WID.world TECHNICAL NOTE SERIES N° 2015/3

Top Incomes in Malaysia 1947 to the Present (With a Note on the Straits Settlements 1916 to 1921)

Anthony B. Atkinson

March 2015

25 March 2015

Top incomes in Malaysia 1947 to the present
(with a note on the Straits Settlements 1916 to 1921)

A B Atkinson

The purpose of these notes is to describe the estimates of the distribution of top incomes in Malaysia derived from tabulations of income tax data. They cover Peninsular Malaya from 1947 to 1988, and Malaysia from 1992 to the present. There is in addition a note on the War Tax levied in the Straits Settlements during and after the First World War.

There are three main ingredients to the estimates:

- The underlying data from tabulations of income taxpayers by ranges of income; these data are available from the introduction of the income tax in the colonial period (first income year 1947) to the present day. The data sources and the methods of interpolation are described in section 1.

- The construction of a control total for the total potential number of taxpayers (tax units) based on demographic data; the purpose of this control total is to allow us to express taxpayers as a percentage of the total, so that the results for the top 1 per cent refer to the top 1 per cent of potential taxpayers (not just the top1 per cent of actual taxpayers). The derivation of the control totals for tax units is described in section 2.

- The construction of a control total for the total income based on national accounts data; the purpose is to express the income of taxpayers as a percentage of the total household income. The derivation of the control totals for income is described in section 3.

A brief description of the results, in terms of income shares and characteristics of the upper tail, is given in section 4. The final section 5 discusses the war tax in the Straits Settlements.

1. The underlying income tax data

Income taxation was employed in many British colonial territories, and the colonial administrators were required to publish detailed reports, which typically included information on the distribution of taxpayers by income range and total incomes. Income tax was introduced into the Federation of Malaya with effect from 1 January 1948. The first *Report of the Income Tax Department*, *Federation of Malaya,*

published in 1950, gave details of the number of taxpayers assessed in 1948 by ranges of assessed income. The same information was published in annual reports (referred to as AR) for subsequent years and reproduced in the *Yearbook of Statistics, Malaysia* (referred to as YS). From these sources, it is possible to locate income tax data for much of the period. The largest gap in the series is from 1976 to 1982. The data sources are listed in Table A.1. The data for 1997 and 1998 have not been used, as the numbers and amounts are not consistent with the ranges. The data for 2004 are affected by the introduction of self-assessment, and have not been used.

The geographical coverage has varied over time. The Federation of Malaya dates from 1948 and included the previously federated states and the states that had been separate protectorates (Johor, Kedala, Kelantan, Perlis and Terengganu). It did not include Singapore. The Federation obtained independence on 31 August 1957. On 16 September 1963, the Federation, together with Singapore, British North Borneo (now Sabah) and Sarawak, formed the state of Malaysia. In August 1965, Singapore was expelled from Malaysia. Malaysia now consists of Sabah, Sarawak and Peninsular Malaya. The income tax data until 1988 cover Peninsular Malaysia, also known as the States of Malaya or as Western Malaysia. From 1992, the income tax data cover all Malaysia (no data have been found for 1989 to 1991).

The income tax data show the number of taxpayers assessed by ranges of assessed income and the total amounts of assessed income per range. The number of ranges was typically around 10 and they extended up to many multiples of the mean. Amounts are in Malay dollars, later (1953 to 1967) Malaya and British North Borneo dollars, and from 1967 Malaysian dollars, called Ringgit, for which the symbol RM is used. The data reflect the administrative process by which they are produced. For example, the data refer to a "year of assessment": e.g. in AR 1949 there is information for the year of assessment 1948, which refers to incomes during the year 1947, referred to as income year (IY) 1947. In this case, the assessments are those made in the 24 months after the end of the income year (i.e. up to 31st December 1949), but in a few cases the figures are given only after 12 months. In the 1999 Budget, the basis for assessment was changed to the current year, and income for 1999 was waived (YS2001, page 249). From 2000, the income year was the year of assessment. From 1994, the published statistics have been taken as relating to chargeable income, rather than assessed income. Chargeable income equals assessed income minus tax exemptions and tax reliefs. An adjustment has been made to provide a continuous series for assessed income, adding the personal relief (RM 8000) to all incomes, but this is only approximate.

The income tax was paid by non-resident as well as resident individuals. In what follows, attention is focused on Malaysian residents. The importance of non-resident income has declined. In 1947, non-resident taxpayers accounted for 8 per cent of the total, although very few were in the upper ranges, and the proportion of total income was 6 per cent. In 2010, non-resident taxpayers accounted for 0.7 per cent of the total, and the proportion of total income was 1.1 per cent.

Use of income tax data is always open to the charge that the data take no account of tax avoidance and tax evasion. These are clearly important considerations. Since the control totals for income are based on national accounts (see below), the estimates made here of the income *shares* understate the true top income shares to the

extent that incomes are not declared. In this sense the estimates provide a lower bound. It is also possible that the changes over time are affected. The introduction of the tax was a process of learning and early reports refer to the extension of powers taken to secure compliance, such as the authorisation to require full disclosure of a person's assets in cases where evasion was suspected under the Income Tax (Amendment) Ordinance of 1953 (AR 1952, page 2). To the extent that compliance has increased, the top shares in later years are closer to the true values. Any downward (upward) trend is therefore under (over) stated.

Since the basic data are in the form of grouped tabulations, and the intervals do not in general coincide with the percentage groups of the population with which we are concerned (such as the top 0.1 per cent), we have to interpolate in order to arrive at the shares of total income. Given that there is information on both the number of persons and the total income in the range, we use the mean-split histogram. The rationale is as follows. Assuming, as seems reasonable in the case of top incomes, that the frequency distribution is non-increasing, then restricted upper and lower bounds can be calculated for the income shares (Gastwirth, 1972). These bounds are limiting forms of the split histogram, with one of the two densities tending to zero or infinity - see Atkinson (2005). Guaranteed to lie between these is the histogram split at the interval mean with sections of positive density on either side. The refined bounds do not apply to percentiles (Atkinson, 2005), so the percentiles have been calculated by Pareto interpolation applied to each interval using the cumulative distribution.

The ranges are in some cases quite broad, and the possible errors of interpolation need to be taken into account. For example, in 2000, taxpayers above RM50,000 constituted 1.48 per cent of the adult population, and those above RM70,000 were 0.85 per cent. If we make no assumption about the distribution, then the "gross" bounds for the share of the top 1 per cent are from 6.04 to 6.12 per cent (these are calculated by assuming either that all incomes are equal to the mean for the range or that people are concentrated at the end points). If we assume that the frequency distribution is non-increasing (which rules out both of the bounds just described), then the restricted bounds are from 6.08 to 6.10, which are close. The mean-split histogram method gives a value for the share of the top 1 per cent of 6.09 per cent. In general, no extrapolation is made into the open upper interval, except in a few cases where the upper interval is close to one of the key percentages. Where the difference is less than 10 per cent, a simple Pareto extrapolation is used to calculate the share. For example, in 1963, the top interval contains 0.0054 per cent of tax units, and an estimate has been made of the share of the top 0.005 per cent. This has not however been done for 1961, when the top interval contained 0.0057 per cent of adults.

2. Control totals for tax units

In constructing the control totals, the aim is to estimate the total of tax units as defined in the income tax legislation, for the purposes of comparison with the figures recorded in the tabulations. The tax unit was the single adult individual or the married couple. This was taken to be the total population aged 15 and over minus the number of married women.

The construction of the control totals for population (and for income) is greatly aided by the document *Malaysia Economics Statistics - Time Series 2011* (MES), available from the website of Department of Statistics Malaysia. The population aged 15 and over is from MES, Table 21.1 for the full period 1947 to 2010. This is a consistent series, as is explained on page 225 of the document, allowing for the significant revisions that have been made to the demographic data in the *Yearbook of Statistics* (YS). For example, YS 2011 (Table 3.7) introduced new population estimates from 2010 based on the *Population and Housing Census of Malaysia 2010*, and adjusted for under-enumeration. The figures for 2011-2013 are from *The Malaysian Economy in Figures 2013*, Table 2.3. There are however differences in the geographic coverage. Prior to 1963, the figures refer only to Peninsular Malaysia. For years prior to the 1970 census, the age composition of the population was only available for earlier census years (1947 and 1957), and the proportions are linearly interpolated between these three census years. The number of married women is for census years (1947, 1957, 1970, 1980, 1991 and 2000) from *Time Series Data Population and Housing Census, 1911-2000*, Table 20, and 2010 from the *Population and Housing Census,* Report, section 9. The figures for intervening years are linearly interpolated.

The resulting series is shown in Table A.2.

3. Control totals for income

The construction of a control total for total household income (at current prices) proceeds here by first considering a measure of national income and then seeking to link total household income to national income.

The series for Gross National Income is from Table 1.1 of the document *Malaysia Economics Statistics - Time Series 2011* (MES), available from the website of Department of Statistics Malaysia. The series has been linked at 2010 to that given in *Statistical Yearbook 2013,* Table 17.1. The data cover Malaysia from 1963. For the years from 1963 to 1988 an estimate has been made of the total for Peninsular Malaya by applying the ratio of GDP for Peninsular Malaya for 1963 to 1975 from Hutton (1980, page 29) to the MES total for Malaysia (and assuming a percentage of 83 per cent for 1976 to 1988). For the years 1955 to 1962, the data cover Peninsular Malaya, but for 1947 to 1953 (1954 is missing and has been interpolated) the data cover Malaya and Singapore, based on the work of Benham (1951) and the World Bank. The estimates for Peninsular Malaya are obtained by subtracting the estimated national income of Singapore (from Atkinson, 2010, Table 5A.2). The

GNI does not correspond to total household income. Subtractions have to be made for depreciation, for retained corporate profits, for non-profit institutions, and additions made for government interest and transfers received by households. In the absence of household income information in the national accounts, it is not easy to make this adjustment and the estimates can only be based on rather limited evidence. Snodgrass (1980) has made estimates for five years between 1957 and 1970. These indicate that total personal income was 75 per cent of GNI in 1957 and 64 per cent in

1970. The latter figure is lower than that given by Chander et al (1980, Table 1), which implies 75 per cent in 1970. For more recent years, reference can be made to the incomes reported in household surveys. However, as noted by Snodgrass, and discussed further by Anand (1983), the household survey figures tend "to be about 25 per cent less than those derived from the national accounts" (Anand, 1983, page 39). Coming to a recent year, we see that the survey income total in the *Household Income and Basic Amenities Survey 2009* is 58.8 per cent of GNI (Table 1.2) and that for 1995 (Chart 1 in the same source) is 68.5 per cent.

An alternative perspective is provided by the expenditure side of the national accounts, where *Malaysia Economics Statistics - Time Series 2011* (MES), Table 1.5, gives total private final consumption expenditure. This falls short of total household income to the extent that there are positive savings and exceeds it to the extent that non-household spending is included and that there are items, such as imputed rents, not included in the income definition. At the beginning of the 1950s, total consumption was around 75 per cent of GDI. It then fell steadily as a percentage and by the end of the period (after 2000) was less than 50 per cent. This is consistent with the substantial rise in the savings rate that has taken place: with a savings rate of 30 per cent, an income share of 65 per cent leads to a consumption share of 45.5 per cent.

In the light of this limited evidence, it has been assumed that total household income is equal to 75 per cent of GNI in 1947-1950, then falling linearly to 65 per cent in 2010. In 2010, this implied that the average income of those not covered by the income tax statistics was 54 per cent of the average income of those covered.

It should be stressed that the resulting income totals, shown in Table A.2, are surrounded by considerable uncertainty. The adjustment from total GDI to household income is approximate and the early estimates of GNI by Benham were stated by him to "involve a considerable amount of guesswork" (1951, page 1).

4. Results

In considering the results, it should be borne in mind that there are a number of qualifications. In particular, the estimates would vary with different assumptions about the income control totals and about the adjustment for different income definitions. Figure 1 shows the estimated top income shares over the period 1947 to 2010. There is a spike around 1950 reflecting the boom in commodity prices, as observed in other countries with primary product exports. There was then an upward trend that continued after independence to the mid-1960s. This was followed by a definite fall from the 1970s. The fall in top shares over this period mirrors the more general decline in income inequality that has been emphasised: "Malaysia's declining inequality from 1973 to 1989 is unique in Asia" (Snodgrass, 2002, page 26). The past two decades have seen top shares falling after the Asian financial crisis and then recovering.

The shares depend on the control total for income, and the qualifications surrounding these figures have been emphasised in the previous section. The same qualification does not apply to Figure 2, which shows the shape of the distribution as measured by the inverse Pareto coefficient. This depends on the population control

totals but not the income totals. The inverse Pareto coefficient, β, shows the mean income above y as a multiple of y. It is calculated from the relative shares, S_i, of two groups, applying the formula that $(S_i/S_j) = [(1-F_i)/(1-F_j)]^{(1/\beta)}$, where F is the cumulative distribution. If the distribution were precisely Pareto in form, then this would be independent of y. As however may be seen from Figure 2, the implied values of β vary depending on the choice of i and j. In the early 1960s, β is close to 2 for the share of the top 0.5 in the top 1 per cent, around 1.75 for the share of the top 0.1 in the top 0.5 per cent, and close to 1.6 for the top 0.05 in the top 0.1 per cent. Moreover, there is no clear sign that income concentration, measured in this way, declined in the 1970s and early 198os. In recent decades, top income concentration has increased.

5. Income tax in the Straits Settlements 1916-1921

The Straits Settlements were a colony established in 1876 that continued in existence until the Second World War. It consisted of Malacaa, Penang, Singapore, Christmas Island and (from 1906) Labuan. The colony had begun consideration of an income tax in 1910, when a bill was introduced into the Legislative Council (*Interim administration report on the war tax ordinance, 1917,* Page 251). There was however strong public opposition and it was withdrawn. The proposal for an income tax was revived in 1916 when there was public support for increasing the colony's contribution to Imperial War Expenditure, and it was enacted as a temporary measure in 1917 with effect from the income year 1916. It was in operation for the income years 1916 to 1921 and provides evidence about the distribution of income in the Straits Settlements for these years.

The population figure for those aged 15 and over is derived from *Report on the census of the Straits Settlements taken on 1st March 1901*, page 15. From this is subtracted the number of married men (pages 55, 93 and 142), to arrive at a total of 311,270 tax units in 1901. This figure has been assumed to increase annually at the rate of growth recorded for Peninsular Malaya between 1911 and 1921 (2.2 per cent per year) from *Malaysia Economics Statistics - Time Series 2011*, Table 21.1. No attempt has been made to estimate a total income figure.

Around 1 per cent of tax units were assessed for income tax. Tabulations were published for income years 1916 to 1921:

Income year	AR = Administration Report of the Income Tax Department	published in Annual Departmental Reports
1916	AR 1917-18	Assessments
1917	AR 1918-19	Assessments
1918	AR 1919-20	Annex B
1919	AR 1920-21	Annex B
1920	AR 1921-2	Annex B
1921	AR 1922-3	Annex B

The results are summarised in terms of the beta (inverse Pareto) coefficients. These have been estimated by a linear regression of the logarithm of (1-F) on the cumulative total income. Applying the set out above, the slope is β. The resulting values of β are

1916	2.03
1917	2.20
1918	2.20
1919	2.11
1920	1.92
1921	1.82

The degree of concentration appears to have fallen significantly. The value of 1.82 in 1921 is not very different from the values around 1.7 shown for the late 1940s in Peninsular Malaya in Figure 2.

References

Anand, S, 1983, *Inequality and poverty in Malaysia*, Oxford University Press, Oxford.

Benham, F, 1951, *The national income of Malaya, 1947-49*, Government Printing Office, Singapore.

Chander, R, Gnasegarah, S, Pyatt, G and Round, J I, 1980, "Social accounts and the distribution of income: The Malaysian economy in 1970", *Review of Income and Wealth*, series 26: 67-85.

Hutton, J P, 1980, "Income tax elasticity and the distribution of income, with an application to Peninsular Malaysia", *South East Asian Economic Review*, vol 1: 13-34.

Snodgrass, D R, 1980, *Inequality and economic development in Malaysia*, Oxford University Press, Oxford.

Snodgrass, D R, 2002, "Economic growth and income inequality: The Malaysian experience" in M G Asher, D Newman and T P Snyder, editors, *Public policy in Asia,* Quorum Books, Westport.

Table A.1 Sources of income tax data

SA = Statistical Abstracts showing details of assessments
AR = Annual Report of the Department of Inland Revenue
YS = Yearbook of Statistics

Income year			assessed to	Geographical Coverage
1947	IT AR 1949	Statement E	31-Dec-49	Fed of Malaya
1948	SA 1950	Statement E	31-Dec-50	Fed of Malaya
1949	SA 1951	Statement E	31-Dec-51	Fed of Malaya
1950	SA 1952	Statement E	31-Dec-52	Fed of Malaya
1951	SA 1953	Statement E	31-Dec-53	Fed of Malaya
1952	SA 1954	Statement E	31-Dec-54	Fed of Malaya
1953	SA 1955	Statement E	31-Dec-55	Fed of Malaya
1954	SA 1956	Statement E	31-Dec-56	Fed of Malaya
1955	SA 1957	Statement E	31-Dec-57	Fed of Malaya
1956	SA 1958	Statement E	31-Dec-58	Fed of Malaya
1957	IT AR 1959	Abstract E	31-Dec-59	Fed of Malaya
1958	IT AR 1960	Abstract E	31-Dec-60	Fed of Malaya
1959	IT AR 1961	Abstract E	31-Dec-61	Fed of Malaya
1960	IT AR 1962	Abstract E	31-Dec-62	Fed of Malaya
1961	IT AR 1963	Abstract E	31-Dec-63	States of Malaya
1962	IT AR 1964	Abstract E	31-Dec-64	States of Malaya
1963	IT AR 1965	Abstract E	31-Dec-65	West Malaysia
1964	IT AR 1966	Abstract E	31-Dec-66	West Malaysia
1965	IT AR 1967	Abstract E	31-Dec-67	West Malaysia
1966	IT AR 1967	Abstract I	31-Mar-68	West Malaysia
1967	IT AR 1968	Abstract E	31-Mar-69	West Malaysia
1968	IT AR 1969	Abstract E	31-Mar-70	West Malaysia
1969	IT AR 1970	Abstract E	31-Mar-71	West Malaysia
1970	IT AR 1971	Abstract E	31-Mar-72	West Malaysia
1971	IT AR 1972	Abstract E	31-Mar-73	Pen Malaysia
1972	IT AR 1973	Abstract E	31-Mar-74	Pen Malaysia
1973	IT AR 1974	Abstract E	31-Mar-75	Pen Malaysia
1974	IT AR 1975	Abstract E	31-Mar-76	Pen Malaysia
1975	IT AR 1976	Abstract E	31-Mar-77	Pen Malaysia
1976				
1977				
1978				
1979				
1980				
1981				
1982				
1983	YS 1985	Table 12.11	31-Mar-85	West Malaysia
1984	YS 1986	Table 12.10	31-Mar-86	West Malaysia
1985	YS 1987	Table 12.10	31-Mar-87	West Malaysia
1986	YS 1988	Table 12.10	31-Mar-88	West Malaysia
1987				
1988	YS 1990	Table 12.10	31-Mar-90	West Malaysia
1989				
1990				
1991				
1992	YS 1993	Table 11.10	31-Mar-94	Malaysia
1993	YS 1994	Table 11.10	31-Mar-95	Malaysia
1994	YS 1997	Table 12.10	31-Mar-96	Malaysia
1995	YS 1998	Table 12.10	31-Mar-97	Malaysia
1996				

1997	YS 1999, data on numbers and amounts not consistent with ranges	Table 12.10	31-Mar-99	Malaysia
1998	YS 2000, data on numbers and amounts not consistent with ranges	Table 12.10	31-Mar-00	Malaysia
1999	income for 1999 waived			
2000	YS 2002	Table 12.11	31-Mar-02	Malaysia
2001	YS 2003	Table 12.11	31-Dec-02	Malaysia
2002	YS 2004	Table 13.11	31-Dec-03	Malaysia
2003	YS 2005	Table 13.11	31-Dec-04	Malaysia
2004	YS 2007, affected by the introduction of self-assessment (YS 2007, page 285)	Table 13.11	31-Dec-05	Malaysia
2005	YS 2008	Table 13.11	31-Dec-06	Malaysia
2006				
2007				
2008				
2009	YS 2010	Table 13.11	31-Dec-10	Malaysia
2010	YS 2011	Table 16.12	31-Dec-11	Malaysia
2011	YS 2012	Table 16.12	31-Dec-12	Malaysia
2012	YS 2013	Table 16.12	31- Dec-13	Malaysia

Table A.2 Control totals Peninsular Malaya and Malayasia

	Tax Units 000s		Total household income RM million	
	Peninsular Malaya	Malaysia	Peninsular Malaya	Malaysia
1947	2,053		1,957	
1948	2,067		1,936	
1949	2,086		1,721	
1950	2,126		2,658	
1951	2,150		4,149	
1952	2,197		3,563	
1953	2,255		3,133	
1954	2,304		3,331	
1955	2,348		3,527	
1956	2,399		3,583	
1957	2,385		3,653	
1958	2,478		3,505	
1959	2,559		3,907	
1960	2,520		4,118	
1961	2,616		4,121	
1962	2,717		4,369	
1963	2,817		4,632	
1964	2,911		4,923	
1965	3,014		5,336	
1966	3,126		5,603	
1967	3,232		5,869	
1968	3,329		6,056	
1969	3,429		6,872	
1970	3,623		7,484	
1971	3,758		7,694	
1972	3,890		8,429	
1973	4,025		10,817	
1974	4,162		12,948	
1975	4,304		12,694	
1976	4,448		15,829	
1977	4,598		18,177	
1978	4,746		21,124	
1979	4,901		25,831	
1980	5,061		29,858	
1981	5,217		32,228	
1982	5,367		34,526	
1983	5,522		38,091	
1984	5,666		42,689	
1985	5,808		41,312	
1986	5,959		38,264	
1987	6,106		43,479	
1988	6,253		49,747	
1989		7,782		68,041
1990		7,987		77,912
1991		8,208		87,474
1992		8,853		97,020
1993		8,737		110,770
1994		9,016		126,137
1995		9,304		143,164
1996		9,600		162,900
1997		9,906		179,133

1998	10,219	179,508
1999	10,545	187,052
2000	10,867	218,328
2001	11,406	217,426
2002	11,834	237,574
2003	12,266	262,174
2004	12,698	296,766
2005	13,132	328,181
2006	13,420	365,859
2007	13,707	411,383
2008	13,997	470,033
2009	14,286	433,553
2010	14,539	480,643
2011	15,047	538,333
2012	15,276	564,745

Figure 1 Shares in total income Malaysia

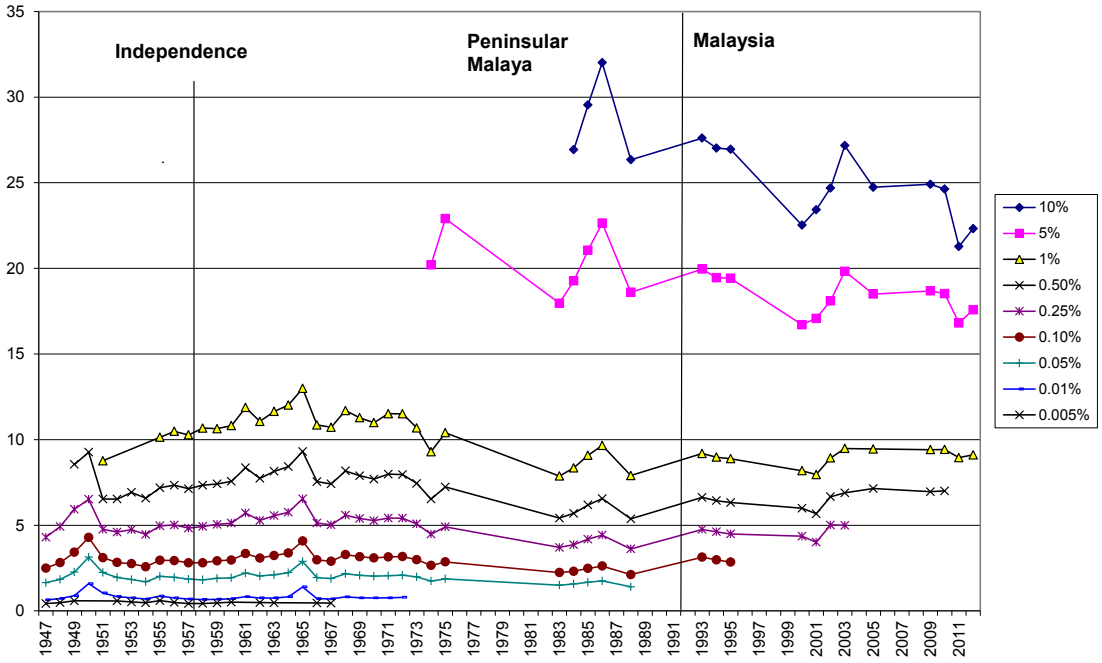

Figure 2 Inverse Pareto coefficients Malaysia