Appendix II

Youth Non-Participation in Songkhla Province of Thailand

Youth Non-Participation in Songkla Province of Thailand

Sunaree Suwanro and Phattrawan Tongkumchum*

Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Songkhla, 94000 Thailand.

> *Corresponding author e-mail:tphattra@bunga.pn.psu.ac.th

Abstract

The objective of this study was to examine the effects of demographic and geographic factors on the prevalence of youth non-participation at work and at school using statistical models. The analysis focused on youth aged 15-17 years both absent from school and not employed in the workforce in Songkhla province. The determinants were gender, religion (Muslim or other), and geographical location based on subdistricts of Songkhla province. Data were obtained from the 2000 Population and Housing Census of Thailand. A non-linear regression model similar to the Lee-Carter model widely used by demographers to forecast human mortality was used to define an index of disparity between gender-religion groups that varied with location in the province.

The results showed substantial differences both in the levels of non-participation between the gender-religious groups, and also widely differing disparity.

The technique used in this study can identify locations where levels of both non-participation and disparity are high, and thus likely to lead to breakdowns in social order.

Keywords: Demographic factors, School non-attendance, Songkhla Province, Logistic regression

Introduction

Social scientists are concerned with inequalities in societies. These include gender, age-specific, cultural, religious, regional differences in education, health, employment, wealth, and political freedom, and one important application of social science theory is concerned with the merits of different strategies for alleviating such inequalities. For example, "affirmative action" has been used by universities to justify accepting less-qualified applicants from disadvantaged backgrounds in preference to other better-qualified applicants. This policy has also been used by government offices to provide employment opportunities to less-qualified applicants with disadvantaged backgrounds. Relevant recent studies include those by Healy et al (2003) in Thailand, Ho (2006) in Hong Kong, Basu and Basu (2005) in Australia, Neupane (2010) in Nepal, and Jhingran and Sankar (2009) in India.

The merits of such strategies continue to be debated. Apple (2004) stated that education systems that stream students too early can perpetuate class inequalities by channeling them into inflexible career paths. Davies and Guppy (2006) argued that schooling contributes to the reproduction of class differences, using the concept of

"cultural capital" developed by Bourdieu (1997) and applied to a study of Japanese children by Bassani (2003). According to Stehr (1999), inequality regimes are now substantially changing based on the extent to which knowledge is a key stratification element, a theme developed further by Sato (2010).

Although Thailand could be regarded as one of the developing nations where social inequality exists to a lesser extent than in other countries, and where all its citizens have similar educational and employment opportunities regardless of gender, region, wealth and ability, there are pockets where substantial disparities exist. For example, in their study of demographic factors affecting education, Thongchumnum and Choonpradub (2008) found substantial differences by gender and religion in secondary school completion rates in Pattani province.

Research into social inequality uses data from sample surveys and registries. Son (2007) used labour force surveys in Vietnam in 2002 and Thailand in 2004 to develop mathematical models that explain the disparity in welfare between male and female workers. Kauppinen (2007) sampled data from an employment statistics register of 15-year-olds living in Helsinki during 1990-1994 to study the effects of neighbourhood-level factors on residents' secondary education completion. Sittichai et al (2009) used data from the registrar's office at Pattani campus to fit a logistic regression model to university discontinuation rates, finding substantial differences by gender, religion and faculty. However, such data collection can be expensive and time-consuming, whereas census data is usually available at minimal cost.

The data available from the 2000 Population and Housing Census of Thailand (National Statistical Office, 2001) include individual records containing employment status (employed, not employed, or unknown), school attendance status (attending, not attending, unknown), gender, religion, district of residence, and age group. These data provide an opportunity to define an outcome that is a measure of youth disadvantage for a demographic group, defined as the proportion of residents in the demographic group (defined by specifying the combination of gender, age group, religion and district) not attending school and not employed. Since most school students are less than 18 years of age and most of those employed are at least 15 years of age, we chose the 15-17 year-olds as the relevant age group for this study.

Our objective was thus to investigate the variation in these *non-participation* (that is, not attending school and not employed) rates with respect to two factors, location (district or larger region) and gender-religion in a region of Thailand where such social inequality has been found or is believed to exist. We chose Songkla as the study region for two reasons, first because an earlier study by Thongchumnum et al (2008) had identified this province as one where secondary education completion provides no benefit in gaining employment, and second because it is the largest province in Thailand with a substantial proportion of both Muslim and non-Muslim residents.

We chose non-participation as the outcome of interest with a view to determining its extent of social disparity, that is, the extent to which its incidence rate varies with location, gender and religion. We believe that it is important to know the answer to this question because there is a body of evidence in the social science

literature that disparity in education and employment opportunities between localities, between religious groups and between males and females can seed social unrest. However, if these disparities can be defined and measured using appropriate methods, preventative strategies can be developed.

The population registration and housing survey in 2000 (National Statistical Office 2006) identified 1,255,662 residents in Songkla province, of whom 23.1% stated their religious affiliation to be Muslim.

In 2000 only 36.6% of Songkhla's population had completed secondary school. Moreover, 13.5% of all 15-17 year olds were both unemployed and out of school. Numerous studies, including Cowell (2008), have used incomes for measuring inequality within a society, usually justified by findings that dependent children share the socio-economic conditions and well-being of their parents (Avramov, 2002). In addition, Hossain and Keus (2004) concluded that the family size, education and standard of living index is inversely related with the incidence of poverty. Disadvantage is often generational and Gesemann (2007) concluded that education is the key to integration.

Philmore and Goodson (2008) used education and employment as indicators of integration within society. Focusing on participation by young people in education and/or employment can help predict likely future levels of integration in the community, as well as probable disparity and tensions. Based on a study of Muslims in Germany, Figen (2007) emphasized that without participation in education there cannot be satisfactory participation in the society's decision making. In another recent report on the integration of young Muslims in Germany, Gesemann (2007) stated that unsatisfactory achievement in schooling and high unemployment present a warning signal that feelings of exclusion and marginalization will be exploited among disadvantaged youths to create conflict.

Although the study by Cowell (2008) combined income and purchasing power, there are few other published studies using such indicators. Avranov (2002) looked at income, education and housing as composite indicators of disadvantage. This comparison of one's relative condition might be relevant to perceptions and feelings about being 'not at school and not employed'. Also, in Songkhla province in southern Thailand there are differences in religion that represent a demarcation within the population. Perceptions of the relative position of oneself and also of one's religious or gender group could be related to subsequent behaviour and unrest. A composite school-employment index could be useful for identifying locations of social inequality in the region studied.

This study examined religion and gender, as well as place of residence in the province, as factors among 15-17 year-olds in being not at work and not at school.

Methods

Study design and variables

The study was cross-sectional, based on population data selected from the 2000 Population and Housing Census of Thailand. Omitted were persons who did not state their age, and persons aged less than 15 years or greater than 17 years, giving a total study sample of 70,621. The determinants were defined as gender, religion

(Muslim or non-Muslim) and region (sub-district or aggregated sub-district) of residence. The adverse outcome was the binary variable defined as giving the answer "no" to both the questions "attending school" and "employed" on the form for the 2000 Population and Housing Census of Thailand.

Definition of regions

Songkhla has 127 sub-districts. Some sub-districts had relatively small populations for either Muslim or non-Muslim residents. A further 62 of the 127 sub-districts had fewer than 50 Muslims aged 15-17 and seven had fewer than 50 non-Muslims aged 15-17. To ensure that the statistical analysis was not compromised by small sample sizes, adjoining sub-districts were combined where necessary to form larger regions, each with a minimum total population of 1,600 persons for both Muslim and non-Muslim residents. For example Thungwang sub-district had a total of only 780 Muslims, so it was combined with Kotaeo sub-district. This reduced the number of residence locations from 127 to 52, and as a result all such regions had at least 50 respondents aged 15-17 in each religion group. Figure 1 shows how these regions were defined, with dotted lines indicating the sub-district borders and continuous lines separating the larger regions.

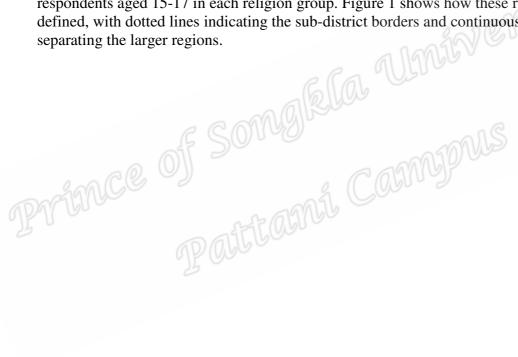
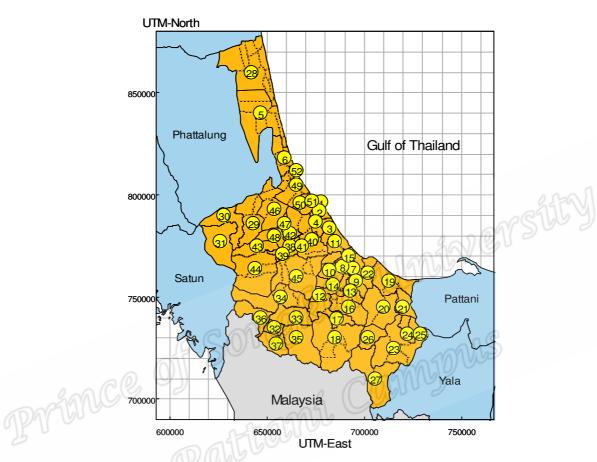


Figure 1
Map of 127 sub-districts in Songkhla, with combinations creating 52 statistical regions



1 BoYang	14 KlongPia+TalingChan	27	KaoDang+BaHoi	40	N HatYai+N Namom
2 KhaoRupChang	15 NamKhao+KhunTatWai	28	Ranot	41	KhoHong
3 KoTaeo+Tungwang	16 E Natawi	29	ERattaphum	42	KlongHae+KlongUtapao
4 Phawong+Koyo	17 W-C Nathawi	30	THaChamuang	43	Chalung
5 N SatingPa+Krasasin	18 S Nathawi	31	KhaoPhra	44	TungTamSao+WKhk
6 S SatingPa+BangKia	19 Tepa+PakBang+KoSaba	32	Sadao	45	EKhk+SHatYai+SNamom
7 BanNa	20 LamPlai+WangYai	33	Prik	46	KhuanNiang
8 PaChing	21 ThaMuang	34	N Sadao	47	N Bangklam
9 SapanMaiKaen+Sakom	22 SaKom	35	SamNakTaeo	48	ThaChang
10 NaWa	23 C Sabayoi	36	PadangBesa	49	N SingHaNakorn
11 ThaMoSai	24 Pian	37	SamnakKham	50	SathingMo
12 NaThap+Chanong	25 BanNot	38	HatYai	51	HuaKhao
13 Khu+Khae	26 Khuha	39	KhuanLang	52	MuangNgam

Statistical methods

Box plots are schematic data summaries used extensively in social science research for comparing distributions of social outcomes with respect to a democratic study factor (see, for example Burke et al, 2009). In preliminary data analysis we used these plots to compare non-participation rates in the four study groups.

Logistic regression (see, for example, Kleinbaum and Klein, 2002) is an appropriate method for comparing binary outcome proportions with respect to multiple determinants such as region and demographic group. Denoting by E[Y] the expected value of a random variable Y, the simplest model takes the form

$$E[Y_{ij}] = \alpha_i + \beta_j.$$

In this model

$$Y_{ij} = \ln \left(\frac{p_{ij}}{1 - p_{ij}} \right),$$
(2)

and p_{ij} denotes the probability of an adverse outcome in region i and religion-gender group j, where j takes values 1 for Muslim males, 2 for Muslim females, 3 for non-Muslim males and 4 for non-Muslim females. The terms α_i and β_j thus represent effects associated with region i and demographic group j, where the demographic effects are scaled to have mean 0 to avoid overparametrization. Equations (1) and (2) give the probability of the adverse outcome as

$$p_{ij} = \frac{1}{1 + \exp(-\alpha_i - \beta_j)}.$$
(3)

The logistic regression model was fitted to the counts in cells defined by combinations of demographic group and region, and the adequacy of the model was assessed by examining the linearity in the plot of deviance residuals against normal quantiles (Venables and Ripley 2002, Chapter 7).

To allow for possible interactions between region and demographic group, model (1) may be extended to a more general multiplicative model of the form

$$E[Y_{ij}] = \alpha_i + \gamma_i \beta_j.$$
(4)

In this model the demographic parameters are scaled to have unit variance as well as mean 0. The additional parameters γ_i provide a measure of the disparity in the adverse event rate between the different demographic groups in region i. Thus if region i has $\gamma_i = 0$, it means that there is no difference in the school adverse event rates between demographic groups in this region, whereas if γ_i is large in magnitude there is a high disparity between these groups. We call this measure the *disparity index* for the specified region.

Model (4) is non-linear and thus cannot be fitted simply using regression. However, Theil (1983) showed that the least squares estimates of the β_j parameters in model (4) are the elements of the eigenvector of the matrix $Y_c^T Y_c$ corresponding to its largest eigenvalue, where Y_c is the matrix with elements $y_{ij} - \overline{y}_i$ and Y^T denotes the transpose of Y. The corresponding least squares estimates of the γ_i parameters are then expressed in terms of the eigenvector components that define the *disparity index*.

$$\gamma_i = \sum_{i=1}^4 \beta_j \left(y_{ij} - \overline{y}_i \right). \tag{5}$$

Since the vector β_j is scaled to have mean 0 and standard deviation 1, it has only two free parameters. If these parameters are regarded as fixed, the model (4) can be fitted using standard linear regression, which provides both estimates and standard errors for the remaining parameters. The total number of parameters is thus 2n where n is the number of regions.

Model (4) thus contains a pair of parameters (α_i, γ_i) for each region, where α_i is the proportion of non-participating subjects and γ_i is the disparity index measuring the extent to which different demographic groups have different non-participation rates. This model has been used extensively for mortality forecasting in population science, where it is known as the Lee-Carter model (see, for example, Lee and Carter 1992). In this research y_{ij} is the logarithm of the mortality rate in a population where the indexes i and j refer to age group and year, respectively.

To allow for values of p_{ij} equal to 0 or 1, a small constant d was added to the numerator and denominator in Equation (2) before log-transforming. Having estimated the vector $\boldsymbol{\beta}_{j}$ in model (4) by least squares we then used logistic regression to estimate the parameters (α_{i}, γ_{i}) for each region. Sum contrasts (Venables and Ripley 2002, Chapter 6) were used instead of the standard treatment contrasts, so that the standard error for each α_{i} parameter provides a confidence interval for the difference between the non-participation rates in region i and the overall mean participation rate in the study area.

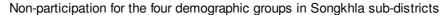
The final results are displayed as a scatter plot of the estimated parameters (α_i, γ_i) for the regions. This plot shows the pattern of both non-participation and disparity in the study area. Using the standard errors estimated from the logistic regression model, the regions may then be classified into groups according to whether the confidence intervals for the non-participation rates exceed, contain, or fall below the overall mean, and according to whether the confidence intervals for the disparity indexes exceed zero, contain zero, or fall below zero.

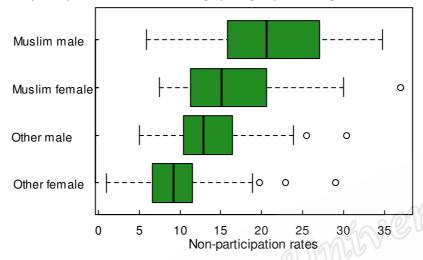
Results

Preliminary analysis

Figure 2 shows box plots of the non-participation rates in the four demographic groups. This graph shows substantial differences between the non-participation rates in the four demographic groups. Muslims had higher non-participation rates, with the maximum 37.0% (for females in North Singha Nakorn) and median values slightly greater than 20% for males and 15.1% for females. The minimum was 1.0% (among non-Muslim females) and the median values for non-Muslims were 9.2% for females and 12.9% for males. Maximum values for non-Muslims were 29% for females (in Sakom) and 30.4% for males (in BanNot).

Figure 2 Non-participation rates of demographic groups





Logistic regression modeling

When logistic regression was fitted to the non-participation rates, the plot of residuals from the additive model (2) showed a clear departure from linearity, whereas model (4) showed an improved fit, so model (4) was used for the analysis. The least squares estimates of β_j were 1.077 for Muslim males, 0.453 for Muslim females, -0.285 for non-Muslim males, and -1.246 for non-Muslim females. Figure 3 shows plots of observed counts and percentages against fitted values as well as residuals against normal quantiles in its upper panels, with a scatter plot of the non-participation rates versus disparity indexes for the regions in the lower panel. While these plots indicate that the fit of the model is reasonable good and the

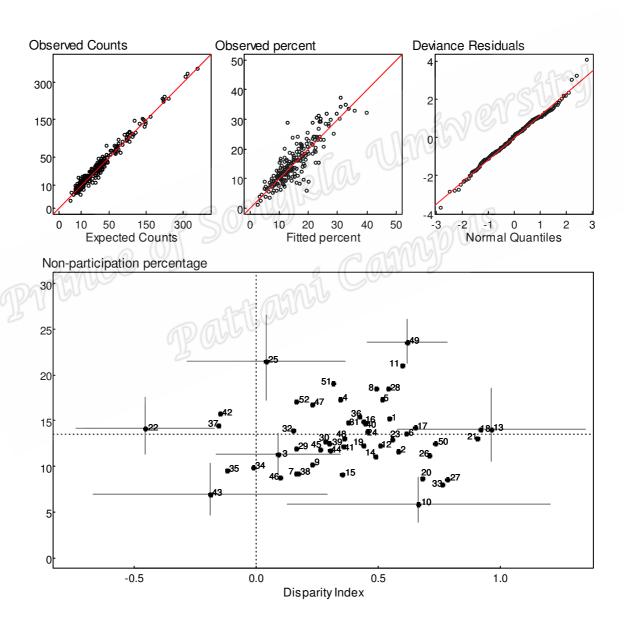
The lower panel shows a scatter plot of non-participation rates and disparity indices for the 52 statistical regions in Songkhla province. On the x-axis, a score of zero for the disparity index would mean that, for the particular region, all four gender-religion groups would have the same non-participation levels, a positive value indicates greater non-participation rates for Muslim youth, and a negative value indicates greater non-participation rates for non-Muslims. On the y-axis, a placement near the dotted line would indicate that, for that particular region, the overall level of non-participation would be close to the overall average for all regions. Regions with points above or below the dotted line have non-participation rates that are higher or lower, respectively, than average.

The plot also shows confidence intervals that can be used to make valid statistical conclusions. The vertical line segments denote 95% confidence intervals for the non-participation rates (when compared with the overall mean), whereas the horizontal line segments denote 95% confidence intervals for the disparity indexes (compared to zero disparity). To reduce clutter, the confidence intervals are shown only for selected regions. Using these confidence intervals the regions may be

classified into nine groups according to whether (a) the confidence interval for the non-participation percentage is wholly above the mean, crosses the mean, or is wholly below the mean, and (b) whether the confidence interval for the disparity index is wholly above 0, contains 0, or is below 0.

Figure 3

Multiplicative model results for youth non-participation rates in Songkhla regions



This classification thus gives rise to nine possible clusters as follows.

1: Higher than average non participation, Muslim disadvantage (7

regions).

2: Higher than average non participation, no evidence of disparity (3

regions).

- 3: Higher than average non participation, Muslim advantage (no region).
- 4: Average non-participation, Muslim disadvantage (21 regions).
- 5: Average non-participation, no evidence of disparity (7 regions).
- 6: Average non-participation, Muslim advantage (1 region).
- 7: Lower than average non-participation, Muslim disadvantage (4 regions).
- 8: Lower than average non-participation, no evidence of disparity (9

regions).

9: Lower than average non-participation, Muslim advantage (no region).

In seven regions, 4 (Phawong+Koyo), 5 (North SatingPa+Krasasin), 8 (Paching), 11 (ThaMoSai), 28(Ranot), 49 (N SinghaNakorn) and 51 (HuaKhao) Muslims had higher than average non-participation and disparity disadvantage. In three regions, 25 (BanNot), 47 (N Bangklam) and 51 (HuaKhao) Muslims had higher than average non-participation but no evidence of disparity. Muslims in the Sakom region had average non-participation and a high negative disparity index (i.e. a disparity advantage).

Figure 4 shows a corresponding thematic map, where disparity is labeled as MD for Muslim disadvantage and MA for Muslim advantage.

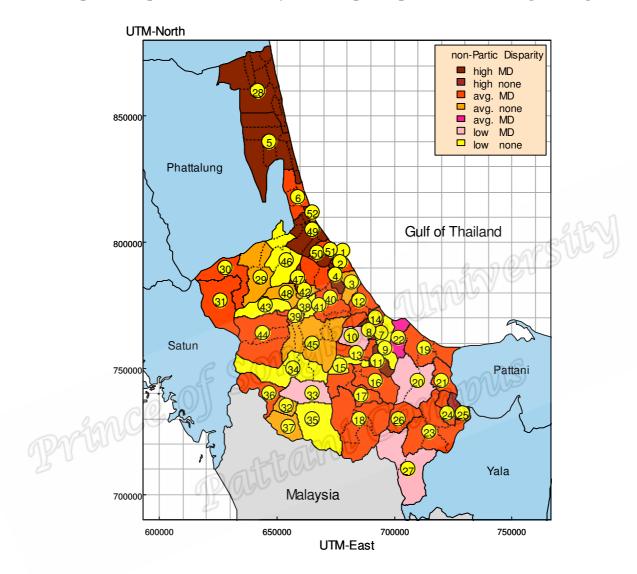


Figure 4
Map of multiplicative model for youth non-participation rates in Songkhla regions

Conclusion and Discussion

This study investigated the factors associated with failure to be at school or work among 15-17 year olds in Songkhla Province. Of the total (70,621) population of 15-17 year olds in Songkhla in 2000, 13.5% were 'not at work and not at study', and therefore at risk of the consequence of limited life chances such as not acquiring a good education, employment, income, and health.

Lack of education among children can lead to further social problems such as crime (Buonanno and Montolio, 2008), illicit drugs, HIV infection, and other social problems (United Nations, 2000).

There were significant variations in non-participation rates between geographical regions, varying from 37% of Muslim females in North Singha Nakorn) to 1.0% of non-Muslim females in Na Wa. Since both gender and religious group

differences were found, it is possible that the general level of inequality in a region might itself have influenced the outcome, i.e., the proportion of non-participating youth. Therefore, disparity indexes were used as a measure of inequality. The construction of a Disparity Index used the average (scaled) set of percentages for the four religion-gender groups and enabled the use of a model that placed locations (regions) on a scatter graph (Figure 3).

In summary, the use of a logistic model and the map shown in Figure 4, with a disparity index of levels of demographic differences, can highlight locations where there is a relatively high level of social inequality. Avranov (2002) suggested that perception of relative disadvantage rather than an 'absolute' deprivation, causes dissatisfaction and that depends in part on the extent to which others in the same location experience the disadvantage. This suggests a need for further studies investigating the association between social unrest and social disparity.

While the study was largely a statistical exercise and further research is needed to validate the method it developed, such studies can be useful in providing measurable indices of social inequality, which in turn can provide factual bases for social science theory and social planning in societies.

Acknowledgements

We are grateful to Don McNeil, Apiradee Lim, and Grieg Rundle for their assistance and suggestions. Thanks are also due to the National Statistical Office of Thailand for providing the data.

References

- Apple, W. 2004. *Ideology and curriculum* (3rd ed.). New York: RoutledgeFalmer.
- Avramov, D. 2002. *People, Demography and Social Exclusion*. Population Studies No.37. Council of Europe Publishing. Belgium.
- Bassani, C. 2003. Social Capital Theory in the context of Japanese Children. [Online] Available from: http://japanesestudies.org.uk/articles/Bassani.html. [Accessed 2010 November 10]
- Basu, S., Basu, P. 2005. Regional Disparity in Australia: Analysis of Gender Development Index. *International Review of Business Research Papers*. 1(2): 56-66.
- Bourdieu, P. 1997. The forms of capital. In AH Halsey, H. Lauder, P. Brown, & S. Wells (Eds.), *Education: Culture, economy, and society*. Oxford, UK: Oxford University Press.
- Buonanno, P., Montolio, D. 2008. Identifying the socio-economic and demographic determinants of crime across Spanish provinces. *International Review of Law and Economics*. 28:89-97.

- Burke, M. B., Miguel. E,. Satyanath, S., Dykema, J., Lobell, D. 2009. Warming in creases the risk of civil war in Africa. *PNAS*. 106(49): 20670-20674. [Online]. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.0907998106 [Accessed 2010 November 23]
- Cowell, F. A. 2008. Measuring Inequality in LSE Perspectives in Economic Analysis. Oxford University Press. [Online]. Available from: http://darp.lse.ac.uk/papersdb/Cowell_measuringinequality3.pdf [Accessed 2009 February 9].
- Davies, S., Guppy, N. 2006. *The schooled society: An introduction to the sociology of education*. Don Mills, Oxford University Press.
- Figen, U. 2007. Social Participation of Turkish and Arabic Immigrants in the Neighborhood: Case Study of Moabit West, Berlin. *Journal of Identity and Migration Studies*. 1(2): 44-66.
- Gesemann, F. 2007. The Integration of Young Muslims in Germany Education and Training as Key Areas for Social Integration. [Online] Available from: http://www.migrationeducation.org/22.1.html?&rid=82&cHash=6af0a687 4a [Accessed 2009 February 7].
- Healy, A. J., Jitsuchon, S., Vajarajupta, Y. 2003. Spatially Disaggregated Estimates of Poverty and Equality in Thailand. [Online] Available from: http://siteresources.worldbank.org/INTPGI/Resources/342674-1092157888460/Healy.DisaggregatedThailand.pdf [Accessed 2010 November 18]
- Ho, E. S. 2006. Social Disparity of Family Involvement in Hong Kong: Effect of Family Resources and Family Network. *The School Community Journal*. 16(2): 7-26.
- Hossain, M., Keus, E. 2004. Socio-economic Condition and Poverty Situation in Rural Southern Part of Bangladesh: A household Level Study. *Journal of Population and Social Studies*. 12(2):135-163.
- Jhingran, D., Sankar, D. 2009. Addressing Educational Disparity Using District Level Education Development Indices for Equitable Resource Allocation in India. Policy Research Working Paper No.4955. The World Bank South Asia Region. India.
- Kauppinen, T. M. 2007. Neighborhood effects in a European city: Secondary education of young people in Helsinki. *Social Science Research*. 36: 421-444.
- Kleinbaum, D. G. & Klein, M. 2002. *Logistic Regression: A Self-Learning Text* (2nd ed). Springer-Verlag. New York.
- Lee, R. D. and Carter, L, R. 1992. Modeling and forecasting US mortality. *Journal of American Statistical Association*, 87(419):659-671.
- National Statistical Office, 2006. Classification of Population: Gender, Housing and Housing Types Regions and Provinces in 2000. Bangkok.

- National Statistical Office, 2001. *The 2000 Population and Housing Census: Changwat Songkhla*. Bangkok.
- Neupane, P. 2010. Impacts of Gender Disparity in Education on Human Resource Development in Asia: A case of Nepal. 14th UNESCO-APEID International Conference 21-23 October 2010 Bangkok Thailand.
- Phillimore, J. and Goodson, L. 2008. Making a Place in the Global City: The Relevance of Indicators of Integration. *Journal of Refugee Studies*. 21(3): 305-325.
- Sato, Y. 2010. Disparity Society Theory and Social Stratification Theory: An attempt to Respond to Challenges by Disparity Society Theory. [Online] Available from: http://www.waseda-giari.jp/sysimg/rresults/456_report_2-1.pdf. [Accessed 2010 November 18].
- Sittichai, R. Tongkumchum, P., McNeil, N. 2009. Discontinuation among University Students in Southern Thailand. *International Education Studies*. 2(2): 17-20.
- Son, H. H. 2007. Occupational Segregation and Gender Discrimination in Labor Market: Thailand and Viet Nam. ERD Working Paper No. 108. Asian Development Bank Philippines.
- Stehr, N. 1999. The Future of Social Inequality. Society. 36(5): 54-59.
- Theil, H. 1983. *Linear Algebra and matrix methods in econometrics*. North-Holland Publishing Company, Amsterdam, the Netherlands.
- Thongchumnum, P., Choonpradub, C. 2008. Trends in Secondary School Education Completion in Pattani Province of Thailand. *Songklanakarin J. of Social Science & Humanities*. 14(4): 575-584.
- Thongchumnum, P., Suwanro, S. Choonpradub, C. 2008. Demographic Factors Affecting Employment in Pattani and Songkla Provinces of Thailand. *Asian Social Science*. 4(11): 169-176.
- United Nations. 2000. Youth in Thailand: A Review of the Youth Situation and National Policies and Programs. New York.
- Venables, W. and Ripley, B. 2002. *Statistics Complements to Modern Applied Statistics with S Fourth edition*. New York.