The Social Sciences 12 (9): 1695-1704, 2017
ISSN: 1818-5800

Geographical and Socio-Economic Analysis in Peninsular Malaysia

[1]Soo-Fen Fam, [2]Noriszura Ismail and [2]Abdul Aziz Jemain
[1]ExCITED, C-TeD, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[2]Universiti Kebangsaan Malaysia, Selangor, Malaysia

Abstract: This study aims to develop a pioneer composite area-based index of socioeconomic deprivation, namely the General Index of Deprivation (GID) using Principal Component Analysis (PCA). The proposed GID which is based on combined resources of census data, administrative registration data, vital statistics and insurance data of eighty-one Administrative Districts (ADs) in Peninsular Malaysia in 2000 can be used to provide a greater understanding and interpretation of the distribution of socio-economic patterns across the ADs. This study also measures the overall and the local clustering in socio-economic deprivation across ADs in Peninsular Malaysia using Global and Local Moran's I. Further, visualizations of the patterns of socio-economic deprivation based on the proposed GID and the locations of spatial clusters based on the Local Moran's I are implemented by using choroplate maps. The results of GID indicate that the ADs can be ranked and classified into four quartiles; the most affluent, the moderately affluent, the moderately deprived and the most deprived. The majority of affluent areas were located in the westcoast of Peninsular Malaysia whereas the most deprived areas were mainly scattered in the Northeast of Peninsular Malaysia. The results of Global Moran's I suggests significant and positive global spatial autocorrelation across spatial weights of the ADs. In addition, the results on Local Moran's I show strong spatial disparities of socioeconomic deprivation in several ADs, indicating the importance of considering geographic localization and spatial condition of each AD for allocating resources and implementing efficient policies in Peninsular Malaysia.

Key words: Deprivation index, Moran's I, principal component analysis, spatial autocorrelation, spatial

INTRODUCTION

The concept of deprivation has its origin in Britain in the late 1960s where it emerged from issues of economic and social inequality (Norris, 1979). In particular, deprivation can be defined as a state of observable and demonstrable disadvantage relative to the local community, the wider society or the nation to which an individual, family or group belongs (Townsend, 1987). Deprivation is also a "status of material and social harm" which affects a person, a family or a society and can be observed by comparing the characteristics of those "harmed" with the characteristics of society as a whole (Abu and Abu, 2009). In terms of classification, deprivation can be divided into two general forms; material and social. Material deprivation is a relative lack of goods, resources or services that are widely available in a society whereas social deprivation refers to isolation or exclusion due to inclusion in a certain class, race, gender or other social division (Townsend, 1993; Dominguez *et al*., 2001).

Since the 1980s, several indices of area-based deprivation for the measurement of economic or social disadvantages of urban areas have been proposed (Townsend, 1987) where such indices were developed from the proportion of households in a defined small geographical unit with a combination of circumstances indicating low living standards or high need for services or both (Bartley and Blane, 1994). The main advantage of using area-based deprivation index is that the index is often used to measure socioeconomic conditions of specific geographical areas rather than individuals, due to the absence of individual level data and hence, it is indicative of the whole society that live in the same area (Car, 1988). In addition, many researchers favored area-based deprivation indices because they are relatively simple and inexpensive and are generally composed of easily available census variables (Bartley and Blane, 1994; Testi and Ivaldi, 2009).

Numerous studies have been carried out regarding deprivation analysis issues and indices, mostly in developed countries such as Europe, United States and Canada. However, only a few studies were found in developing countries either low or middle income nations (McIntyre *et al*., 2002; Yuan and Xu, 2011) and Malaysia which is a developing country located in Southeast Asia with an estimated population of 28.3 million in 2010 is of no exception. In Malaysia, several socio-economic

Corresponding Author: Soo-Fen Fam, ExCITED, C-TeD, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

indicators, either single indicator or composite indices, have been developed and applied to depict socio-economic patterns across geographical areas. The two composite indices which are currently receiving significant interest from government agencies and policymakers are the Malaysian Urban Indicators network (MURNInet), proposed by the Department of Town and Country Planning for Gauging town sustainability and the Human Development Index (HDI), used by the UNDP for measuring key dimensions of human capabilities. However, these indices do not measure socio-economic status in terms of small geographical area units and to the best of our knowledge, studies on index of socio-economic deprivation depicting the socioeconomic status across small geographical areas in Peninsular Malaysia have not been carried out. Extensive research regarding the development of index of socio-economic deprivation is of utmost important to government agencies and policymakers in Malaysia, especially in identifying priority areas and targeting programmes to improve socio-economic conditions and inequalities.

Several variables or indicators for deprivation have been proposed in literatures and the commonly used indicators are income level (McIntyre *et al*., 2002; Pampalon *et al*., 2009; Noble *et al*., 2010), education level, social class (Carstairs and Morris, 1989), unemployment rate (Pampalon *et al*., 2009; Noble *et al*., 2010), percentage of families owning homes (Testi and Ivaldi, 2009; Choi *et al*., 2011) and percentage of families owning vehicles (McIntyre *et al*., 2002; Havard *et al*., 2008; Bellani and D'Ambrosio, 2011; Choi *et al*., 2011). Even though a practical and simple measure of socio-economic level can be obtained by considering a single variable indicator (McIntyre *et al*., 2002; Sanchez *et al*., 2008), the indicator has limited usefulness when it comes to measuring a concept as complex as deprivation (Sanchez *et al*., 2008). Therefore, the development of a general index of socio-economic deprivation which is based on a combination of simple indicators of both material and social deprivations is therefore crucial in depicting the socio-economic status of a small geographical area relative to other areas in Peninsular Malaysia. However, one of the major difficulties in constructing a composite index is the determination of an appropriate mean to aggregate multidimensional variables into a composite index and this obstacle should be overcome in the process of developing such index.

Several techniques can be implemented to construct composite indices of area-based deprivation and the two common methods are simple additive method and Principal Component Analysis (PCA). Simple additive method was largely used in the 1980s by researchers from Europe such as (Jarman, 1983; Townsend, 1987; Carstairs and Morris, 1989) who developed composite indices of material deprivation. Since the 1990s however, several studies have reverted to the use of PCA to construct area-based indices and such examples can be found in the deprivation index of New Zealand namely 91, 96, 2001 and Nzdep2006, constructed for three principal purposes; resource allocation, research and advocacy (Crampton *et al*., 1997; Salmond *et al*., 1998; Salmond and Crampton, 2002; Salmond *et al*., 2007). In another example, PCA was applied to develop material and social deprivation indices in Canada to facilitate the monitoring of social inequalities in health planning (Pampalon *et al*., 2009). In France, Havard *et al*. (2008) proposed a small area index of deprivation to detect social disparities in the distribution of health outcomes at small-area levels and Rey created an ecological deprivation index and evaluated its association with mortality. Another example can be found in United States where an area-based socioeconomic index was constructed to examine patterns of socioeconomic areas in all-cancer mortality among the US men (Singh *et al*., 2002).

The objective of this study is to develop a pioneer composite area-based index of socioeconomic deprivation, namely the General Index of Deprivation (GID) which depicts the socioeconomic status across small geographical areas in Peninsular Malaysia using PCA method. In addition, this study provides a greater understanding and interpretation of the distribution of socioeconomic patterns Across Administrative Districts (ADs) in Peninsular Malaysia by relating the ADs to the newly developed GID. And finally, this study analyze spatial characteristics in socio-economic deprivation across ADs in Peninsular Malaysia, in terms of overall and local clustering, using global and local Moran's I.

Even though the Townsend's deprivation index has been derived and discussed in our previous paper (Fam *et al*., 2011a, b), the selected indicators were related to material deprivation and do not cover the multidimensional aspects of deprivation. In addition, the data was solely based on decennial census and may provide outdated information. Furthermore, the indicators were assigned equal weight regardless of their relative importance to the index. The four indicators of the Townsend's deprivation index in Peninsular Malaysia were unemployment rate (proxy for social position deprivation), percentage of household not owning a car (proxy for income deprivation), percentage of household not owning a house (proxy for poor living condition) and percentage of household not owning piped water (proxy for scarcity of basic human needs and housing amenities conditions).

In this study, PCA which is a multivariate technique employed in a composite indexing (Felipe and Resende, 1996; Kallmann, 1997) is applied to develop GID in Peninsular Malaysia. The main advantage of using PCA is that it does not involve subjective opinions for assigning weights to its variables, unlike some methods which consider policymaker's or expert's opinion for the assignment of weights. Another advantage is that the method is easy to implement since the data is allowed to determine the weights by itself where the optimal weight accounts for the largest proportion of variance (Ram, 1982). In particular, PCA does not impose requirements for normality and homoscedasticity which is rarely encountered in social science data and hence, transformation of data is not required. As examples, in their study of constructing socioeconomic deprivation index in France, Havard *et al*. (2008) pointed out that the newly developed index which is obtained via PCA is simpler to develop than the Index of Multiple Deprivation (IMD) and at the same time, overcomes the limitations of Carstair's and Townsend's indices. In addition, PCA is computationally easier compared to other statistical methods and can be implemented to data which are easily collected from household surveys (Jobson, 1992). Another main advantage of applying PCA is that it allows new variables to be added with minimal cost (Ram, 1982). However, to ensure the accuracy of index, component variables should be additive and differential weighting should be assigned to the variables (Gordon, 1995).

MATERIALS AND METHODS

Data: The data for developing GID is provided by Department of Statistics Malaysia which is the national statistics office responsible for collecting and disseminating the nation's statistical data and Insurance Services Malaysia (ISM) which is the national office responsible for collecting insurance data. Specifically, the data comprises of several resources including census data, administrative registration data, vital statistics, insurance data and safety data of eighty-one Administrative Districts (ADs) in Peninsular Malaysia in 2000. Abiding by the Statistics Act 1965 and Census Act 1960, the census data was collected in terms of zone level.

Geography, Homogeneity and Spatial scale effect: In Peninsular Malaysia, the smallest geographical unit in census data is enumeration block, each consisting of 80 to 120 living quarters. Amalgamation of a few enumeration blocks forms one zone, averaging about 500 living quarters in each zone. The nesting process continues where nested zones become a canton, nested cantons form an administrative district, nested administrative districts turn into state and lastly, nested states form the Peninsular Malaysia. Further, illustration on the nesting of blocks, zones, cantons, administrative districts and states in Peninsular Malaysia is provided by Fam *et al*. (2011a, b).

In an ecological research, the homogeneity of socioeconomic characteristics is essential, indicating that the selection of a basic geographical unit should be as small as possible (Pampalon *et al*. 2009). Even though a zone which is a small geographical area with consistent number of living quarters is a suitable candidate as a basic geographical unit in Peninsular Malaysia, its inconsistent boundary does not allow the execution of spatial temporal analysis. Therefore, an administrative district which has a more consistent boundary is chosen as the basic geographical unit in this study. Furthermore, all resources of data such as census data, administrative registration data, vital statistics and insurance data are available in terms of district level, allowing the data from different resources to be matched and combined for further geographical and socioeconomic analysis. It should be noted that East Malaysia is excluded from this study due to varied culture and geographical differences.

Variables: The objective of this study is to develop the composite area-based index of socioeconomic deprivation, namely the General Index of Deprivation (GID) which depicts the socioeconomic status across areas in Peninsular Malaysia. In short, the proposed GID is an index derived from a selected series of domains of socioeconomic deprivation of a small area level by combining the optimal weights, obtained using PCA. Based on the literature review of deprivations and the availability of data for each AD in Peninsular Malaysia, twenty-three variables are initially considered for the procedure of deriving the optimal weights of GID. The twenty-three variables also fulfill the following criterions:

- Relevance to the concept of deprivation in the Malaysian context
- Compatible and applicable over time

The twenty three variables and their domains are shown in Table 1. In particular, seven domains are considered; family, economic and material domain, basic utilities domain, employment domain, education domain, health domain, health facilities domain and well-being domain. Therefore, the proposed GID is expected to explain multiple dimensions of socioeconomic status of all ADs in Peninsular Malaysia.

Construction of GID: PCA, a multivariate technique originally introduced by Pearson (1901) and independently developed by Hotelling (1933), linearly transforms a large set of correlated variables into a substantially smaller set of uncorrelated variables called principal components which represents a large proportion of information of the original data set. The basic idea behind PCA is that a smaller set of uncorrelated variables is much easier to be understood and used for further analyses compared to a larger set of correlated variables. Even though the maximum number of principal components extracted is equal to the dimensionality of variables of the original data set in general, only a few principal components account for most variation of the original set of variables. In particular, the first principal component accounts for the largest proportion of variation; the second principal component is uncorrelated with the first component and accounts for the largest proportion of variation not accounted by the first principal component and so on.

PCA is a statistical method which uses a covariance matrix or a correlation matrix to extract components. In many cases, different variables often have different units and scales and therefore, the correlation matrix is preferred by researchers as it overcomes this problem by scaling each covariance with its associated standard deviations.

Since, PCA works optimally when variables are correlated (Vyas and Kumaranayake, 2006), Pearson correlations are initially calculated to examine correlations between the twenty three variables. In particular, highly correlated variables or variables with correlation coefficient significant at one percent level are retained as suggested by Alderman and Morris (1967). In addition, if the variables showed extremely high correlations (>0.95), variables with relatively lower correlation are dropped to avoid multicollinearity. After retaining highly correlative variables and checking for multicollinearity, the data table with eighty one Ads and p variables, $p \leq 23$ is then formed. From the data table, a single component which represents the relationship between socio-economic deprivation and selected variables and at the same time explains maximal variability in the data without losing too much information is obtained. The ith principal component, y_1, i = 1, 2, 3, ..., p which is independent of all other principal components and can be calculated using linear combination of p variables can be equated as:

$$y_i = e_{i1}z_1 + e_{i2}z_2 + ... + e_{ip}z_p$$

Where:

e_{ij}, j = 1, 2, 3, ..., p = The coefficient derived from PCA
Z_j = The standardized variable

In our study, the first principal component, i = 1 is selected for GID as it explained the largest proportion of data variation. The ADs are then ranked and classified into four quartiles (Q1-Q4) according to their respective GID where each quartile has an approximately equal number of ADs (around twenty ADs per quartile). In particular, Quartile 1 represents the most affluent area, Quartile 2 the moderately affluent area, Quartile 3 the moderately deprived area and Quartile 4 the most deprived area (Table 1).

Exploratory spatial analysis: The Global Moran's I, a measure of overall clustering in the proposed GID can be calculated as:

$$I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(a_i - \bar{a})(a_j - \bar{a})}{\sum_{i=1}^{n}(a_i - \bar{a})^2}$$

Where:

a_i = The observation or an attribute value in district
i and w_{ij} = The spatial weight of the link between district i and district j

The weight, w_{ij} which is assigned to district j if it is a neighbor of district i based on queen contiguity with borders and common vertices sharing can be written as:

$$w_{ij} = \begin{cases} 1 & \text{if districts i and j are connected} \\ 0 & \text{if districts i and j are not connected} \end{cases}$$

Inference for Global Moran's I can be performed by using permutation based on the asymptotic distribution of I. The mean and variance of I statistic are derived via the assumption of randomizing the attribute values to the district's lattice. In particular, values of I larger than the expected value, E(I) = 1/n-1, indicate positive spatial autocorrelation and vice versa.

Even though the Global Moran's I can be used to measure the overall clustering, it does not provide the location of each cluster and the type of spatial correlation of each AD. Therefore, the Local Moran's I is applied to measure local spatial association in the proposed GID, besides identifying the location of each cluster and the type of spatial correlation of each AD. In particular, a significant positive association indicates an AD of high deprivation with neighbors of high deprivation and vice versa whereas a significant negative association indicates an AD of low deprivation with neighbors of high deprivation and vice versa. The Local Moran's I can be measured as:

Table 1: Initial variables considered for GID

Variables	Definitions
Family, economic and material deprivation domain	
X_1	Total dependency ratio (100 population aged 15-64 years)
X_2	Percentage of female headed household
X_3	Percentage of private household without a car
X_4	Percentage of private household do not own a house
X_5	Percentage of private household without a radio
X_6	Percentage of private household without a television
X_7	Percentage of private household without a fixed telephone line
X_7	Percentage of private household without a refrigerator
X_8	Percentage of private household without a washing machine
Basic utilities deprivation domain	
X_9	Percentage of private household without treated piped water
X_{10}	Percentage of private household without a toilet with flush system
X_{11}	Percentage of private household without 24 h of electricity supply a day
X_{12}	Percentage of private household without garbage collection facility
Employment deprivation domain	
X_{13}	Percentage of unemployed population aged 15-64 years who are economically active
Education deprivation domain	
X_{14}	Percentage of population without education
X_{15}	Percentage of population without receiving upper secondary education (<10 years of education)
X_{16}	Percentage of population aged 10 years and above without receiving SPM certificate
Health deprivation domain	
X_{17}	Relative risk of infant mortality
X_{18}	Relative risk of neonatal mortality
X_{19}	Relative risk of stillbirth
Health facilities deprivation domain	
X_{20}	Hospital bed ratio (10,000 population)
Well-being deprivation domain	
X_{21}	Percentage of population aged 5-20 years caught, convicted and placed in corrective institutions including those placed under parental custody (10,000 population)
X_{22}	Percentage of policyholders making claims of private car thefts in 2000-2003 (10,000 population)

$$I = \frac{(a_i - \bar{a})}{\sum_{i=1}^{n}(a_i - \bar{a})^2} \sum_{j=1}^{n} w_{ij}(a_j - \bar{a})$$

Where:

a = Positive

I = Indicates spatial clustering of similar values (either high or low)

whereas a negative I indicates spatial clustering of opposite values between an AD and its neighbors. In the presence of global spatial autocorrelation, inference for Local Moran's I can be performed by using conditional permutation based on the asymptotic distribution of I. In other words, the conditional approach requires the value at district i to be fixed but the remaining values to be randomly permuted over all locations (Anselin, 1995).

RESULTS AND DISCUSSION

The Pearson correlation matrix for the twenty-three variables indicates that only fourteen variables are highly correlated at 1% significant level. In addition, the matrix shows a case of variable's duplication where X_{16} and X_{17} have unusually high correlations (0.996). Therefore, variable X_{17} is retained as it has a relatively high correlation with other variables compared to X_{16}. Finally, thirteen variables are retained and they are respectively renamed as z_1, z_2, z_3, ..., z_{13}. The number of domains reduces from seven to five and the final five domains selected are family, economic and material domain, basic utilities domain, employment domain, education domain and health domain. The Pearson correlation matrix for the thirteen variables in the PCA is shown in Table 2.

Table 3 provides the principal components with their eigenvalues and percentage of variance of the thirteen retained variables. The first principal component accounts for 70.08% of the overall variation whereas the other twelve components explain the remaining 10.12-0.18% indicating that a significant dimensional reduction is achieved if information from the first principal component is used.

The coefficients of the thirteen variables of the first principal component, after standardization are provided in Table 4. It can be seen that all coefficients are positive and similar, indicating that the thirteen variables contribute similar weights to the first principal component. The highest weight is 0.307, contributed by Z_5 (percentage of private household without a fixed telephone line) followed by Z_6 (percentage of private household without a refrigerator) and Z_1 (total dependency ratio). The 2000 GID which is defined as the first principal component, can be calculated using the following equation:

Table 2: Lower triangular correlation matrix for thirteen variables in PCA

Variables	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7	Z_8	Z_9	Z_{10}	Z_{11}	Z_{12}	Z_{13}
Z_1	1.00	-	-	-	-	-	-	-	-	-	-	-	-
Z_2	0.86	1.00	-	-	-	-	-	-	-	-	-	-	-
Z_3	0.68	0.71	1.00	-	-	-	-	-	-	-	-	-	-
Z_4	0.52	0.52	0.84	1.00	-	-	-	-	-	-	-	-	-
Z_5	0.82	0.76	0.85	0.76	1.00	-	-	-	-	-	-	-	-
Z_6	0.73	0.71	0.87	0.91	0.90	1.00	-	-	-	-	-	-	-
Z_7	0.72	0.76	0.75	0.75	0.80	0.89	1.00	-	-	-	-	-	-
Z_8	0.84	0.69	0.63	0.41	0.71	0.62	0.58	1.00	-	-	-	-	-
Z_9	0.83	0.83	0.66	0.57	0.76	0.76	0.80	0.62	1.00	-	-	-	-
Z_{10}	0.69	0.82	0.37	0.31	0.49	0.47	0.58	0.47	0.73	1.00	-	-	-
Z_{11}	0.71	0.55	0.55	0.59	0.75	0.76	0.73	0.63	0.68	0.36	1.00	-	-
Z_{12}	0.85	0.88	0.64	0.40	0.76	0.63	0.61	0.69	0.81	0.67	0.60	1.00	-
Z_{13}	0.67	0.59	0.49	0.59	0.65	0.65	0.60	0.51	0.61	0.50	0.68	0.57	1.00

Table 3: Principal components, eigenvalues and percentages of variance

Principal component	Eigenvalue	Variance (%)
1	9.110	70.080
2	1.316	10.122
3	0.688	5.290
4	0.635	4.883
5	0.377	2.903
6	0.289	2.220
7	0.146	1.122
8	0.143	1.097
9	0.103	0.792
10	0.080	0.617
11	0.053	0.406
12	0.038	0.293
13	0.023	0.177

Table 4: Coefficients of the thirteen variables of the first principal component

Variables	Coefficient	Values
Z_1	Total dependency ratio (100 population aged 15-64 years)	0.303
Z_2	Percentage of private household without a car	0.295
Z_3	Percentage of private household without a radio	0.277
Z_4	Percentage of private household without a television	0.250
Z_5	Percentage of private household without a fixed telephone line	0.307
Z_6	Percentage of private household without a refrigerator	0.303
Z_7	Percentage of private household without a washing machine	0.293
Z_8	Percentage of unemployed population aged 15-64 years who are economically active	0.256
Z_9	Percentage of private household without treated piped water	0.295
Z_{10}	Percentage of private household without a toilet with flush system	0.226
Z_{11}	Percentage of population without education	0.246
Z_{12}	Percentage of population aged 10 years and above without receiving SPM certificate	0.262
Z_{13}	Relative risk of infant mortality	0.278

$$\begin{aligned}\text{2000 GID} = {} & 0.303z_1+0.295z_2+0.277z_3+0.250z_4+\\ & 0.307z_5+0.303z_6+0.293z_7+0.256z_8+\\ & 0.295z_9+0.226z_{10}+0.246z_{11}+0.262z_{12}+\\ & 0.278z_{13}\end{aligned}$$

The 2000 GID is then applied to the dataset of eighty one ADs in Peninsular Malaysia. The ADs are ranked and classified into four quartiles, Quartile 1 ranging from -4.908 to -2.138, Quartile 2 ranging from -2.138 to -0.444, Quartile 3 ranging from -0.444 to -2.037 and Quartile 4 ranging from 2.037 to 7.867, corresponding, respectively to the most affluent, the moderately affluent, the moderately deprived and the most deprived ADs. Further, illustration is showed in Fig. 1 where the lighter shades indicate areas with relatively lower level of deprivation and the darker shades indicate areas with relatively higher level of deprivation.

For testing whether the overall spatial patterns exhibited in 2000 GID show significant spatial structure, the Global Moran's I is calculated and it is equal to 0.632 with $p<2.2e^{-16}$. Therefore, the null hypothesis of randomness or no clustering is rejected and the result suggests significant and positive spatial autocorrelation across all spatial weights of the ADs. Figure 2 show the local spatial association in 2000 GID using the Local Moran's I which can be categorized into three spatial clusters; deprived, affluent and non-significant. The shaded areas represent the ADs with positive spatial autocorrelation between the ADs and their neighbors whereas the non-shaded areas represent the ADs with insignificant spatial autocorrelation. In particular, the deprived cluster is represented by the darker shade whereas the affluent cluster is represented by the lighter shade. The majority of ADs in Kelantan and Terengganu were categorized under the deprived cluster, whilst the affluent cluster mainly consisted of ADs in Selangor and Wilayah Persekutuan.

The measures of local spatial association in 2000 GID using the Local Moran's I statistic indicates the existence of strong spatial disparities of deprivation. In particular, the spatial association of each AD in Peninsular Malaysia can be clustered into three spatial groups, namely the deprived, the affluent and the non-significant where the deprived cluster consists of deprived ADs with deprived neighbors, the affluent cluster contains affluent ADs with affluent neighbors and the non-significant cluster has ADs with insignificant spatial autocorrelations or random spatial patterns. Further, results indicate that twenty four ADs out of eighty one ADs have significant positive

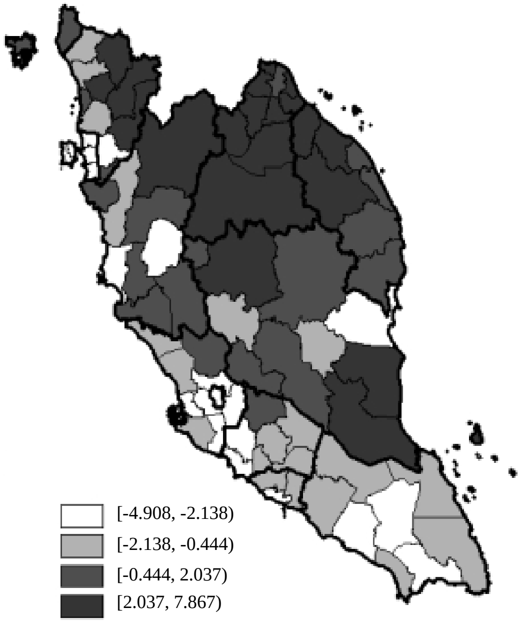

Fig. 1: Quartiles analysis of 2000 GID in Peninsular Malaysia

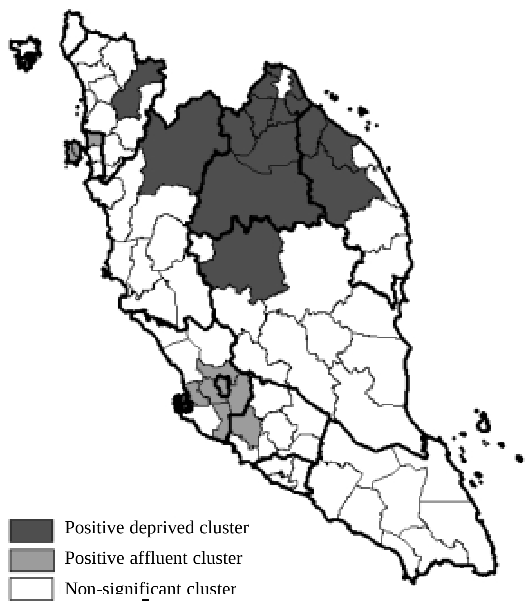

Fig. 2: Local spatial association of 2000 GID in Peninsular Malaysia

spatial autocorrelation at 5% significant level where fourteen ADs were categorized under the deprived cluster and ten ADs fell under the affluent cluster. It is interesting to note that the cluster characterized as "diamonds in the rough" where affluent ADs are surrounded by deprived ADs and the cluster characterized as "doughnuts" where deprived ADs are surrounded by affluent ADs do not exist in Peninsular Malaysia. These characteristics of cluster have been described in a study by Ertur and Koch (2006). By using the variables derived from census data and administrative data in year 2000, a pioneer composite deprivation index for Peninsular Malaysia, GID has been developed at the administrative district level by applying PCA, a technique that is well accepted at international level. The results of PCA indicate that various demographic and socio-economic variables contribute to deprivation in Peninsular Malaysia. In general, the analysis suggests that the percentage of private household without a fixed telephone line is the most important variables contributing to the deprivation in Peninsular Malaysia in census year 2000. Total dependency ratio and other household needs such as refrigerator, car washing machine and radio also contribute significantly to the deprivation in Peninsular Malaysia. Education, relative risk of infant mortality and access to basic utilities such as treated piped water also play an important role. These variables indicate that the citizens of deprived ADs are mainly deprived in terms of communication facilities (fixed telephone line and radio), food preservation facilities (refrigerator), transportation conveniences (car) and quality lifestyles (washing machine).

In 2000, the majority of affluent areas, indicated by the first quartile were located in the west coast of Peninsular Malaysia which also comprised metropolitan and urban areas. The main economic activities in these areas were services and manufacturing and thus, contributing to the overall growth in employment which resulted in higher purchasing power to own material facilities and basic utilities. The ADs under this quartile were Kuala Lumpur which is the capital of Malaysia, Petaling Jaya, Johor Bharu, Melaka Tengah and all ADs in the state of Pulau Pinang. In particular, Timur Laut, an AD in Pulau Pinang was the most affluent AD with the lowest GID.

The most deprived areas or the areas of the fourth quartile were mainly scattered in the Northeast of Peninsular Malaysia where the majority were located in Kelantan and Terengganu which also comprised rural areas. The main economic drivers of these areas were agriculture and fishery. The most deprived AD which was indicated by the highest GID was Gua Musang, a remote AD in Kelantan. In addition, about ninety percents of the ADs in Kedah state were deprived where only one AD in Kedah, namely Yan was not categorized under Quartile 4. Instead, it was categorized under Quartile 3 which was the moderately deprived area. For the state of Terengganu, three out of seven ADs were categorized under Quartile 4 whereas the other four ADs fell into Quartile 3.

The GID proposed in this study can be used as a practical and effective tool for implementing policies, initiating academic researches either for pilot or in-depth studies and allocating resources donated by agencies, companies and non-profit organizations. It should be noted that addressing the deprived populations involves both economic and social conditions and therefore, the newly proposed GID which accounts for multiple domains of deprivation is an appropriate index for depicting such conditions. Mitigation and management of conditions leading to deprivation will support economic growth, ensure productive human resources and expand the market of consumers who are willing and able to spend on goods and services and will eventually, resulted in reduced crime and higher level of quality of life. However, it is important to take into account geographic localization and spatial condition of each AD in the process of allocating resources and implementing efficient policies since there is a strong evidence of global and spatial autocorrelation in socioeconomic deprivation across the ADs. The reason for this is that the expected effects of allocating resources and implementing policies could be over- or under-estimated, depending on the spatial interaction pattern of the ADs and their neighbors.

CONCLUSION

This study develops a pioneer composite area-based index of socioeconomic deprivation for Peninsular Malaysia, called the General Index of Deprivation (GID), using PCA method. The proposed 2000 GID can be used to provide a greater understanding and interpretation of the distribution of socioeconomic patterns across Administrative Districts (ADs) in Peninsular Malaysia. In addition, this study calculates the global and the local Moran's I index to measure overall and local clustering in socioeconomic deprivation across ADs in Peninsular Malaysia.

The AD which has a more consistent boundary is chosen as the basic geographical unit of socioeconomic analysis in this study. Furthermore, all resources of data such as census data, administrative registration data, vital statistics and insurance data are available in terms of AD level, allowing the data from different resources to be matched and combined for further geographical and socioeconomic analysis. After retaining highly correlated variables and checking for multicollinearity, thirteen variables out of twenty-three variables from multidimensional domains of deprivation are retained for the development of 2000 GID using PCA. Based on the PCA results, the first principal component which accounts for 70.08% of the overall variation is selected as the 2000 GID. The coefficients of the GID are positive and similar, indicating that the thirteen variables contribute similar weights. The 2000 GID is then applied to the eighty-one ADs in Peninsular Malaysia and the ADs are ranked and classified into four quartiles corresponding to the most affluent, the moderately affluent, the moderately deprived and the most deprived ADs. The 2000 GIDs of the eighty-one ADs in Peninsular Malaysia imply that the majority of affluent areas were located in the westcoast of Peninsular Malaysia where the main economic activities were services and manufacturing whereas the most deprived areas were mainly scattered in the northeast of Peninsular Malaysia where the main economic drivers were agriculture and fishery.

The Global Moran's I is 0.632 with $p<2.2e^{-16}$, indicating significant and positive spatial autocorrelation in socioeconomic deprivation across all spatial weights of the ADs. Further, results on Local Moran's I indicate that the spatial association of each AD in Peninsular Malaysia can be clustered into three spatial clusters; deprived, affluent and non-significant.

In Malaysia, several initiatives have been implemented by government and policymakers to reduce poverty and materially deprived areas across the nation. In particular, raising living standards of low-income households has been selected as one of the government's six important pillars or National Key Result Areas (NKRAs) which is included in a special programme called the Government Transformation Programme (GTP), newly launched in 2010. Examples of initiatives and activities to reduce poverty and materially deprived areas across the nation which have been implemented in 2009-2011 by the government through the GTP were the allocation of federal welfare assistance to those qualified on the first day of each month, the provision of support to increase home ownership by offering about 44,000 low-cost houses, the delivery of special programmes under 1 Azam which create and provide jobs, trainings and funds, the creation of employment opportunities via setting up Jimat 1 Malaysia stores and the increment of number of urban, rural and mobile medical clinics to cater for low-income households. Therefore, the GID proposed in this study provides an effective tool for identifying the most deprived areas in either national, state or district levels and can be used as a starting point to consider an in-depth study of a smaller area level of deprivation. The GID can also be used as a practical tool in the activities and initiatives of managing the polarity between the economically well-off and the deprived to ensure that no Malaysian is left behind. However, geographic localization and spatial condition of each AD is important to be taken into account in the process of allocating resources and implementing efficient policies, since there is a strong evidence of global and spatial autocorrelation in socioeconomic deprivation across the ADs.

LIMITATIONS

The GID developed in this study has its limitations. Since, it is a measure of socioeconomic conditions observed at neighborhood levels, the index could be used as an ecology measure and not an individual measure. In addition, the index indicates relative measures of socioeconomic status and therefore cannot be used to provide information on absolute levels of poverty within a community despite its useful measure of considering inequality between the ADs. Even though, the PCA applied in this study has several advantages, the fact that the principal components are artificially constructed indices is still being debated among researchers. Critics of PCA argue that the technique is arbitrary or in other words, the method of selecting the number of components and the variables to be included is not well defined (Vyas and Kumaranayake, 2006).

IMPLICATIONS

This study can be extended in several directions in the future. Firstly, utilization of more indicators should be pursued to provide greater depth and insight to the results as long as the data for all ADs in Peninsular Malaysia are available and they are compatible and applicable over time besides being relevant to the concept of deprivation in the Malaysian context. Secondly, besides census and administrative registration data, the more updated data from other reliable resources should be considered. Thirdly, in addition to the measurement of overall and local clustering based on the Global and the Local Moran's I, further measurement of inequality and polarization of districts and regions should be tested. Finally, the GID proposed in this study can be further analyzed by testing its relationship with other variation of health outcomes and the successfulness of entrepreneurship.

ACKNOWLEDGEMENTS

The researchers are grateful to the Department of Statistics, Malaysia (DOSM) for providing the census data for this study. The financial and technical support for this study was sponsored by the Centre of Technopreneurship Development (C-TeD), Universiti Teknikal Malaysia Melaka (UteM).

REFERENCES

Abu, K.S.S. and A.S.J.A. Abu, 2009. Socio economic deprivation indices. Am. J. Sci. Res., 2: 37-51.

Alderman, I. and C.T. Morris, 1967. Society, Politics and Economic Development: A Quantitative Approach. Johns Hopkins University Press, Baltimore, Maryland, Pages: 307.

Anselin, L., 1995. Local indicators of spatial association-LISA. Geograph. Anal., 27: 93-115.

Bartley, M. and D. Blane, 1994. Commentary: Appropriateness of deprivation indices must be ensured. BMJ., 309: 1479-1479.

Bellani, L. and C. Ambrosio, 2011. Deprivation, social exclusion and subjective well-being. Soc. Indic. Res., 104: 67-86.

Car, H.R., 1988. Revising the RAWP Formula: Indexing Deprivation and Modelling Demand. York University, Toronto, Ontario,.

Carstairs, V. and R. Morris, 1989. Deprivation: Explaining differences in mortality between Scotland and England and Wales. Br. Med. J., 229: 886-889.

Choi, M.H., K.S. Cheong, B.M. Cho, I.K. Hwang and C.H. Kim *et al*., 2011. Deprivation and mortality at the town level in Busan, Korea: An ecological study. J. Preventive Med. Public Health, 44: 242-248.

Crampton, P., C. Salmond and F. Sutton, 1997. NZDep91 index of deprivation instruction book. Health Services Research Centre, Mancheste, England.

Dominguez, B.M.F., C. Borrell, J. Benach and M.I. Pasarin, 2001. Measures of material deprivation in small area studies. Gaceta Sanitaria SESPAS., 15: 23-33.

Ertur, C. and W. Koch, 2006. Regional disparities in the European Union and the enlargement process: An exploratory spatial data analysis, 1995-2000. Ann. Reg. Sci., 40: 723-765.

Fam, S., A.A. Jemain and Z.W.Z. Wan, 2011a. Spatial analysis of socioeconomic deprivation in peninsular Malaysia. Int. J. Arts Sci., 4: 241-255.

Fam, S.F., A.A. Jemain and K. Ibrahim, 2011b. The association between material deprivation and relative risk of infant mortality in peninsular malaysia. Int. J. Arts Sci., 4: 257-269.

Felipe, J. and M. Resende, 1996. A multivariate approach to the measurement of development: Asia and Latin America. J. Dev. Areas, 30: 183-210.

Gordon, D., 1995. Census based deprivation indices: Their weighting and validation. J. Epidemiol. Community Health, 49: S39-S44.

Havard, S., S. Deguen, J. Bodin, K. Louis and O. Laurent *et al*., 2008. A small-area index of socioeconomic deprivation to capture health inequalities in France. Social Sci. Med., 67: 2007-2016.

Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. J. Educ. Psychol., 24: 498-520.

Jarman, B., 1983. Identification of underprivileged areas. Br. Med. J., 287: 1705-1706.

Jobson, J.D., 1992. Applied Multivariate Data Analysis: Categorical and Multivariate Methods. Springer, Berlin, German,.

Kallmann, K., 1997. An Introduction to Social Indicators and Social Reporting: A Digest of the Literature. University of KwaZulu-Natal, Durban, South Africa,.

McIntyre, D., D. Muirhead and L. Gilson, 2002. Geographic patterns of deprivation in South Africa: Informing health equity analyses and public resource allocation strategies. Health Policy Plan., 17: 30-39.

Noble, M., H. Barnes, G. Wright and B. Roberts, 2010. Small area indices of multiple deprivation in South Africa. Social Indic. Res., 95: 281-297.

Norris, G., 1979. Defining urban deprivation, in urban deprivation and the inner city. Croom Helm, London, England.

Pampalon, R., D. Hamel, P. Gamache and G. Raymond, 2009. A deprivation index for health planning in Canada. Chron. Dis. Can., 29: 178-190.

Pearson, K., 1901. On lines and planes of closest fit tosystems of points in space. Phil. Mag., 2: 559-572.

Ram, R., 1982. Composite indices of physical quality of life, basic needs fulfilment and income: A principal component representation. J. Dev. Econ., 11: 227-247.

Salmond, C. and P. Crampton, 2002. NZDep2001 index of deprivation. Department of Public Health, Wellington, New Zealand.

Salmond, C., P. Crampton and F. Sutton, 1998. NZDep96 index of deprivation instruction book. Health Services Research Centre, Mancheste, England.

Salmond, C., P. Crampton and J. Atkinson, 2007. NZDep2006 Index of Deprivation. University of Otago, Dunedin, New Zealand,.

Sanchez, C.C., R.R. Ocana and A.A. Fernandez, 2008. Deprivation index for small areas in Spain. Social Indic. Res., 89: 259-273.

Singh, G.K., B.A. Miller, B.F. Hankey, E.J. Feuer and L.W. Pickle, 2002. Changing area socioeconomic patterns in US cancer mortality, 1950-1998: Part I-all cancers among men. J. Nat. Cancer Inst., 94: 904-915.

Testi, A. and E. Ivaldi, 2009. Material versus social deprivation and health: A case study of an urban area. Eur. J. Health Econ., 10: 323-328.

Townsend, P., 1987. Deprivation. J. Social Policy, 16: 125-146.

Townsend, P., 1993. The International Analysis of Poverty. Harvester Wheatsheaf, Great Britain, United Kingdom isBN:9780745013756, Pages: 291.

Vyas, S. and L. Kumaranayake, 2006. Constructing socio-economic status indices: How to use principal components analysis. Health Policy Plan., 21: 459-468.

Yuan, Y., F. Wu and X. Xu, 2011. Multiple deprivations in transitional Chinese cities: A case study of Guangzhou. Urban Affairs Rev., 47: 721-736.