (1969b) argued that the unexpected negative effect of education on labour migration in Egypt might be due to two causes. First, an increase in educational attainment of a potential migrant will increase that person's productivity in the area of origin as well as in the area of destination. Hence, a potential migrant with a high level of education will evaluate the net effect of migration, and decide not to migrate if migration brings no extra gains in productivity. The second cause of a possible negative effect of education in the area of origin is simultaneity bias. If a large flow of migration occurs among more educated persons, then it may cause the level of educational attainment in the region of origin to decrease during the period of measurement, whereas the level of educational attainment in the destination region would be likely to increase.

Because regions differ in the availability of job opportunities, it is important to include a variable as a proxy for the probability that the potential migrant will find a job at the destination within a given period of time. Todaro (1969) suggested the use of the unemployment rate at the destination as a proxy for this probability. Although Todaro's model of migration was specific for two sectors in less developed countries, Greenwood (1975) argued that it could be applied to interregional migration in any country.

As with educational attainment, however, the effect of unemployment on migration can be unexpected. There are three possible explanations for a counterintuitive effect of unemployment. First, simultaneity bias may occur because the variables explaining migration are also likely to be influenced by migration; that is, migration is affected by unemployment but unemployment is also affected by migration (Greenwood 1975). Second, as found by Greenwood (1969a) for the case of labour migration in the United States, this 'wrong' effect of unemployment occurs because unemployment rates in rural areas tend to be lower than those in urban areas. Third, for the case of internal migration in Jamaica, Adams (1969) explains that people are simply attracted to high-income regions, despite the reality that the probability of earning a better income is not very great.

The lower unemployment rates in rural areas than in urban areas are probably due to disguised unemployment in rural areas in the form of underemployment (Greenwood 1969a). For the case of Indonesia, Dhanani (2004) stated that the open unemployment rate (the 'true' unemployment rate, where people had no work but were willing to work and looking for a job) was higher in urban than in rural areas because of the higher proportion of urban than rural youth actively seeking work.

In Indonesia, the definition of unemployment refers to those who have not worked for a minimum of one hour during the reference period (one week prior to the survey), but are actively seeking a job (BPS 2014). This definition excludes underemployed persons who are working less than a 'normal' threshold of working hours—35 working hours per week—but are seeking additional work. This group is overrepresented in rural areas. The National Labour Force Survey (Sakernas) for August 2014 showed that whereas the *unemployment* rate was 7.12% in urban areas and just 4.81% in rural areas, the share of *underemployment* was 4.99% in urban areas but 10.80% in rural areas. Rural youths are likely to believe that they have a better chance of getting a job if they migrate to an urban area. Therefore, in the Indonesian case, a lower unemployment rate could be associated with a higher rate of migration.

Next to push and pull factors, another way of extending the basic gravity model is to add more indicators of the costs of moving. One such indicator is contiguity among regions. Regions that share a common border (such as Jakarta and Bodetabek—see appendix table A1) should have significantly lower moving costs than regions that are not contiguous, while relatively inaccessible destinations (regions with oceans or seas as borders) should have fewer in-migrants due to the increased costs of transportation (Kim and Cohen 2010; see Van Lottum and Marks 2012 for the case of Indonesia). Accounting for contiguity is useful when the measurement of distance relates to a fixed point in each region (for example, a centroid). However, improvements in technology, communication, information, and transportation may reduce the physical costs of migration (Bunea 2012; Greenwood 1997).

Because information may reduce the physical costs of migration, prior information about a potential destination plays an important role in a potential migrant's decision-making process. Such migrants are more likely to move to an area about which they have prior information, rather than to an area about which they have no prior information. Information about potential destinations can be acquired from people who have previously migrated there. This so-called network effect describes the linkages between the potential migrants in the regions of origin and their relatives and friends who have already settled as migrants in the destination areas. The potential migrants' relatives and friends are supposed to facilitate their migration.

This migration network then leads to the accumulation of social capital. Social capital accumulation is defined as an accumulation of migration-related information and resources gained from relatives and friends who have already migrated. The theory of cumulative causation of migration was introduced by Massey (1990), who extended Myrdal's (1957) concept of circular and cumulative causation. Cumulative causation theory postulates that once a migration flow begins, it continues to grow (Fussell and Massey 2004). The idea underlying this concept is that migration creates changes in social and economic structures, leading to more migration. The underlying mechanism proposed in this theory is that migration occurs due to the accumulation of social capital gained from a migration network.

Actual measures of network effects are usually scarce or not available. A popular proxy to measure the network effects of migration is the migrant stock. The migrant stock is defined as the accumulated number of previous in-migrants to the destination who have migrated from a particular region of origin (Beine, Bertoli, and Moraga 2014; Fan 2005; Greenwood 1969a, 1975; Peeters 2012).

DATA AND METHOD

The migration data for this study were derived from the Population Censuses of 2000 and 2010 and the Intercensal Population Survey of 2005. Unlike Van Lottum and Marks (2012), who defined migration as lifetime migration, we defined interregional migration as a change in the place of residence during a five-year period (recent migration). The advantage of using recent migration rather than lifetime migration is that it reflects population dynamics more accurately.

In contrast to the studies by Darmawan and Chotib (2007) and Van Lottum and Marks (2012) analysing interprovincial migration and the paper by Wajdi (2010) analysing interisland migration, we divided Indonesia into metropolitan and non-metropolitan areas. We used the information on metropolitan agglomeration in Government Regulation 26/2008 on National Spatial Planning,[1] and from the World Bank (2012), to distinguish between the two types of area. The 13 regions included in our analysis are summarised in appendix table A1 (see also appendix figures A1 and A2). Table 1 shows the explanatory variables used in the analysis.

Following Conley and Topa (2002), we calculated geographical distance, D_{ij}, as the direct distance in kilometres between the centroids of origin i and destination j. Although this measure does not consider physical barriers such as rivers or highways, it does represent with reasonable accuracy the average distance travelled by migrants.

Per-capita GDP in constant 2000 prices was compiled from various BPS (Statistics Indonesia) publications. We used GDP including oil and gas to account for the full capacity of the economy, and checked whether the results were different from those using GDP without oil and gas.

Differences in economic structure as another proxy for the costs of moving were represented by the percentage of workers in agriculture and the percentage of highly educated workers. We calculated the sectoral employment rate and the unemployment rate based on the National Labour Force Surveys for 2000, 2005, and 2010.

The final variable, migrant stock (S_{ij}) at time t, was defined as the proportion of i to j migration flows to the total out-migration from region i at time $t-5$, that is, the total number of migrants who migrated from i to j divided by the total number of migrants from i to all possible destinations, $(M_{ij\,t-5} / \sum_j M_{ij\,t-5})$. The migrant stock for 2005 was calculated based on the Population Census for 2000, and the migrant stock for 2010 was calculated based on the Intercensal Population Survey for 2005.

In our analysis, we employed three gravity models of migration. The first model is a basic gravity model and is specified as follows, in a linearised form:

$$\ln\left(M_{ij}\right) = \beta_0 + \beta_1 \ln\left(P_i\right) + \beta_2 \ln\left(P_j\right) + \beta_3 \ln\left(D_{ij}\right) + e_{ij}. \tag{2}$$

M_{ij} represents the gross interregional migration flows in Indonesia from the origin i to the destination j. P_i and P_j respectively denote the population at origin i and the population at destination j, while D_{ij} is the geographical distance between origin i and destination j. In accordance with the general principles of the basic gravity

1. Peraturan Pemerintah Nomor 26 Tahun 2008 tentang Rencana Tata Ruang Wilayah Nasional, available at http://www.minerba.esdm.go.id/library/sijh/pp26-2008.pdf.

TABLE 1 *Summary of Data Sources for the Explanatory Variables*

Explanatory variable	Data source
Size of population at origin (P_i) Size of population at destination (P_j)	Authors' calculation based on Population Census 2000 & 2010 & Intercensal Population Survey (Supas) 2005
Geographical distance between origin & destination (D_{ij})	Authors' calculation (see text for details)
Regional GDP including oil & gas per capita at origin ($GDPcap_i$) Regional GDP including oil & gas per capita at destination ($GDPcap_j$)	Authors' compilation based on various BPS (Statistics Indonesia) publications
Percentage of agricultural workers at origin ($AGRI_i$) Percentage of agricultural workers at destination ($AGRI_j$) Percentage of highly educated workers at origin (E_i) Percentage of highly educated workers at destination (E_j) Unemployment rate at origin (U_i) Unemployment rate at destination (U_j)	Authors' calculation based on National Labour Force Survey (Sakernas) 2000, 2005 & 2010
Contiguity (dummy variable): Regions share common border (dC_{ij}) Regions separated mostly by land (dL_{ij}) Regions separated by sea/ocean (reference category)	Authors' elaboration
Migrant stock (S_{ij})	Authors' calculation based on Population Census 2000 & Intercensal Population Survey 2005

model, we expected that β_1 and β_2 would have positive signs, while β_3 would have a negative sign.

Our second model is a modified gravity model and is specified as follows:

$$\begin{aligned}\ln\left(M_{ij}\right) = \beta_0 + \beta_1 \ln\left(P_i\right) + \beta_2 \ln\left(P_j\right) + \beta_3 \ln\left(D_{ij}\right) + \beta_4 \ln\left(GDP_i\right) + \beta_5 \ln\left(GDP_j\right) \\ + \beta_6 \ln\left(AGRI_i\right) + \beta_7 \ln\left(AGRI_j\right) + \beta_8 \ln\left(E_i\right) + \beta_9 \ln\left(E_j\right) + \beta_{10} \ln\left(U_i\right) \\ + \beta_{11} \ln\left(U_j\right) + \beta_{12} \ln\left(dC_{ij}\right) + \beta_{13} \ln\left(dL_{ij}\right) + e_{ij}. \qquad (3)\end{aligned}$$

Because migrants are attracted to destinations that are more developed than their regions of origin, real per-capita GDP was expected to have a negative effect at the origin ($\beta_4 < 0$) and a positive effect at the destination ($\beta_5 > 0$). And because migrants are more likely to migrate from a traditional agricultural sector to a modern sector, the coefficient for the share of agricultural workers was expected to have a positive sign at the origin ($\beta_6 > 0$) and a negative sign at the destination ($\beta_7 < 0$). The coefficients for the percentage of highly educated workers were expected to be positive at both the origin and the destination ($\beta_8 > 0$ and $\beta_9 > 0$). The coefficient for the unemployment rate at the origin was expected to have a positive effect on out-migration ($\beta_{10} > 0$) and a negative effect on in-migration to that region ($\beta_{11} < 0$). Unlike Van Lottum and Marks (2012), who indicated only whether a province shared the same border with another province, following Mayer and Zignago (2011), we included dummy variables to capture the effect of being geographically contiguous, being separated mostly by land, or being separated by sea (the reference category). In this model, *dC* takes a value of 1 if origin *i* and destination *j* share a border (for example, Jakarta and Bodetabek, or Kedungsepur and the Rest of Central Java and Yogyakarta) and 0 if they do not; *dL* takes a value of 1 if origin *i* and destination *j* are separated mostly by land (for example, Jakarta and Bandung Raya) and 0 if they are not. We expected the coefficient for *dC* to have a positive sign ($\beta_{12} > 0$) and the coefficient for *dL* to have a negative sign ($\beta_{13} < 0$).

In order to explore the network effect on interregional migration in Indonesia, we also estimated a gravity model to which we added the migrant stock (S_{ij}) as a proxy for social networks and the availability of information. The migrant stock was also supposed to capture the cumulative effects of past migration. If today's migration patterns reflect the forces of the past to a great extent, this variable would have a strong effect. We estimated this model separately because we were concerned that adding the migrant stock variable might cause problems of endogeneity and multi-collinearity that could lead to overspecification of the model (see, for example, Greenwood 1969b).

With the addition of S_{ij} in equation (3), our third gravity model is specified as follows:

$$\begin{aligned}\ln(M_{ij}) = {} & \beta_0 + \beta_1 \ln(P_i) + \beta_2 \ln(P_j) + \beta_3 \ln(D_{ij}) + \beta_4 \ln(GDP_i) + \beta_5 \ln(GDP_j) \\ & + \beta_6 \ln(AGRI_i) + \beta_7 \ln(AGRI_j) + \beta_8 \ln(E_i) + \beta_9 \ln(E_j) + \beta_{10} \ln(U_i) \\ & + \beta_{11} \ln(U_j) + \beta_{12} \ln(DC_{ij}) + \beta_{13} \ln(DL_{ij}) + \beta_{14} \ln(S_{ij}) + e_{ij}. \end{aligned} \tag{4}$$

Because the availability of information provided by relatives and friends who had previously migrated would reduce migration costs, we expected the coefficient for migrant stock (S_{ij}) to be positive, and if today's migration patterns reflected the forces of the past to a great extent, then this variable could be expected to be highly significant ($\beta_{14} > 0$). Table 2 summarises the explanatory variables and the expected results for each variable.

We estimated the coefficient for our three models using Poisson regression. Poisson regression was chosen over ordinary least squares (OLS) regression to avoid four specific problems identified with estimation of gravity models using OLS, assuming a log-normal distribution of migration flows (Flowerdew and Aitkin 1982). The first problem with OLS is bias in the estimation results due to the logarithmic fitting. Before estimating the parameters in an OLS regression, the migration flows need to be converted into logarithmic values, but in Poisson, this conversion is not necessary. The second is the failure of the model to meet the normality assumption of OLS. In Poisson, there is no normality assumption. Third is the problem of unequal variance in the error terms; this too is not applicable to Poisson. The final problem is unstable results due to zero flows. The zero-flows

TABLE 2 *Summary of Explanatory Variables and Expected Results*

Explanatory variable	Parameter	Expected result
Size of population at origin (P_i)	β_1	Positive
Size of population at destination (P_j)	β_2	Positive
Geographical distance between origin & destination (D_{ij})	β_3	Negative
Regional GDP per capita at origin (GDP_i)	β_4	Negative
Regional GDP per capita at destination (GDP_j)	β_5	Positive
Percentage of agricultural workers at origin ($AGRI_i$)	β_6	Positive
Percentage of agricultural workers at destination ($AGRI_j$)	β_7	Negative
Percentage of highly educated workers at origin (E_i)	β_8	Positive
Percentage of highly educated workers at destination (E_j)	β_9	Positive
Unemployment rate at origin (U_i)	β_{10}	Positive
Unemployment rate at destination (U_j)	β_{11}	Negative
Contiguity (dummy variable):		
Regions share common border (dC_{ij})	β_{12}	Positive
Regions separated mostly by land (dL_{ij})	β_{13}	Negative
Regions separated by sea/ocean	Reference	
Migrant stock (S_{ij})	β_{14}	Positive

problem in OLS is usually treated by changing zero flows into a small number (normally 1) or simply by dropping the observations that contain zero flows. However, this treatment may cause estimation bias. The use of censored regression—for example, Tobit regression—may also cause estimation bias, because both OLS and Tobit regressions have normality as a key assumption that theoretically includes negative values (Brown and Dunn 2011), while Poisson is a count distribution.

The Poisson model, on the other hand, also has some drawbacks. One is a relatively low deviance statistic (as a measurement of the performance of the model) when the number of explanatory variables is small. Therefore, Flowerdew and Aitkin (1982) suggest adding more independent variables to the basic gravity model to improve the estimation performance of the Poisson model. Another drawback of Poisson models is overdispersion. In a Poisson model, the variance is equal to the mean. When the variance in the data is larger than the mean, the standard errors of the coefficients are biased downward. This drawback can be handled in part by ensuring a robust estimation of standard errors (see, for example, Hilbe 1999). However, Silva and Tenreyro (2011a) have shown that when this solution is used for Poisson estimation, a convergence problem may occur, leading to failure to find the right estimates. As a consequence, the estimation will be very sensitive to numerical problems, which may produce spurious and misleading results.

In light of these considerations, we used the Poisson pseudo-maximum likelihood (PPML) estimator proposed by Silva and Tenreyro (2006). A simulation study by Silva and Tenreyro (2011b) confirms that the PPML estimator is generally good, even in the case of overdispersion. Furthermore, the PPML estimator produces a robust estimation, even when the dependent variable has a large number of zeroes. In our study, a comparison of a classical Poisson regression (results not shown) and the PPML regression revealed that the estimated effects were exactly the same, but the standard errors of the PPML regression were larger.

RESULTS

Table 3 provides the descriptive statistics for the variables. It shows that migration flows fell in the period 2000–2005 but increased in the period 2005–10. In terms of development indicators, we found that regional GDP and the share of highly educated workers increased over the entire period (2000–2010), while the share of workers in the agricultural sector declined. The unemployment rate increased in 2000–2005, but then decreased in 2005–10.

Table 4 provides the results from our three models: the basic gravity model (model 1) and the two modified gravity models (models 2 and 3). Overall comparison of the three models shows that, as expected and as Flowerdew and Aitkin (1982) suggested, adding more independent variables to the basic gravity model considerably improves the performance of the Poisson model. Compared with the basic gravity model, the R^2 of the first modified gravity model (model 2) increased from 0.2350 to 0.6681 in 1995–2000, from 0.3590 to 0.6287 in 2000–2005, and from 0.3588 to 0.7171 in 2005–10. The inclusion of the migrant stock variable in model 3 as a representation of social networks led to a further increase in R^2 from 0.6287 to 0.9071 in 2000–2005 and from 0.7171 to 0.9247 in 2005–10. Thus, the modified gravity model indeed proved better than the basic gravity model in explaining migration flows in Indonesia. The model including the migrant stock variable (model 3) predicted migrant flows very well, but might be overspecified.

As expected, the coefficients for population size at the origin showed positive signs, although some of them were statistically insignificant (the basic model and model 2 in 2000). The positive signs for this coefficient indicated that there was more migration from larger regions (in population terms) because they had more capacity to send migrants. The coefficients for population size at the destination were also positive and statistically significant. Most of the destination population-size parameters were close to 1, indicating that the share of migrants received by each destination was approximately proportional to its population size. Unlike Van Lottum and Marks (2012), who consistently found that population at the origin was more important in explaining migration than population at the destination, we found a slightly larger effect of population at the destination, except in the basic gravity models for 2005 and 2010. This difference in findings could be caused in part by the difference in the definition of migration (recent migration in our study versus lifetime migration in Van Lottum and Marks's study), but it could also indicate that population at the destination had become an increasingly important pull factor for migration. This latter interpretation would be consistent with another finding of Van Lottum and Marks (2012), that population at the destination became an increasingly important factor in migration during 1971–2000.

For the basic gravity model, the effect of distance on migration was negative and highly significant. As expected, adding more variables to this model led to a decrease in the effect of distance on migration (Greenwood 1969a; Levy and Wadycki 1974; Schwartz 1973). In their study of Venezuela, Levy and Wadycki (1974) found that adding more variables to a basic gravity model reduced the estimated coefficient of distance by almost 50% (from −1.04 to −0.42). Our results also showed

TABLE 3 *Descriptive Statistics for the Variables, 2000–2010*

Variable	2000				2005				2010			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Migration flows from i to j (M_{ij}) (no. of migrants)	42,019	94,370	728	780,314	26,978	50,699	0	453,769	37,766	72,507	528	691,383
Size of population at origin (P_i) (million people)	15.39	9.85	3.71	35.85	16.40	10.07	4.17	37.00	18.28	11.61	4.49	46.15
Size of population at destination (P_j) (million people)	15.39	9.85	3.71	35.85	16.40	10.07	4.17	37.00	18.28	11.61	4.49	46.15
Geographical distance between origin & destination (D_{ij}) (km)	1,256	1,022	20	4,407	1,256	1,022	20	4,407	1,256	1,022	20	4,407
Regional GDP per capita at origin ($GDPcap_i$) (Rp million)	8.16	6.12	3.60	27.30	9.34	7.49	4.04	33.18	11.23	9.23	4.96	40.75
Regional GDP per capita at destination ($GDPcap_j$) (Rp million)	8.16	6.12	3.60	27.30	9.34	7.49	4.04	33.18	11.23	9.23	4.96	40.75
Natural log of percentage of workers in agriculture at origin [$\ln(AGRI_i)$]	37.84	17.52	0.23	60.59	36.05	19.58	0.24	60.12	32.09	17.16	0.61	53.12
Natural log of percentage of workers in agriculture at destination [$\ln(AGRI_j)$]	37.84	17.52	0.23	60.59	36.05	19.58	0.24	60.12	32.09	17.16	0.61	53.12
Natural log of percentage of highly educated workers at origin [$\ln(E_i)$]	6.15	3.92	2.63	15.10	6.97	3.30	2.47	14.55	8.66	2.93	5.01	15.92
Natural log of percentage of highly educated workers at destination [$\ln(E_j)$]	6.15	3.92	2.63	15.10	6.97	3.30	2.47	14.55	8.66	2.93	5.01	15.92
Unemployment rate at origin (U_i)	7.18	3.10	3.29	13.22	12.04	3.75	7.66	20.32	7.83	2.71	3.55	11.84
Unemployment rate at destination (U_j)	7.18	3.10	3.29	13.22	12.04	3.75	7.66	20.32	7.83	2.71	3.55	11.84
Contiguity (dummy variable):												
Regions share common border (dC_{ij})	0.10	0.30	0.00	1.00	0.10	0.30	0.00	1.00	0.10	0.30	0.00	1.00
Regions separated mostly by land (dL_{ij})	0.27	0.44	0.00	1.00	0.27	0.44	0.00	1.00	0.27	0.44	0.00	1.00
Regions separated by sea/ocean (reference)												
Migrant stock (S_{ij})					8.33	14.13	0.29	85.38	8.33	13.19	0.00	87.88

Source: Authors' statistical results.

Note: SD = standard deviation. Number of observations for each period = 156.

TABLE 4 *Poisson Regression Results for Basic and Modified Gravity Models, 2000–2010*

	Model 1 (basic gravity model)			Model 2 (1st modified gravity model)			Model 3 (2nd modified gravity model)	
Explanatory variable	2000	2005	2010	2000	2005	2010	2005	2010
Constant	11.80***	10.80***	10.72***	1.17	3.74	3.01	0.92	2.98**
	(0.89)	(0.91)	(0.89)	(3.24)	(5.10)	(5.35)	(1.99)	(1.48)
Natural log of population size at origin [ln(P_i)]	0.48	0.70***	0.70***	0.34	1.03***	0.72***	0.51***	0.17**
	(0.38)	(0.19)	(0.18)	(0.22)	(0.22)	(0.18)	(0.13)	(0.09)
Natural log of population size at destination [ln(P_j)]	0.58***	0.56***	0.65***	1.14***	1.24***	1.21***	1.25***	1.03***
	(0.19)	(0.16)	(0.14)	(0.21)	(0.22)	(0.18)	(0.14)	(0.10)
Natural log of geographical distance [ln(D_{ij})]	−0.63***	−0.65***	−0.64***	−0.01	−0.44**	−0.40**	−0.31***	−0.06
	(0.10)	(0.09)	(0.10)	(0.17)	(0.22)	(0.16)	(0.09)	(0.04)
Natural log of GDP per capita at origin [ln(GDP_i)]				−0.05	0.29	0.07	0.34*	−0.10
				(0.27)	(0.34)	(0.29)	(0.18)	(0.12)
Natural log of GDP per capita at destination [ln(GDP_j)]				1.49***	0.78**	0.38	0.56***	0.03
				(0.42)	(0.33)	(0.32)	(0.12)	(0.16)
Natural log of % of workers in agriculture at origin [ln($AGRI_i$)]				−0.19	−0.11	−0.25	0.04	−0.12*
				(0.12)	(0.14)	(0.18)	(0.07)	(0.07)
Natural log of % of workers in agriculture at destination [ln($AGRI_j$)]				0.41***	0.19	0.29	0.10*	−0.03
				(0.14)	(0.15)	(0.19)	(0.06)	(0.06)
Natural log of % of highly educated workers at origin [ln(E_i)]				−0.26	0.33	−0.14	0.33*	−0.04
				(0.31)	(0.33)	(0.53)	(0.20)	(0.29)
Natural log of % of highly educated workers at destination [ln(E_j)]				0.93***	0.95***	2.21***	0.94***	1.52***
				(0.25)	(0.25)	(0.52)	(0.16)	(0.31)
Natural log of unemployment rate at origin [ln(U_i)]				0.09	−0.84**	−0.59**	−0.06	−0.21
				(0.32)	(0.43)	(0.25)	(0.23)	(0.14)
Natural log of unemployment rate at destination [ln(U_j)]				−0.01	0.01	−0.03	0.14	0.34***
				(0.26)	(0.42)	(0.31)	(0.20)	(0.12)
Dummy for regions that share a border [dC_{ij}]				2.32***	0.81	0.84	−0.47	0.32**
				(0.67)	(0.83)	(0.68)	(0.32)	(0.17)
Dummy for regions separated mostly by land [dL_{ij}]				0.72*	0.25	0.19	−0.05	0.05
				(0.40)	(0.44)	(0.43)	(0.19)	(0.19)
Natural log of migrant stock [ln(S_{ij})]							0.82***	0.79***
							(0.06)	(0.05)
R^2	0.2350	0.3590	0.3588	0.6681	0.6287	0.7171	0.9071	0.9247

Source: Authors' statistical results. *Note:* Standard errors are in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

diminishing negative effects of distance after adding more variables. For example, in the basic gravity model for the year 2000, the effect was −0.63 and statistically significant at 1%. In model 2, the effect was −0.01 and statistically insignificant. In the year 2010, the estimated coefficient for distance in the basic gravity model was −0.64 (statistically significant at 1%), decreasing to −0.40 (statistically significant at 5%) in model 2 and falling further to −0.06 (statistically insignificant) in model 3.

According to our descriptive findings, the average distance of migration increased from 607 kilometres in 2000 to 631 kilometres in 2005 and 673 kilometres in 2010. We did not, however, find strong indications in the models that the effect of distance had weakened over time. In the basic model, the effect of distance was about the same in 2000, 2005, and 2010; in model 2, it was more strongly negative in 2005 and 2010 than in 2000; and only in the third model was the effect less negative in 2010 than in 2005. Thus, we found only weak support for our hypothesis that the effect of distance would diminish over time. Nevertheless, the findings are more in line with our hypothesis than with the earlier finding of Van Lottum and Marks (2012) that the effect of distance on migration increases over time—possibly because of the different definitions used for migration (recent migration in our study versus lifetime migration in theirs). Our findings on the distance decay effect are also in line with studies from China (Fan 2005; Poncet 2006; Shen 2012).

A negative effect of GDP at the origin and a positive effect of GDP at the destination would clearly indicate that a lack of economic development in origin regions triggers migration towards more developed regions. The coefficients for per-capita GDP at the destination showed the expected signs, although they seemed to decrease through time and were no longer statistically significant in the 2010 models. However, in none of the models was the effect of GDP at the origin significantly negative. In some models the effect was positive, but the evidence for a positive effect was weak (the coefficient for 2005 was significant in model 3, but only at 10%). This effect of GDP at the origin is not in line with Massey's (1988) argument that migration may be positively related to the level of economic development in the region of origin, not just the region of destination. However, because the GDP coefficients at the destination were larger than those at the origin, and most of the GDP coefficients at the origin were statistically insignificant, the findings might indicate that, in terms of regional development, the pull forces of destination areas are stronger than the push forces of origin areas. The use of GDP excluding oil and gas showed the same signs and statistical significance as the use of GDP with oil and gas, but with slightly different values for the beta coefficients (results not shown).

Another proxy for economic development, the share of agricultural workers, showed mostly insignificant effects on migration at both the origin and the destination. However, the signs of the coefficients for this variable were mostly negative at the origin and positive at the destination. These findings are in line with a study by Butzer, Mundlak, and Larson (2003) that shows that migration rates from agricultural to non-agricultural areas in Indonesia are relatively low compared with those of other countries, implying that labour surpluses have not been reallocated at a fast pace to other sectors of the economy. A partial explanation of this finding could be that the share of agricultural workers may change, owing not to migration but to a shift from the agricultural sector to non-agricultural sectors within one region.

The estimated coefficients for education at the destination were as expected (positive and statistically significant). This finding is in accordance with the theoretical expectation that potential migrants are attracted to regions with high educational attainment. For the origin, however, only one model displayed the expected positive and statistically significant coefficient (model 3 for year 2005). The negative effect on migration of education at the origin is in line with the findings of Greenwood (1969b) for Egypt, Lucas (1985) for Botswana, and Quinn and Rubb (2005) for Mexico.

The estimated coefficients for the unemployment rate at the origin were mainly negative but statistically insignificant, while the coefficients for the unemployment rate at the destination were mostly positive and insignificant. This could be because of disguised unemployment in rural areas in the form of underemployment (Greenwood 1969a).

As expected, nearly all of the coefficients for the contiguity dummy *dC* were positive, but they were significant only in model 2 for 2000 and in model 3 for 2010. The same pattern can be observed for *dL*, with only one significantly positive coefficient (the coefficient in model 2 for 2000). Thus, the evidence for the effect of borders on migration in Indonesia is weak, despite the signs being as expected. These findings are also in line with Van Lottum and Marks (2012), who found a decreasing importance of contiguity in explaining migration in Indonesia.

The migrant stock variable, a proxy for social network size, showed a strong positive and statistically significant effect on interregional migration in Indonesia. In line with Greenwood's (1969a) findings, the inclusion of the migrant stock variable in our model reduced the effects of several other variables, for example, distance and unemployment rate at the origin. The migrant stock variable captures not only the network effect of migration, but also the past cumulative effects of migration forces. This is suggested by the decreasing effect of distance and borders (proxies for the costs of moving), which is consistent with previous studies in other countries that used the same framework (Fan 2005; Greenwood 1969a). The findings therefore indicate a positive impact of social networks on migration and support the importance of cumulative causation in migration.

CONCLUSION AND DISCUSSION

Previous research on interregional migration in Indonesia found a strong indication of concentration (urbanisation and overurbanisation) but also a weak indication of deconcentration (suburbanisation and metropolitan to non-metropolitan migration) (Wajdi, Van Wissen, and Mulder 2015). In this study, we employed basic and modified (extended) gravity models, using data from the 2000 and 2010 Population Censuses and the 2005 Intercensal Population Survey (Supas), to explore the determinants of interregional migration in Indonesia. We aimed to test Long's (1985) hypothesis that shifts in population settlement patterns (population redistribution)—that is, concentration and deconcentration of the population—have a strong relationship with economic development. In particular, we wanted to test the hypothesis that economic development is positively correlated with population concentration.

In line with the classical gravity model, we find a positive effect of the population size of the destination on migration. Distance is negatively related to the size of migration flows. Based on the positive and significant effect of GDP per capita at the destination, we conclude that migration is directed towards more developed regions. This finding confirms Long's thesis that population redistribution has a positive relationship with economic development, as indicated by population concentration in more developed regions during the early stages of development. In our models, we do not see any sign of a deconcentration of population in the later stages of development, although previous research for specific areas such as Jakarta and Bodetabek did find evidence of deconcentration (Wajdi, Van Wissen, and Mulder 2015). This could be due to the short timespan used in this research, whereas Long's research focused on trends over longer timespans. It would therefore be advisable to conduct such research using longer timespans, for example, by using the latest (2015) intercensal data to identify the latest trends.

We used data from the 2005 Intercensal Population Survey together with data from the 2000 and 2010 censuses to allow more detailed analysis of migration flows during 2000–2010. The intercensal survey is a national survey designed to permit estimation up to the district level (415 districts), and to provide demographic data between census dates. It should be borne in mind, however, that its sample size is relatively small and that the trends suggested by the data could be caused by sampling error.

The findings from the socioeconomic variables used in this study (GDP, unemployment rate, educational attainment, and share of agricultural sector) support the notion that regional disparities in development are an important factor in interregional migration in Indonesia. They also suggest that interregional migration is predominantly a response to pull rather than push forces, and that, over time, the influence of push forces has decreased whereas that of pull factors has increased. These findings suggest the importance of creating pull forces in new areas, especially outside Java. The Indonesian government under President Joko Widodo is encouraging the development of the Tol Laut (Sea Highway), that is, the strengthening of sea transportation by connecting seaports with each other. This will increase the volume of trade and improve the distribution of products that fulfil consumer demand. Educational facilities—not just senior high school but also tertiary facilities—also need to be expanded. A success story in building higher-education facilities that attract migrants away from metropolitan areas can be found in Depok, where Universitas Indonesia is located, and in Jatinangor in West Java, where Universitas Padjajaran is located.

One variable that is rarely studied in migration studies in Indonesia, migrant stock as a representation of the social network effect on migration, shows a statistically significant effect. This is in line with findings from studies in developed nations. The positive and statistically significant effect of this variable indicates a positive impact of social networks on migration and also a cumulative causation of migration that captures the collective effects of past migration forces. The strong ties of migrants with relatives and friends at the origin could have both positive and negative effects. One positive effect is the flow of remittances to the origin, which could be useful in lifting those areas' living standards. One negative effect is that migrants may be attracted to regions that are already developed, causing more congestion in those areas. It is therefore necessary to pay attention to the development of infrastructure and human resources in the origin areas, not just the destination areas.

This article has shown that gravity models are useful in explaining migration in Indonesia, in both theoretical and methodological terms. By including a set of variables that represents variations in regional characteristics and disparities in regional economic development, and by introducing a migrant stock variable, our analysis helps to explain migration flows in Indonesia. The study also provides additional evidence that existing migration theories and experience from other countries are relevant for conceptualising population movements in Indonesia.

A major advantage of gravity models (and other models that focus on macro-level migration flows) is that they allow the inclusion of characteristics of both destination regions and origin regions. Although the gravity model of migration has led to more advanced modelling of migration, however, it is not suitable for the inclusion of micro factors (Greenwood 1997). The focus on macro-level migration flows rather than the micro-level behaviour of individuals can be seen as the main drawback of our models. In order to more fully understand the migration phenomenon in Indonesia, researchers should be encouraged to investigate inter-regional migration further, using models that facilitate a micro-level approach.

REFERENCES

Adams, Dale W. 1969. 'Rural Migration and Agricultural Development in Colombia'. *Economic Development and Cultural Change* 17 (4): 527–39.

Alatas, Secha. 1993. 'Macro Patterns of Internal Migration in Indonesia, 1971–1990'. *Indonesian Journal of Demography* 20 (40): 21–47.

Anderson, James. 1979. 'A Theoretical Foundation for the Gravity Equation'. *American Economic Review* 69 (1): 106–16.

Bappenas (Ministry of National Development Planning). 2015. *Laporan Nasional Evaluasi Kinerja Pembangunan Daerah di 33 Provinsi Tahun 2014* [National Report Evaluating Regional Development Performance in 33 Provinces in 2014]. Jakarta: Bappenas.

Beals, Ralph E., Mildred B. Levy, and Leon N. Moses. 1967. 'Rationality and Migration in Ghana'. *Review of Economics and Statistics* 49 (4): 480–86.

Beine, Michel, Simone Bertoli, and Jesús Fernández-Huertas Moraga. 2014. 'A Practitioners' Guide to Gravity Models of International Migration'. Documento de Trabajo 2014-03, FEDEA, Madrid. http://documentos.fedea.net/pubs/dt/2014/dt-2014-03.pdf.

Bodvarsson, Örn B., and Hendrik van den Berg. 2013. *The Economics of Immigration Theory and Policy*. 2nd ed. New York: Springer.

Borjas, George J. 2001. 'Economics of Migration'. In *International Encyclopedia of the Social & Behavioral Sciences*, edited by Neil J. Smelser and Paul B. Baltes, 9803–9. Oxford: Pergamon.

———. 2013. 'Labor Mobility'. In *Labor Economics*. 6th ed., 318–66. New York: McGraw-Hill/Irwin.

BPS (Statistics Indonesia). 2014. *Indikator pasar tenaga kerja indonesia agustus 2014* [Labour market indicators in Indonesia August 2014]. Jakarta: BPS. https://bps.go.id/index.php/publikasi/899.

Brown, Judith E., and Peter K. Dunn. 2011. 'Comparisons of Tobit, Linear, and Poisson–Gamma Regression Models: An Application of Time Use Data'. *Sociological Methods & Research* 40 (3): 511–35.

Bunea, Daniela. 2012. 'Modern Gravity Models of Internal Migration: The Case of Romania'. *Theoretical and Applied Economics* 19 (4): 127–44.

Butzer, Rita, Yair Mundlak, and Donald F. Larson. 2003. 'Intersectoral Migration in Southeast Asia: Evidence from Indonesia, Thailand, and the Philippines'. Policy Research Working Paper 2949, World Bank, Washington, DC, January. https://openknowledge.worldbank.org/bitstream/handle/10986/19171/multi0page.pdf?sequence=1.

Chotib. 1998. 'Skedul model migrasi dari DKI Jakarta/luar DKI Jakarta: Analisis data Supas 1995 dengan pendekatan demografi multiregional' [Model migration schedule for Jakarta and outside Jakarta: Analysis of data from Intercensal Population Survey 1995 using a multi-regional demographic approach]. Master's diss., Universitas Indonesia.

Conley, Timothy G., and Giorgio Topa. 2002. 'Socio-Economic Distance and Spatial Patterns in Unemployment'. *Journal of Applied Econometrics* 17 (4): 303–27.

Dahl, Gordon B. 2002. 'Mobility and the Return to Education: Testing a Roy Model with Multiple Markets'. *Econometrica* 70 (6): 2367–420.

Darmawan, Beny, and Chotib. 2007. 'Perkiraan pola migrasi antarprovinsi di Indonesia berdasarkan indeks ketertarikan ekonomi' [Estimated migration patterns among provinces in Indonesia based on the index of economic attraction]. Paper presented at Parallel Session IIIC: Poverty, Population and Health, Demographic Institute, Faculty of Economics, Universitas Indonesia, Depok.

Dhanani, Shafiq. 2004. 'Unemployment and Underemployment in Indonesia, 1976–2000: Paradoxes and Issues'. SES Working Paper, InFocus Programme on Socio-Economic Security (SES), ILO, Geneva, January. http://www.ilo.int/public/english/protection/ses/download/docs/indonesia.pdf.

Etzo, Ivan. 2008. 'Internal Migration: A Review of the Literature'. MPRA Paper 8783, Munich Personal RePEc Archive (MRPA), May. http://mpra.ub.uni-muenchen.de/8783/1/MPRA_paper_8783.pdf.

Fan, C. Cindy. 2005. 'Modeling Interprovincial Migration in China, 1985–2000'. *Eurasian Geography and Economics* 46 (3): 165–84.

Firman, Tommy. 1994. 'Migrasi antarprovinsi dan pengembangan wilayah di Indonesia' [Interprovincial migration and the development of the regions in Indonesia]. *Prisma* 7 (XXIII): 3–15.

Flowerdew, Robin, and Murray Aitkin. 1982. 'A Method of Fitting the Gravity Model Based on the Poisson Distribution'. *Journal of Regional Science* 22 (2): 191–202.

Fussell, Elizabeth, and Douglas S. Massey. 2004. 'The Limits to Cumulative Causation: International Migration from Mexican Urban Areas'. *Demography* 41 (1): 151–71.

Girsberger, Esther Mirjam. 2015. 'Essays on Migration, Education and Work Opportunities'. PhD diss., European University Institute.

Greenwood, Michael J. 1969a. 'An Analysis of the Determinants of Geographic Labor Mobility in the United States'. *Review of Economics and Statistics* 51 (2): 189–94.

———. 1969b. 'The Determinants of Labor Migration in Egypt'. *Journal of Regional Science* 9 (2): 283–90.

———. 1975. 'Research on Internal Migration in the United States: A Survey'. *Journal of Economic Literature* 13 (2): 397–433.

———. 1997. 'Internal Migration in Developed Countries'. In *Handbook of Population and Family Economics, Volume 1B*, edited by Mark R. Rosenzweig and Oded Stark, 647–720. Amsterdam: Elsevier BV.

Greenwood, Michael J., and John M. McDowell. 1991. 'Differential Economic Opportunity, Transferability of Skills, and Immigration to the United States and Canada'. *Review of Economics and Statistics* 73 (4): 612.

Handiyatmo, Dendi. 2009. 'Penggunaan jenis transportasi oleh pelaku mobilitas ulang alik di enam kawasan metropolitan: Analisis data Supas 2005' [Choice of modes of transportation by commuters in six metropolitan areas: Analysis of data from Intercensal Population Survey 2005]. Master's diss., Universitas Indonesia.

Hilbe, Joseph. 1999. 'Robust Variance Estimators for MLE Poisson and Negative Binomial Regression'. *Stata Technical Bulletin* 8 (45): 26–28.

Kim, Keuntae, and Joel E. Cohen. 2010. 'Determinants of International Migration Flows to and from Industrialized Countries: A Panel Data Approach beyond Gravity'. *International Migration Review* 44 (4): 899–932.

Lessem, Rebecca. 2009. 'Labor Migration in Malaysia'. Unpublished manuscript, University of Wisconsin-Madison.

Levy, Mildred B., and Walter J. Wadycki. 1974. 'Education and the Decision to Migrate: An Econometric Analysis of Migration in Venezuela'. *Econometrica: Journal of the Econometric Society* 42 (2): 377–88.

Long, John F. 1985. 'Migration and the Phases of Population Redistribution'. *Journal of Development Economics* 17 (1–2): 29–42.

Lucas, Robert E. B. 1985. 'Migration amongst the Batswana'. *Economic Journal* 95 (378) 358–82.

Massey, Douglas S. 1988. 'Economic Development and International Migration in Comparative Perspective'. *Population and Development Review* 14 (3): 383–413.

———. 1990. 'Social Structure, Household Strategies, and the Cumulative Causation of Migration'. *Population Index* 56 (1): 3–26.

Mayer, Thierry, and Soledad Zignago. 2011. 'Notes on CEPII's Distances Measures: The GeoDist Database'. CEPII Working Paper 2011-25, CEPII, Paris. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1994531.

Minami, Ryoshin. 1967. 'Population Migration away from Agriculture in Japan'. *Economic Development and Cultural Change* 15 (2): 183–201.

Myrdal, Gunnar. 1957. *Rich Lands and Poor*. New York: Harper and Row.

Niedercorn, John H., and Burley V. Bechdolt, Jr. 1969. 'An Economic Derivation of the Gravity Law of Spatial Interaction'. *Journal of Regional Science* 9 (2): 273–82.

Öberg, Sture. 1997. 'Theories on Inter-Regional Migration: An Overview'. In *People, Jobs and Mobility in the New Europe*, edited by Hans H. Blotevogel and Anthony J. Fielding, 3–22. Chichester: Wiley.

Peeters, Ludo. 2012. 'Gravity and Spatial Structure: The Case of Interstate Migration in Mexico'. *Journal of Regional Science* 52 (5): 819–56.

Poncet, Sandra. 2006. 'Provincial Migration Dynamics in China: Borders, Costs and Economic Motivations'. *Regional Science and Urban Economics* 36 (3): 385–98.

Quinn, Michael A., and Stephen Rubb. 2005. 'The Importance of Education–Occupation Matching in Migration Decisions'. *Demography* 42 (1): 153–67.

Ravenstein, Ernst G. 1885. 'The Laws of Migration'. *Journal of the Statistical Society of London* 48 (2): 167–235.

Rogers, Andrei, Salut Muhidin, Lisa Jordan, and Megan Lea. 2004. 'Indirect Estimates of Age-Specific Interregional Migration Flows in Indonesia Based on the Mobility Propensities of Infants'. Working Paper POP2004-0001, Research Program on Population Processes, Institute of Behavioral Science, University of Colorado.

Rozelle, Scott, J. Edward Taylor, and Alan DeBrauw. 1999. 'Migration, Remittances, and Agricultural Productivity in China'. *American Economic Review* 89 (2): 287–91.

Sahara, Ida. 2010. 'Pola waktu tempuh pekerja dalam melakukan mobilitas ulang-alik di kota metropolitan Indonesia tahun 2008' [Commute times for workers in executing commuting mobility in metropolitan Indonesia in 2008]. Master's diss., Universitas Indonesia.

Sahota, Gian S. 1968. 'An Economic Analysis of Internal Migration in Brazil'. *Journal of Political Economy* 76 (2): 218–45.

Schwartz, Aba. 1973. 'Interpreting the Effect of Distance on Migration'. *Journal of Political Economy* 81 (5): 1153–69.

Shen, Jianfa. 2012. 'Changing Patterns and Determinants of Interprovincial Migration in China 1985–2000'. *Population, Space and Place* 18 (3): 384–402.

Silva, J. M. C. Santos, and Silvana Tenreyro. 2006. 'The Log of Gravity'. *Review of Economics and Statistics* 88 (4): 641–58.

———. 2011a. 'Poisson: Some Convergence Issues'. *Stata Journal* 11 (2): 207–12.

———. 2011b. 'Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator'. *Economics Letters* 112 (2): 220–22.

Sjaastad, Larry A. 1962. 'The Costs and Returns of Human Migration'. *Journal of Political Economy* 70 (5): 80–93.

Spreitzhofer, Günter. 2005. 'Post-Suharto's Jabotabek Region: New Issues of Demographic and Socio-Economic Change in Western Java'. *Geografia: Malaysian Journal of Society and Space* 1 (1): 1–10.

Stark, Oded, and Shlomo Yitzhaki. 1988. 'Labour Migration as a Response to Relative Deprivation'. *Journal of Population Economics* 1 (1): 57–70.
Todaro, Michael P. 1969. 'A Model of Labor Migration and Urban Unemployment in Less Developed Countries'. *American Economic Review* 59 (1): 138–48.
———. 1980. 'Internal Migration in Developing Countries: A Survey'. In *Population and Economic Change in Developing Countries*, edited by Richard A. Easterlin, 361–402. Chicago: University of Chicago Press.
Van Lottum, Jelle, and Daan Marks. 2012. 'The Determinants of Internal Migration in a Developing Country: Quantitative Evidence for Indonesia, 1930–2000'. *Applied Economics* 44 (34): 4485–94.
Wajdi, Nashrul M. 2010. 'Migrasi antarpulau di Indonesia: analisis model skedul migrasi dan model gravitasi hybrida' [Interisland migration in Indonesia: Model migration schedules and hybrid gravity model analysis]. Master's diss., Universitas Indonesia.
Wajdi, Nashrul, Leo J. G. van Wissen, and Clara H. Mulder. 2015. 'Interregional Migration Flows in Indonesia'. *Sojourn: Journal of Social Issues in Southeast Asia* 30 (2): 371–422.
World Bank. 2012. *Indonesia. The Rise of Metropolitan Regions: Towards Inclusive and Sustainable Regional Development*. Washington, DC: World Bank.
Zipf, George Kingsley. 1946. 'The P_1P_2/D Hypothesis: On the Intercity Movement of Persons'. *American Sociological Review* 11 (6): 677–86.

TABLE A1 *Summary of the Division of Indonesia into 13 Regions*

Region	Remarks
1. Jakarta	Jakarta (the Special Capital Region of Jakarta, or DKI Jakarta) is Indonesia's capital and the country's biggest city in terms of both population size and economy. In 2005, Jakarta was the world's 11th largest city, and one of 16 megacities in developing countries (Spreitzhofer 2005; World Bank 2012). Jakarta consists of 1 district (*kabupaten*) (Kepulauan Seribu) and 5 municipalities (*kota*) (East, West, North, South & Central Jakarta).
2. Bodetabek (Bogor, Depok, Tangerang & Bekasi)	Bodetabek is the metropolitan area surrounding Jakarta. It consists of 3 districts (Bekasi, Bogor, and Tangerang) and 4 municipalities (Bekasi, Bogor, Depok, and Tangerang). This area is part of the greater Jakarta metropolitan region (Jabodetabek).
3. Bandung Raya	This metropolitan area is located in West Java province. It consists of 2 districts (Bandung and West Bandung) and 2 municipalities (Bandung and Cimahi).
4. Rest of West Java & Banten	This area comprises the rest of West Java and Banten provinces, except Jakarta, Bodetabek & Bandung Raya.
5. Kedungsepur (Kendal, Demak, Semarang, Salatiga & Grobogan)	This metropolitan area is located in Central Java province. It consists of 4 districts (Demak, Grobogan, Kendal & Semarang) and 2 municipalities (Salatiga and Semarang).
6. Rest of Central Java & Yogyakarta	This area comprises the rest of Central Java and Yogyakarta, except Kedungsepur. Yogyakarta is not considered to be a separate metropolitan area because most of its population works in the agricultural sector (Handiyatmo 2009; Sahara 2010).
7. Gerbangkertosusila (Gresik, Bangkalan, Mojokerto, Surabaya, Sidoarjo & Lamongan)	This metropolitan area is located in East Java province. It consists of 5 districts (Bangkalan, Gresik, Lamongan, Mojokerto & Sidoarjo) and 2 municipalities (Mojokerto and Surabaya).
8. Rest of East Java	This area comprises the rest of East Java province, except Gerbangkertosusilo.
9. Mebidangro (Medan, Binjai, Deli Serdang & Tanah Karo)	This metropolitan area is located in Sumatra. It consists of 2 districts (Deli Serdang and Karo) and 2 municipalities (Medan and Binjai).
10. Rest of Sumatra	This area comprises the rest of Sumatra, except Mebidangro.
11. Kalimantan	Kalimantan is one of Indonesia's 5 biggest islands. It consists of 5 provinces.
12. Sulawesi	Sulawesi is one of Indonesia's 5 biggest islands. It consists of 6 provinces.
13. Rest of Indonesia	This area comprises the remaining 7 provinces, namely Bali, East Nusa Tenggara, West Nusa Tenggara, Maluku, North Maluku, Papua & West Papua. Papua (also known as Irian Jaya) is the biggest island in this region & accounts for 21.8% of Indonesia's land area.

Source: Authors' compilation.

FIGURE A1 *Map of Indonesia (with inset map of Mebidangro)*

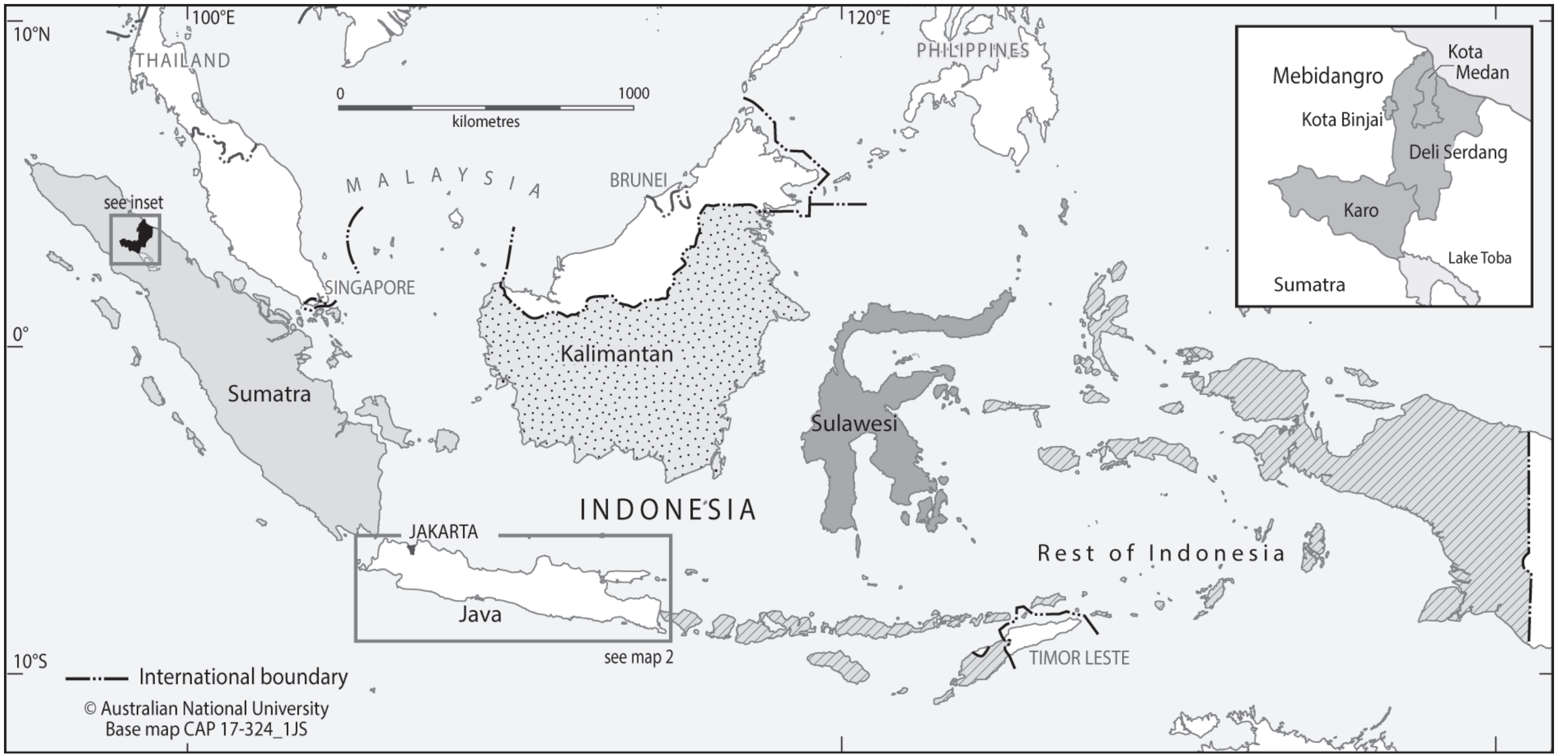


FIGURE A2 *Map of Java*