INTER-PROVINCIAL MIGRATION AND REGIONAL DEVELOPMENT IN INDONESIA 1995-2000: A Data Analysis on Population Census 2000 with Application of *Spatial Interaction Model*

Chotib
Demographic Institute
Faculty of Economics University of Indonesia
Email address: chotib@hotmail.com; chotib@ldfeui.org

ABSTRACT

There are many migration theories have been applied by various interdisciplinary scholars. However, determinant factors of migration phenomenon largely depend upon the context of its study. In the context of maximization of individual welfare and existence of regional development disparities in Indonesia, migration analysis approached by place to place migration model is very important. The result of this research is expected to give more understanding on population distribution in Indonesia, and its relation to regional development with specific various regional characteristics, and policy implementation on economic and population development as well, specifically in the context of regional autonomy era.

The model developed by this study concludes that high regional domestic product in any region tends to increase volume of out migration from this region and decrease volume of in migration into this region. High rate of urbanization in any region tends to "hold on" population move out of this region and "pull-in" population entering into this region. High industrialization in any region tends to increase interaction between two highly industrialized regions. This study also shows that migration stream between two regions is mainly dominated by contiguous regions.

BACKGROUND

The pattern of spatial migration flow from one region to the other region most likely reflects the regional development disparity and interdependency among regions. It also reflects the flow of human resources among regions. As it is generally shown by the developing countries, the pattern of migration flows in Indonesia tend to show a polarization, centralization of migration flows to special regions, especially to big cities and Java Island (Firman, 1994). This polarization of migration flows, for sure, reflects the polarization of development pattern in Indonesia.

The movement or population migration from one area to the other area mostly is based on the economic motivation (Ananta and Wongkaren, 1996). In this case, people who intend to move, or even to stay, have already taken into consideration the advantage and disadvantage that possible faced (McConnell dan Brue, 1989). This planned movement is called by Tjiptoherijanto (2000) as 'voluntary planned migration'. This term is used to differ planned migration from the movement caused by non-economic factors such as political riot, condemnation, natural disaster, and other movement called *"migrasi dukalara"* or impelled movement.

There are many theories developed based on the economic point of view to analyze the migration phenomena. Two major theories are the most relevance to our study, i.e. neo-classical economics and new economics of migration. Theory based on neo-classical economics, whether from the macro or micro approach, most likely focuses on difference of wages and working conditions among regions or countries, and costs of individual consideration to migrate. According to neo-classical theory, population movement or migration is an individual decision based on willingness to reach his/her maximum welfare.

In the other side, the new economics of migration approach considered that population movement or population mobility occurs not related only to labor market but also related to the other factors. The decision to migrate is not just an individual decision but also related to environment conditions, especially family and living condition, and place of destination as well.

In Indonesia, the implementation of migration theories has already been done by researchers from several field of studies, such as sociology, population and economic studies. The weight of each factor related to the migration phenomena, however, still depends on each context of the research conducted. Therefore, migration is put in the context of development processes, the impact of socio-economic conditions, government policies, infrastructure provided and level of technology development reached. In other word, migration is influenced by different factors in different development level.

As a rational process to maximize the individual's welfare and to minimize the different level of development between regions in Indonesia, migration analysis based on 'place to place migration' approach is very important. Through this approach, the analysis is expected to be able to give a better understanding on inter-regional migration in Indonesia related to regional development disparity based on specific characteristics of each region and policy implication implemented on economic development and population as well, especially in the autonomy era.

In the context of development of sciences, this study contributes to enrich regional economics, geography and demography through 'place to place migration model' in one of developing countries, Indonesia.

MIGRATION MODELS IN ECONOMIC DEVELOPMENT

The economic model based on population mobility process, especially the inter-regional employment among regions, first was developed by Sir Arthur Lewis in 1954, and formulized and developed by Fei and Ranis (1961). In Lewis-Fei-Ranis model (LFR), the economy consists of two sectors, they are (a) subsistent traditional sector in rural areas characterized by overflow employment with very low productivity, and (b) modern industrial sector in urban areas with high productivity as a place for traditional employment sector moves slowly to (Todaro, 1976).

.The main focus of the above model is transfer process of employment and work opportunity growth caused by output expansion in modern sector. The quick of output expansion depends on rate of capital accumulation that is possible provided by the re-investment of business profits done the capital owners. The rate of wages in this sector assumes constant and decided as constant multiplier above the rural wages. Lewis (1954) assumed that the level of urban wages at least is 30 percent higher than the average of rural income for attracting employment to migrate.

LFR model is considered simple and able to explain the economic growth to western countries. However, this model implicitly assumes three important things deviated from migration incidents and condition of under development of developing countries. These assumptions are:

First, the rate of employment mobility and the creation of work opportunities in urban sector are proportional to the rate of capital accumulation. The sooner the rate of capital accumulation, the higher the growth of modern sector, and the quicker the rate of employment creation. This assumption is different from the previous assumption and has also a different characteristic, for example, this assumption has the characteristic of labor saving investment.

Second, rural area has the characteristic of labor surplus, while modern sector is full employment. In general, the study shows that urban sector has the characteristic of unemployment and under employment, while the rural sector has a few labor surplus, especially in the planting and harvesting event. A study by Chotib (2001a) shows that in Indonesia the unemployment rate in the urban areas is higher than in the rural , based on various data and period.

Third, the rate of wages is constant until all labor surpluses in the rural areas absorbed. In fact, in most developing countries, the rate of wages in the urban areas tends to increase toward rate of wages in the rural, absolutely or relatively, although open-unemployment has been increasing.

Contrary to the assumption that migration mainly is based on migrant's economic rational consideration without taking care on employment in urban areas, Todaro model postulated that migration occurs as respond to difference of expected income between urban and rural areas. Basic premises of this model is that a migrant as a decision maker has already considered many opportunities available in two places and he/she chooses one of them that maximizing his/her expected gains of migration. This expected gains are being measured through: (a) differences of real income between working in the rural (place of origin) and in the urban (place of destination); and (b) probability for new coming migrant getting a job in the urban area.

Mathematicaly, Todaro model can be explained as follows. If *V (0)* is present value of net income during time planned of migration, . $Y_{u,r}(t)$ is average individual real income in urban area (place of destination) and rural area (place of origin); *n* is long periode of time planned of migration; and *I* is discount rate reflecting the rate of migrant's time preference. Therefore, the decision to migrate or not it depends on whether *V (0)* negative or positive:

$$V(0) = \int \left[p(t)Y_u(t) - Y_r(t) \right] e^{-it} dt - C(0)$$

where *C (0)* reflecting migration costs, and *p (t)* is probability of a migrant obtains urban job with average income at period *t*.

As proxy of probability, *p (t)* is employment rate in urban ($= E_u/L_u$). Besides, rate of migration is influenced also by different rate of wages that definites directly the flow of net expected income. The different wages are formulized by Todaro as ratio between $Y_u/Y_r = W$, and $W > 1$ assumed to be constant.

Decision to migrate is also influenced by costs of migration that could be related to intervening obstacles in migration theory of Lee, stated by Todaro as other factors, *Z,* such as distance, migrant's perception on advantages and disadvantages of available chances in place of origin and place of destination, the existence of contact person at place of destination, convenience of place of destination, etc.

By defining the rate of migration *m* as ratio of the real people migrate, *M,* against the number of employment at place of origin, *Lr*, the basic migration, according to Todaro is as follows :

$$m = F\left(\frac{E_u}{L_u}, W, Z\right)$$

$$F^{'}\left(\frac{E_u}{L_u}\right) > 0, F^{'}(W) > 0, F^{'}(Z) \gtrless 0$$

Todaro and John Harris his friend utilized and enlarged the Todaro's basic framework of thinking to develop the model of internal changes of two factors about migration and unemployment that possible attracts more explicitly the impact of migration on income in the rural (place of origin), rate of output either in urban or rural areas and social welfare as a whole. Two sectors mentioned are urban and rural sectors that are differentiated by production and income.

Harris-Todaro model is enlargement of neo-classic two-trade sector model. Thus, the assumption and pro-position of neo-classic model are still applied, among others are related to uniformity of the proportion of production technology level in the urban and rural areas, neo- classic behavior decision on level of production factor utilization and rate of output, and trade mechanism theory to determine of exchange rate between agricultural and manufacture commodities. The development done is caused by uniformity (homogeneity) migration of that two-trade sector.

Mathematically, Harris-Todaro model can be explained as follows. If each of *Wr* and *Wu* is rate of nominal wages in rural and urban areas, and E_u is the number of employment in urban, and L_u is of rural, then expected urban income is written as:

$$E(W_u) = W_u * \frac{E_u}{L_u}$$

While expected rural income is E (Wr) is Wr. The number of migrants from rural to urban areas is the function of difference between E (Wu) and E(Wr), and written as:

$$M = f(E(W_u) - E(W_r))$$

Condition of urban-rural balance is reached when the difference of expected income between the two sectors is none, that is:

$$E(W_u) = E(W_r) = W_r$$

$$W_u . \frac{E_u}{L_u} = W_r \quad \text{atau} \quad \frac{E_u}{L_u} = \frac{W_r}{W_u}$$

Based on this formula, at the level of balance there is urban unemployment rate as:

$$1 - \frac{E_u}{L_u} = 1 - \frac{W_r}{W_u}$$

In other word, urban unemployment rate is related reversely to the difference of expected income.

SOME APPLICATIONS OF THE MODEL

Place to place migration model applied by Field (1982) actually is development of spatial interaction model in geography. This model is one of spatial analysis model purposing to predict measurement and flow (for example, population, commodities, or information) by using several models related to characteristics of inter-regions.

A study done by Koestoer (1991) on accessibility of the purpose to place of work in Jabotabek areas used spacial-interaction model. This study tries to relate accessibility of place of work between Jakarta and spread of residence in Botabek areas. Two frameworks of estimation is used in this study, they are pattern based on estimation of distance and estimation of time related to accessibility on place of work.

Chotib (2001 b) done the analysis of inter-provincial migration by using spacial-interaction model on 1995 Supas data (1995 Intercensal Population Survey). The analysis shows that the migration flow was influenced positively by the number of population in both places (place of origin and of destination) and influenced negatively by distance (measured by long of straight line between two capitals of provinces and dummy variable: directly or indirectly borders between two provinces)

Some previous studies have done by Yunus (1985), Yosephine (1989) and Utama (1994), analyzed the migration by the similar ways. They have measured distances through the economic point of view, that was measured by ticket fare of plane between provinces.

In place to place migration model, analysis is done by flow of population migration between place of destination *j* and place of origin *i*. Economic migration model relates systematically to the rate of migration from area *i* to *j*, that is relates positively to the pull power of general labor market condition at place of destination *(Ej),* relates negatively to the costs of movement from *i* to *j* (*Cij),* that is:

$$M_{ij} = f(E_i, E_j, C_{ij})$$

<0 >0 <0

Usually, E_i and E_j is approximated by average income level, employment probability and other matters showing economic condition at each area, while C_{ij} is approximated by distance. Place to place migration analysis based on statement that the economic chances at each area has an important role to determine the allocation of spatial population. In the application of this model in Colombia, Field (1982) used the regional income variables and opportunities at labor market as economic variables explaining the occurrence of inter-regional migration. Income variable used are income at place of origin and at place of destination. Employment rates variable (percentage of the working population out of the labor force) at the place of origin and of destination is used as the reflection of labor force condition.

Demographic characteristics of migrants have important role on migration behavior. The older the person, the lower his/her intention to migrate. Similar case, male population tends mobile than the female population.

METHOD OF ANALYSIS AND PROCEDURES

In this study migration function is estimated by stepwise multiple regression model with estimation technique of Ordinary Least. Square (OLS) by using packet of SPSS version 11. Dependent variable of this function is number of migrants from province o (place of origin) to province *d* (place of destination) *($FLOW_{od}$)* in various alternative form of models and independent variables, they area;

a. PDRB_o, PDRB_d, RPIND_o, RPIND_d, EMPLOY_o, EMPLOY_d, URBAN_o, URBAN_d.
b. Dummy variable, such as characteristic of migrant (SEX), directly borders between province *o* and province *d* (CONT) and characteristic of area --urban/rural on province *d* (URRU).

Analysis is done on 30 provinces in Indonesia with the number of migrants from place *i* to place *j* as dependent variable and called as one observation. Migrant data computed from 200 Population Census published by Central Board of Statistics (2001), Serie L2.2. In this publication, migration data obtained from Table 12a presenting recent-migration flow among provinces in Indonesia during 1995-2000. Migrants are differed based on sex and urban/rural areas at recent residence. Unfortunately, that data hove no in formation on characteristic areas (urban/rural) of the province of origin.

The concept of migration used is recent migration that is population whose recent residence is different from residence five years ago. As macro side, it is expected that aggregation rate based on provinces is able to reveal the phenomena in the context of regional development in Indonesia.

Other data used are regional domestic product bruto (PDRB) of each province during 1995-1998. While, the industrialization of each region is proxy of development growth rate through investment in industry and manufacturing sectors in a certain region. That data are obtained by looking at the average industrial share on total PDRB during 1995-1998. Other data also indicate the difference of development between regions, that is average of employment in 1995 and 2000 and urbanization rates of each region during 1995 and 2000.

Distance variable reflects the economic cost of movement from one province to the other one. Economic cost is proxied by dummy variable CONT (contageous) with value "1" for a pair of province having direct border (for instance, West Java and DKI Jakarta Province), and value "0" for a pair of province having indirect border. For a pair of province separated by a strait is decided to have value "1", since transportation on crossing the strait is considered running better. This case can be seen at some pairs of provinces, among others are Provinces of Lampung-Banten, East Java-Bali, Bali-West Nusa Tenggara, and East-West Nusa Tenggara. For provinces separated by an ocean (not a strait) are considered as having direct borders although imaginary line of two provinces is located at the middle of these two provinces.

Model Specification

Inter-province migration model developed is as follows (Model 1):

$$FLOW = \beta_0 + \beta_1 * PDRB\_o + \beta_2 * PDRB\_d + \beta_3 * RPIND\_o + \beta_4 * RPIND\_d + \beta_5 * EMPLOY\_o + \beta_6 * EMPLOY\_d + \beta_7 * URBAN\_o + \beta_8 * URBAN\_d + \beta_9 * SEX + \beta_{10} * URRU + \beta_{11} * URRU + \beta_{12} * CONT$$

Attracment-1 on regression model for Model 1 shows regression coefficient (parameter β for each variable). It shows that the value of R^2 is measurement whether ordinary regression model is applicable or not. With very small value of R^2 (0,019242), these varieties of migration flows are able to be explained by explanatory variables showed by model above 1,9 percent. However, *F* test shows a very significant model (prob=0,0000), that can be assumed that we refuse hypothesis null which stated that there is no explanatory variable be able to explain the varieties of migration flows from their average.

To test each significant of explanatory variable, we have to see the value of *t* prob which varies enough. Significant *t* prob (less than 0.05) is submitting at explanatory variable *PDRB_d, RPIND_d, URBAN_d* and *CONT* PDRB *d*.

Explanatory variable such as *PDRB_o,* although it is not significant, but it is still put in Model 2, since this variable is a pair of *PDRB_d* variable. Similar case for *PDRB_o*. For other variables that are insignificant, they will be thrown away and will be applied alternatively for Model 2 as follow:

$$FLOW = \beta_0 + \beta_1 * PDRB\_o + \beta_2 * PDRB\_d + \beta_3 * RPIND\_o + \beta_4 * RPIND\_d + \beta_5 * URBAN\_o + \beta_6 * URBAN\_d + \beta_7 * CONT$$

By using Model 2 above, the test result whether the model bad or good is not changeable, that is with R^2 value is small (1,8 percent) but it has a great meaning (prob.F=0,0000). Meanwhile, significance test for each explanatory variable is relatively better, almost all explanatory variables have significant *t* prob. Explanatory variables such as *RPIND_d* and *URBAN_d,* althouggh they are not significant, but still put in the model. Theoretically, those two explanatory variables should be put as pair of *RPIND_d* and *URBAN_d*.

As an estimator, model submitted should fulfill some basic assumption and has BLUE (Best Linier Unbiased Estimator) characteristic. The assumption that should be fulfill (Pindyck and Rubinfeld, 1991) are specification of the model have been determined as multiliniear regression model; explanatory variable is nonstochastic and has no relation to precised liniear between two or more explanatory variables; error has expected value of 0; error of different observations is independent, so that they do not correlated; and variable error is normal distribution.

From the above assumption, model can be tested by multicoliniearity test (correlation between explanatory variables), auto-correlation test (on in dependent error) and heteroschedisity test (on constant variable error).

Trespassing on the assumption causes the interpretation of parameter coefficients is difficult, although we still are able to obtain the value of smallest quadratic estimator. Regression coefficient of first and second variables which correlation is high is estimated to measure the change of dependent variable (*Y*) caused there is a variable in the function, by estimating that other modifiers are constant. However, the existence of multicoliniearity causes change of each variable, then the observation of pair of variable which has high correlation will be strongly changed. The impact is we will be difficult to separate the effect of one independent variable on other dependent variables.

To detect the above problem, partial correlation matrix or the values of *VIF* (Variance Inflating Factor), collinearity statistics process is use. Guideline commonly used is that the correlation value between variables is not bigger than 0.8 or 0.9 that indicates a strong linear correlation is occurred. Other guidelines is, if *VIF* value is bigger than 10, it indicates that there is a multicollinearity occurred.

Attachment 2 reveals partial matrix correlation, the value of correlation between explanatory variables. From matrix we can see that there is no value bigger than 0.8 or 0.9. If there were relatively high great value (between PDRB_o and RPIND_o), but it is still lower than 0.8, or 0.7. It could be assumed that this model has no problem with multicolinearilty. This assumption is also supported by *VIF* value between 1 to 3, so that there is no *VIF* value that is higher than 10.

The other basic assumption of regression model using the smallest quadratic metode present here is that there is no correlation between errors. Problem auto-correlation most likely occurs on the study of time series that shows error relates to observation on certain period to the following period. Auto-correlation does not influence instability or consistency of smallest quadratic assumption, but influences efficiency model. Variant estimation parameter that does not tend to have test *t* value of each explanatory variable receives hypothesis null (H_0).

Method usually used to detect auto-correlation occurrence is Durbin-Watson test. This test is done by comparing the upper value (DW_u) and (DW_l) of Durbin-Watson table and the number of observation and explanatory variables plus one. Meanwhile, confidence obtained from the test results includes five regions, they are: (1) less than DW_l; (2) between DW_l and DW_u; (3) between DW and 4 - DW_u; (4) between 4 - DW_l and 4 – DW_l; and (5) more than 4 - DW_l.

If *DW* calculation is located between 1 and 5 then the model shows auto-correlation problem. If *DW* calculation is located in interval 3, or closes to value 2, then model has no experience serious problem of auto-correlation. If *DW* is located between interval 2 and 4 then the test result can not be assumed whether there is auto-correlation problem or not.

Attachment 1 shows *DW* value of Model 1 or Model 2 that are relatively similar, that is 1.9; it means, we receive hypothesis null that indicates there is no serial correlation (auto-correlation).

If error has constant variance, it is called homoscedastic, if error changes that causes variant also changes, then it is heteroscedasity. Heteroscedasity often occurs in cross-section economic data, because it has more variety of individuals than individual in time series data.

To observe whether heteroscedasity occurs or not, it can be done by the following methods (Newbold, Carlson dan Thorne, 2003): If regression model such as

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_K x_{Ki} + \varepsilon_i ,$$

then model for predicted *y* value is:

$$\hat{y}_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \ldots + b_K x_{Ki} ,$$

and the residual of the model is $e_i = y_i - \hat{y}_i$. Then, we estimate by simple regression where dependent variable is quadratic of residual and explanatory variables is predicted *y* value: $e_i^2 = a_0 + a_1 \hat{y}_i$. The value of R^2 (coefficient determination) obtained times the number of observation (nR^2) and compared to $\chi^2_{1,\alpha}$ (critical value of chi-quadratic with degree of freedom 1 and error probability α) at the table. If nR^2 is bigger than chi-quadratic table, then we refuse hypothesis null which states that regression model has constant error variance.

Attachment 3 shows the calculation of the above method, and it results determination coefficient 0.008. When it is multiplied by the number of observation (3,480), the result is 27.84. Meanwhile, the value of chi-quadratic in the table with degree of freedom 1 and alpha 10% is 2.71. It means, we refuse hypothesis null which states that the model has constant error variant. In other word, model has high heteroscedasity. It is possible occurred since the model used cross-section data, moreover with great number of observation (3,480) and be added by few number of explanatory variables.

The evidence of heteroscedasity can be seen also from scatter-plot between predicted value and its residual. At Attachment 3 shows that its pattern is not random and tends to form a straight line. Such kind of pattern shows that heteroscedasity is occurred.

INTERPRETATION OF THE RESULTS

From the above submission, it obviously shows that this model experienced deviation of assumption in homoscedasity, although in multicolinearity and autocorrelation test have no serious problem. As already mentioned before, data used in this study is cross-section with great number of observations and relatively small explanatory variables. Therefore, this model has very small determination coefficient (1.8 %) and experiences serious heteroscedasity.

Apart from problem on econometrics above, coefficient value of each explanatory variables has substantial meaning, as follows: the positive regression coefficient value of PDRB_o and the negative regression coefficient value of PDRB_d show that migration has positive relationship to PDRB of place of origin and has negative relationship to PDRB of place of destination. In other word, the higher PDRB of an area, the greater out-migration flow from that area. Meanwhile, high PDRB of an area makes also low number of migrants coming to that area. This result of regression reflects the contrary statement of a theory saying that population migration tends to flow from regions having low output to regions having great output.

Negative coefficient of *URBAN_o* and positive coefficient of *URBAN_d* show that migration flow has negative correlation to the number of urbanization in areas of origin and has positive correlation to the number of urbanization in areas of destination. In other word, the higher the urbanization number, the lower the migrants going out from that place, and contrary, the higher the number of migrant coming to that place. It is accordance to the theory saying that migrants tends coming to the areas where the urbanization is high.

Explanatory variable *RPIND_o* has positive regression as *RPIND_d* does. It means that migration flows either from or to the high industrial areas. Industrialization of an area tends causing strong interaction to the other areas, especially that having also high industrialization. It is accord with inter-spacial interaction theory that notes migration flows from area of origin to area of destination is straight with the strength of those two areas. Their strength, in this case, is proxied by their industrialization share.

Positive coefficient value of *CONT* variable gives a figure to us that migration flows tend occurred in areas having direct borders with areas surroundings. It is accord with the classical migration theory that notes that migration tends occurred in close distance areas. And, it also accords with the spatial interaction theory that notes that migration intention from area of origin to area of destination is upside down with the distance of those two areas.

CONCLUSION

Model developed in this study gives conclusion to us that high PDRB in an area tends to increase migrants coming out from that area and reduce migrants coming to that area. Meanwhile, high migration rate in an area tends to 'hold' people going out from that area and to 'pull' people coming into that area. Whereas, areas having high industrialization tends to strengthen interaction between them. Migration also tends to flow to areas as having direct borders compared to other areas located farther from areas of origin.

The above analysis shows that this model developed still has high heteroscedasticity so that for an estimation it needs alertness to learn and to interpret coefficient value resulted, although it has significant error probability. Besides, there is need a modification of the model, so that the analysis of the model has not use OLS method but it can be done through logit or probit model by modifying dependent variable number of migrants to probability of migration of an area.

Bibliography

Ananta, Aris dan Turro S. Wongkaren. 1996. In Aris Ananta and Chotib. *Mobilitas Penduduk di Indonesia*. Lembaga Demografi FE-UI & BBKBN, Jakarta

Center Bureau of Statistics. 1996. *Penduduk Indonesia, Hasil Survei Penduduk Antar Sensus 1995*. Serie S2. Jakarta.

------. 1999. *Produk Domestik Regional Bruto Propinsi-propinsi di Indonesia menurut Penggunaannya 1995-1998*. Jakarta.

------. 2001. *Penduduk Indonesia: Hasil Sensus Penduduk Tahun 2000* . Serie L2.2. Jakarta.

Chotib. 2001a. "Dibalik Angka Pengangguran di Tengah Krisis". *KOMPAS*. 13 Mai.

------. 2001b. "Inter-provincial Migration in Indonesia 1990-1995: Application of Spatial Interaction Model Using the 1995 Intercensal Population Survey Data". *Journal of Population*. Vol. 7 no. 1:41-62.

Field, Gary S. 1982. "Place to Place Migration in Columbia". *Economic Development and Cultural Change*. Vol. 30. No. 3.

Firman,Tommy.1994. "Migrasi Antar propinsi dan Pengembangan Wilayah di Indonesia". *Prisma*. No. 7. p. 3-15.

Koestoer, R. H. 1991. "Accessibility to Employment in BOTABEK: A Modelling Approach". *Majalah Demografi Indonesia*. No. 36.

Lee, Everett S. 1977. *Suatu Teori Migrasi* (Migration Theory). Translated. Yogyakarta: Center of Population studies, University of Gajah Mada.

McConnel, Campbell R and Stanley L. Brue. 1989. *Contemporary Labor Economics*. New York: McGraw Hill Book Company.

Newbold, Paul, William L. Carlson, and Betty Thorne. 2003. *Statistics for Business and Economics*. Fifth Edition. New Jersey: Pearson Education, Inc.

Pindyck Robert S. and Daniel L. Rubinfeld. 1991. *Econometric Models and Economic Forecasts*. New York: McGraw Hill, Inc.

Sapta Utama, Bey. 1994. *Migrasi dan Pembangunan Daerah di Indonesia: Aplikasi Model). Thesis,* Faculty of Economic University of Indonesia.

Tjiptoherijanto, Prijono. 2000. *Mobilitas Penduduk dan Pembangunan Ekonomi.* Paper presented in Two-day Simposium, State Ministry of Transmigration and Population. Jakarta 25-26 May.

Todaro, Michael P. 1976. *Internal Migration in Developing Countries*. Geneva: International Labour Office.

Yunus, Nizam. 1985. *Beberapa Faktor yang Mempengaruhi Migrasi antar Propinsi di Indonesia* (Several Factors Influencing Interprovincial Migration in Indonesia). Thesis, Faculty of Economic University of Indonesia.

Yosephine, Susanne. 1989. *Faktor-faktor Penentu Migrasi Masuk dan Migrasi Keluar antar Propinsi di Indonesia*. Thesis, Faculty of Economic University of Indonesia.

Attachment 1:

Model 1

Dependent Variable: FLOW

Method: Least Squares

Date: 01/22/04 Time: 19:35

Sample(adjusted): 1 3480

Included observations: 3480 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-109195.9	67877.41	-1.608723	0.1078
PDRB_O	7.33E-05	5.76E-05	1.271852	0.2035
PDRB_D	-0.000228	5.71E-05	-3.985295	0.0001
RPIND_O	50.91674	107.4211	0.473992	0.6355
RPIND_D	734.0405	105.7405	6.941907	0.0000
EMPLOY_O	385.3783	491.7579	0.783675	0.4333
EMPLOY_D	649.6411	491.7153	1.321173	0.1865
URBAN_O	-25.26238	57.53313	-0.439093	0.6606
URBAN_D	129.2822	57.47466	2.249377	0.0246
SEX	-1492.616	1326.965	-1.124834	0.2607
URRU	247.6781	1326.965	0.186650	0.8519
CONT	5775.496	2098.217	2.752573	0.0059

R-squared	0.019242	Mean dependent var	2369.078
Adjusted R-squared	0.016131	S.D. dependent var	39459.32
S.E. of regression	39139.76	Akaike info criterion	23.99111
Sum squared resid	5.31E+12	Schwarz criterion	24.01233
Log likelihood	-41732.53	F-statistic	6.185463
Durbin-Watson stat	1.979229	Prob(F-statistic)	0.000000

Model 2:

Dependent Variable: FLOW

Method: Least Squares

Date: 01/23/04 Time: 10:41

Sample(adjusted): 2 3480

Included observations: 3479 after adjusting endpoints

Convergence achieved after 4 iterations

Variable	Coeffi-	Std. Error	t-Statistic	Prob.

	cient			
C	-10884.26	3044.588	-3.574954	0.0004
PDRB_O	7.53E-05	5.82E-05	1.293213	0.1960
PDRB_D	-0.000225	5.71E-05	-3.946102	0.0001
URBAN_O	-38.67810	55.59120	-0.695759	0.4866
URBAN_D	105.3570	54.88971	1.919430	0.0550
RPIND_O	36.82804	106.9257	0.344427	0.7305
RPIND_D	714.3011	104.1882	6.855874	0.0000
CONT	5342.470	2097.009	2.547662	0.0109
R-squared	0.018322	Mean dependent var		2366.218
Adjusted R-squared	0.016059	S.D. dependent var		39464.63
S.E. of regression	39146.46	Akaike info criterion		23.99059
Sum squared resid	5.32E+12	Schwarz criterion		24.00651
Log likelihood	-41722.63	F-statistic		8.095699
Durbin-Watson stat	1.999933	Prob(F-statistic)		0.000000

Attachment 2:

Partial Correlation Matrix (Model1)

Coefficient Correlations[a]

Model			CONT	URRU	SEX	INDS_D	INDS_O	EMPLOY_O	EMPLOY_D	URBAN_O	URBAN_D	PDRB_D	PDRB_O
1	Correlations	CONT	1.000	.000	.000	-.028	-.013	.083	.087	.031	.024	.028	.016
		URRU	.000	1.000	.000	.002	.000	.000	.000	.000	.001	-.002	.000
		SEX	.000	.000	1.000	.002	.000	.000	.000	.000	.001	-.002	.000
		INDS_D	-.028	.002	.002	1.000	.037	-.001	.063	.022	.565	-.974	-.036
		INDS_O	-.013	.000	.000	.037	1.000	.064	.001	.567	.020	-.036	-.974
		EMPLOY_O	.083	.000	.000	-.001	.064	1.000	.041	.261	.010	.002	-.042
		EMPLOY_D	.087	.000	.000	.063	.001	.041	1.000	.012	.261	-.041	.000
		URBAN_O	.031	.000	.000	.022	.567	.261	.012	1.000	.037	-.025	-.662
		URBAN_D	.024	.001	.001	.565	.020	.010	.261	.037	1.000	-.660	-.024
		PDRB_D	.028	-.002	-.002	-.974	-.036	.002	-.041	-.025	-.660	1.000	.037
		PDRB_O	.016	.000	.000	-.036	-.974	-.042	.000	-.662	-.024	.037	1.000
	Covariances	CONT	4449887	-157.544	-157.544	-.032	-.015	85342.536	89828.572	4540.721	3576.609	1.118E-02	6.579E-03
		URRU	-157.544	1778864	7.299	1.399E-03	5.149E-05	-2.556	53.967	4.679	123.503	-.001	.000
		SEX	-157.544	7.299	1778864	1.399E-03	5.149E-05	-2.556	53.967	4.679	123.503	-.001	.000
		INDS_D	-.032	1.399E-03	1.399E-03	2.825E-07	1.037E-08	-3.175E-04	1.624E-02	7.958E-04	2.079E-02	.000	.000
		INDS_O	-.015	5.149E-05	5.149E-05	1.037E-08	2.839E-07	1.656E-02	3.096E-04	2.097E-02	7.530E-04	.000	.000
		EMPLOY_O	85342.536	-2.556	-2.556	.000	1.656E-02	237727.664	9852.266	8836.132	344.872	1.813E-04	-.004
		EMPLOY_D	89828.572	53.967	53.967	1.624E-02	3.096E-04	9852.266	237910.334	397.974	8824.204	-.004	.000
		URBAN_O	4540.721	4.679	4.679	7.958E-04	2.097E-02	8836.132	397.974	4817.498	178.768	.000	-.009
		URBAN_D	3576.609	123.503	123.503	2.079E-02	7.530E-04	344.872	8824.204	178.768	4797.856	-.009	.000
		PDRB_D	1.118E-02	-.001	-.001	.000	.000	1.813E-04	-3.828E-03	-3.319E-04	-8.761E-03	3.673E-08	1.359E-09
		PDRB_O	6.579E-03	.000	.000	.000	.000	-3.910E-03	-1.543E-05	-8.826E-03	-3.154E-04	1.359E-09	3.692E-08

a. Dependent Variable: FLOW

VIF Values in Model 1:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	-46173.1	67017.707		-.689	.491					
	PDRB_O	3.296E-04	.000	.169	1.715	.086	.032	.029	.029	.029	34.016
	PDRB_D	-.001	.000	-.296	-3.008	.003	.033	-.051	-.051	.030	33.836
	INDS_O	-.001	.001	-.113	-1.270	.204	.029	-.022	-.021	.036	27.554
	INDS_D	1.737E-03	.001	.289	3.268	.001	.041	.055	.055	.036	27.426
	URBAN_O	-80.471	69.408	-.032	-1.159	.246	.009	-.020	-.020	.368	2.717
	URBAN_D	199.640	69.267	.080	2.882	.004	.037	.049	.049	.370	2.706
	EMPLOY_O	308.417	487.573	.012	.633	.527	.001	.011	.011	.820	1.219
	EMPLOY_D	161.004	487.761	.006	.330	.741	-.018	.006	.006	.820	1.220
	SEX	-1503.066	1333.741	-.019	-1.127	.260	-.019	-.019	-.019	1.000	1.000
	URRU	237.228	1333.741	.003	.178	.859	.003	.003	.003	1.000	1.000
	CONT	5815.260	2109.475	.047	2.757	.006	.045	.047	.047	.982	1.018

a. Dependent Variable: FLOW

Partial Correlation Matrix (Model 2):

	PDRB_O	**PDRB_D**	**RPIND_O**	**RPIND_D**	**URBAN_O**	**URBAN_D**	**CONT**
PDRB_O	1.000000	-0.034250	0.733191	-0.026540	0.642342	-0.022150	-0.031131
PDRB_D	-0.034250	1.000000	-0.025097	0.727448	-0.021622	0.642112	-0.009645
RPIND_O	0.733191	-0.025097	1.000000	-0.034994	0.415940	-0.014343	-0.000813
RPIND_D	-0.026540	0.727448	-0.034994	1.000000	-0.017267	0.414510	0.013417

URBAN_O	0.642342	-0.021622	0.415940	-0.017267	1.000000	-0.034483	-0.038170
URBAN_D	-0.022150	0.642112	-0.014343	0.414510	-0.034483	1.000000	-0.025203
CONT	-0.031131	-0.009645	-0.000813	0.013417	-0.038170	-0.025203	1.000000

VIF Values in Model 2:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	-10905.1	3025.350		-3.605	.000					
	PDRB_O	7.542E-05	.000	.039	1.310	.190	.032	.022	.022	.324	3.085
	PDRB_D	-2.24E-04	.000	-.115	-3.930	.000	.033	-.067	-.066	.330	3.028
	RPIND_O	36.518	105.743	.009	.345	.730	.027	.006	.006	.456	2.191
	RPIND_D	710.414	104.145	.168	6.821	.000	.102	.115	.115	.465	2.150
	URBAN_O	-38.635	55.009	-.016	-.702	.483	.009	-.012	-.012	.580	1.725
	URBAN_D	107.072	54.965	.043	1.948	.051	.037	.033	.033	.581	1.722
	CONT	5420.993	2084.240	.044	2.601	.009	.045	.044	.044	.996	1.004

a. Dependent Variable: FLOW

Attachment 3:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.087[a]	.008	.007	52.1751

a. Predictors: (Constant), Standardized Predicted Value

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	72104.313	1	72104.313	26.487	.000[a]
	Residual	9467961	3478	2722.243		
	Total	9540065	3479			

a. Predictors: (Constant), Standardized Predicted Value

b. Dependent Variable: ZESID2

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.998	.884		1.128	.259
	Standardized Predicted Value	4.553	.885	.087	5.147	.000

a. Dependent Variable: ZESID2

Scatterplot

Dependent Variable: FLOW

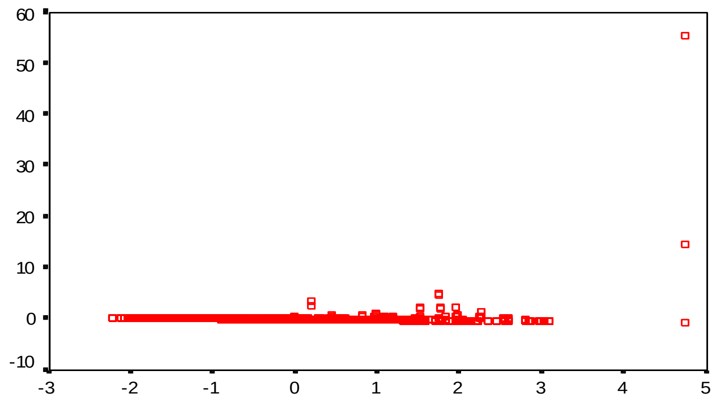

TABLE SUMMARIES

VARIABLES	MODEL 1		MODEL 2	
	Coefficient	t-Prob	Coefficient	t-Prob
C	-109195.9	0.1078	-10884.26	0.0004
PDRB_O	7.33E-05	0.2035	7.53E-05	0.1960
PDRB_D	-0.000228	0.0001	-0.000225	0.0001
RPIND_O	50.91674	0.6355	36.82804	0.7305
RPIND_D	734.0405	0.0000	714.3011	0.0000
EMPLOY_O	385.3783	0.4333	-	-
EMPLOY_D	649.6411	0.1865	-	-
URBAN_O	-25.26238	0.6606	-38.67810	0.4866
URBAN_D	129.2822	0.0246	105.3570	0.0550
SEX	-1492.616	0.2607	-	-
URRU	247.6781	0.8519	-	-
CONT	5775.496	0.0059	5342.470	0.0109
R-squared	0.019242		0.018322	
Adjusted R-squared	0.016131		0.016059	

F-statistic	6.185463 (0.000000)	8.095699 (0.000000)
Durbin-Watson stat	1.979229	1.999933
Multicolinearity	PDRB_O, PDRB_D; RPIND_O, RPIND_D	No Multicolinearity
Autocorrelation	No autocorrelation	No autocorrelation
Heteroscedasticity	High heteroscedastic	High heteroscedastic