ISSN 1440-771X

MONASH University

Australia

Department of Econometrics
and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Inequality Trends and Determinants in Sri Lanka 1980-2002: A Shapley Approach to Decomposition

Ramani Gunatilaka and Duangkamon Chotikapanich

February 2006

Working Paper 6/06

Inequality Trends and Determinants in Sri Lanka 1980-2002:

A Shapley Approach to Decomposition

Ramani Gunatilaka *and* Duangkamon Chotikapanich

Department of Econometrics and Business Statistics
Monash University, Melbourne, Australia

February 20, 2005

Abstract

Sri Lanka liberalised its economy in 1977, paving the way for more rapid economic growth and higher rates of job creation. But tensions over distributional issues still plague the body politic. This paper investigates the evolution of Sri Lanka's income distribution in the period 1980-2002 and uses the Shapley value decomposition methodology to determine underlying causes. The study finds that while average incomes rose across strata, the rich experienced more rapid income growth leading to greater inequality. Inequality change was driven by differential access to infrastructure, education, and occupation status. Demographic factors including ethnicity, and spatial factors, contributed very little. The study recommends policies that ensure more equitable access to income earning assets such as education and infrastructure services and make sure that increases in inequality do not take place along sectoral, regional and ethnic fault lines.

Keywords: income inequality; Sri Lanka; Shapley value decomposition

JEL classification: D31

1. INTRODUCTION

In the mid-1970s, Sri Lanka had 'one of the most regulated economies outside the centrally planned economies' (Cuthbertson and Athukorala (1991) as cited in Athukorala and Jayasuriya (2000)). But inherent contradictions in the import substituting strategy resulted in a stagnant economy, high unemployment and the rationing of consumer goods. These economic ills were exacerbated by the first oil price shock of 1973 which paved the way for a landslide victory of the right-of-centre United National Party at the General Elections of 1977. The party had campaigned on a platform of liberalizing the economy and revitalizing the private sector. Thus, Sri Lanka became the first country in South Asia to liberalize its economy and dismantle the import substituting policy framework that had been in place since the late 1950s.

Economic liberalization was primarily aimed at generating export-oriented industrialization and a higher rate of economic growth and employment creation. In contrast to the preoccupation with social welfare issues of earlier governments, the government of the time was never directly concerned with the distributive effects of economic liberalization, looking instead to higher economic growth to translate into greater social welfare. As expected, economic growth rates picked up, light industrial goods began to claim a bigger share of the country's exports, and unemployment rates began to decline. A further, more intensive wave of reforms in the early 1990s left Sri Lanka one of the most open economies of the developing world by the end of the decade (Athukorala and Rajapatirana 2000).

Nevertheless, the post-liberalization era of higher economic growth and lower unemployment levels also saw continuing social conflict and greater political instability related to distributional issues. Sri Lanka's twin political conflicts of the early 1970s, the first involving ethnic Sinhalese youth and the second involving ethnic Tamil youth, stemmed from state policies that aimed to achieve redistributive justice in an economy made stagnant by a restrictive trade regime (Abeyratne 2004). The shortage of resources ensured that poverty rather than wealth was shared and masses of rural youth remained outside the economic mainstream. Thus, a fertile ground was created for the eruption and sustenance of social conflict.

Economic liberalization exacerbated rather than mitigated these tensions. Some economic liberalization measures negatively affected segments of the population engaged in the production of certain import substituting agricultural and industrial goods (Gunasinghe 1986). Other measures are likely to have increased returns to those with certain endowments of income generating assets such as education and urban residence (Dunham and Jayasuriya 2000; Lakshman 1997).

Encouraged by sub-continental geopolitical forces in the early 1980s, disaffected Tamil youths launched a violent insurrection against the state in a bid to secede. By the late 1980s Sinhalese youth in the rural south had also revolted. Thanks to the lack of external support for the latter movement, the government reasserted control with a mixture of brute force, more rapid economic reform, and an industrial and welfare policy that targeted areas outside the urban metropolis (Dunham and Kelegama 1997). But even today, the secessionist conflict in the north and east shows little signs of abating, the ceasefire notwithstanding, and twenty years of conflict have aggravated regional disparities. Meanwhile, the south remains politically volatile.

Thus, issues relating to income distribution continue to be of enormous economic and political significance in Sri Lanka. Critical questions remain about the evolution of income distribution and its underlying causes after economic liberalization, and a long-term analysis of these issues can contribute much to policy making. Nevertheless, the empirical evidence thus far has been piecemeal. Earlier studies that looked at the issue followed different methodologies and covered different periods (see Dunham and Jayasuriya 2000; Glewwe 1985; Glewwe 1986; Gunewardena 1996; Lakshman 1997). None extended beyond the mid-1990s. Analyses of the determinants of inequality (for example, Glewwe 1986; Gunewardena 1996), explained differences in income distribution between distinct groups of income recipients and between recipients with different sources of income. However, these methods did not permit the simultaneous decomposition of inequality by population subgroups and by income components. Nor did they enable the researchers to quantify the contributions of many other factors to total inequality.

In contrast, the present study applies a consistent, up-to-date methodology to examine the progress of inequality and its causes during the full sweep of 1980 to 2002. Thus, it is able to take advantage of the greater availability of data as well as of innovations in methodology that have emerged since the last rigorous assessment of inequality in Sri Lanka. In particular, it applies the Shapley value decomposition methodology to disaggregate total inequality and its changes into contributory factors.

This paper is organised as follows: Section 2 discusses the data used and presents the findings of the analysis of inequality trends in Sri Lanka over the reference period. Section 3 introduces the Shapley value decomposition methodology used to identify the causes of inequality and presents the results of the analysis. Section 4 concludes by drawing out the policy implications of the study's findings.

2. DATA AND INCOME DISTRIBUTION IN SRI LANKA

The analysis in this paper uses consumption expenditure as proxy for income. This is because consumption expenditure is a more accurate measure of individual and household welfare in developing countries (Deaton and Zaidi 2002). Large informal sectors made up of self-employment, small business and subsistence agriculture make the gathering of accurate income data difficult in developing countries, while means-tested income support programmes can encourage under-reporting of income. Moreover, consumption expenditure is a direct measure of individual and household welfare whereas income streams exhibit transitory fluctuations (Barrett et al. 2000).

Data on household expenditure and demographic and other characteristics of households were drawn from the 1980/81, 1985/86 Labour Force and Socio-Economic Surveys (LFSS) and the 1990/91, 1995/96 and 2002 Household Income and Expenditure Surveys (HIES) conducted by the Department of Census and Statistics, Sri Lanka. Despite the difference in name, the surveys are broadly comparable in design and methodology. However, the 1990/91, 1995/96 and 2002 HIES excluded the administrative districts in the Northern and Eastern Provinces as the conflict situation prevailing there precluded data collection in the region. Hence the present analysis relates only to the seven provinces outside the Northern and Eastern Provinces, which account for roughly 85 per cent of Sri Lanka's population of roughly 19 million people.

The analysis takes into account only households with positive expenditure. We also excluded households that box plot analyses of income data revealed as outliers. Household expenditure data were then adjusted to take into account equivalence scales, economies of scale, and temporal and spatial differences in the cost of living. Details of how we adjusted the data are as follows.

Individual expenditure was adjusted to take into account the different costs of children relative to adults and the economies of scale in consumption within a household by using equivalence scales. If household consumption expenditure is y_i and the adult equivalent size of the household is m_i, then the unit of analysis that we use is per adult equivalent consumption or y_i / m_i where m_i is calculated as follows:

$$m_i = \left(\phi_1 n_{a,i} + \phi_2 n_{c,i} \right)^{\theta}. \qquad (1)$$

In the formulation above, the number of adults is $n_{a,i}$ and the number of children is $n_{c,i}$. The term θ is a measure of economies of scale within the household and can take any value $0 \le \theta \le 1$. The term ϕ_1 is the cost of an adult member and Deaton and Zaidi (2002) who recommend this formula for the setting of equivalence scales in developing countries set its value as unity. The term ϕ_2 is the cost of a child relative to an adult and can take any value $0 \leq \phi_2 \leq 1$. In this paper we arbitrarily set ϕ_2 and θ as 0.6 and 0.9, respectively. We set out the reasons why we adopted these values in the Appendix.

The data on household expenditure were adjusted for temporal and spatial differences in the cost of living using the set of regional price indices for the five survey years developed by Gunatilaka (2005). The price indices were constructed by applying the Country Product Dummy (CPD) method developed by Summers (1973) and Rao (2003).

Major population surveys have many sampling units which have different probabilities of being selected and most analyses of survey data use these sampling weights to adjust the data. However, we have not used sampling weights in the present analysis because as Deaton (1997) points out, it would have biased the OLS estimates used for the regression-based decomposition analysis which is the major part of this paper.

Sri Lanka's income distribution over the period of 1980 to 2002 can be seen at a glance from Table 1 which sets out mean income and changes in mean income per quintile of income distribution for each survey year. There are three sections in the table. The first section gives the quintile mean income with standard errors in parentheses. These standard errors are calculated using the bootstrap method derived from Biewen (2002). The second and third sections give changes in mean income in absolute terms and as percentage changes. Standard errors for the estimates of changes in mean income and the z-statistics are derived using Biewen's (2000) methodology which assumes that the samples of any two years are independent. This is a valid assumption as the samples of the Sri Lankan household surveys differ from one year to the next.

It can be seen from Table 1 that mean incomes in Sri Lanka grew across quintiles over the reference period. All changes were significant at the 5 per cent level. But variations in the rates of income growth between quintiles are marked with average incomes in the higher quintiles growing faster than incomes in the lower quintiles in all sub-periods other than the period 1985-1990. Considering the period as a whole, mean income among the poorest quintile of income distribution grew by 50 per cent while the top most quintile experienced a substantially higher mean income growth rate of 88 per cent.

The impact of these changes on the distribution can be seen in the movements in the Lorenz curve in Figure 1. The Lorenz curve represents the functional relationship between the cumulative proportion of income and the cumulative proportion of income units, assuming that income units are arranged in ascending order of income.

Apart from a decline in inequality in 1990, the long-term trend has been for inequality to rise over the period. Although differences between the Lorenz curves are hard to discern, particularly at the ends of the distribution, the 2002 distribution appears furthest away from the line of equality. The 1990 distribution appears the most equal. In order to distinguish the curves in between and for a clearer representation of the extent to which the Lorenz curve has moved away from the line of equality over the reference period, we plot the transformed Lorenz curves in Figure 2 as differences between the line of equality and the associated Lorenz curves (see Deaton 1997). While the general shape of the transformed Lorenz curve remains the same for each distribution, the 1990 and 1980 distributions appear the most equal. The 1995 distribution is more unequal than the 1985 distribution, and the 2002 distribution is the most unequal.

Nevertheless, Lorenz curves only provide a partial ranking of distributions. When two Lorenz curves intersect like the 1980 and 1990 distributions, we cannot say which is more equal. In contrast, the Gini coefficient provides a complete ordering of distributions and is conveniently defined in terms of the Lorenz curve as one minus twice the area between the line of equality and the Lorenz curve. Table 2 presents the Gini coefficients for the five survey years calculated using the Lerman and Yitzhaki (1989) procedure based on the following formula,

$$G = \frac{2}{\bar{y}} \operatorname{cov}(y, \pi), \tag{2}$$

where $\operatorname{cov}(y, \pi)$ denotes the covariance between y and π, and $\bar{y}$ is the mean of y. Table 2 also presents bootstrapped standard errors derived using the same methodology as were the standard errors for mean income and changes presented in Table 1.

Changes in the Gini coefficient confirm that inequality rose between 1980 and 1985 and declined between 1985 and 1990. From 1990 onwards inequality rose steadily so that during the 22 year period, the Gini rose by some 14.6 per cent. The bootstrapped standard errors in the middle panel of the table show that the changes were significant at the 5 per cent significance level for a two-tailed test. The largest and most significant increase in inequality during the 1980-2002 period occurred between 1990 and 1995. Thus, the long-term trend has been for inequality to increase in Sri Lanka since economic liberalisation, despite a brief reversal around 1990.

We combine these two phenomena – rising mean incomes and rising inequality – and examine their impact on social welfare narrowly defined to include just these two components, by using generalized Lorenz curve analysis. The generalized Lorenz curve enables the comparison of different distributions with different means and thus different aggregates and is obtained by scaling up the Lorenz curve by its mean (Shorrocks 1983). Figure 3 plots the generalized Lorenz curves for the five survey years. It can be seen that by and large, the ranking of the distributions in terms of equality alone which we derived from the Lorenz curve analysis in Figure 1, has been almost totally reversed in terms of social welfare in the generalized Lorenz curves. The most unequal distribution, the 2002 distribution, emerges unambiguously as the most desirable in terms of welfare. The 1980 distribution, which vied with the 1990 distribution as being most equitable, emerges as the least desirable.

To sum up the findings on inequality trends, we note that growth rates of mean income by quintile revealed that all groups experienced consistent increases in mean income over the survey years. Income inequality rose between 1980 and 1985, declined between 1985 and 1990 and rose steadily thereafter. Generalized Lorenz curve analysis showed that due to the growth in mean incomes, welfare narrowly defined to include only notions of levels and spread, appears to have progressively increased in Sri Lanka with every survey year.

What factors gave rise to these far reaching changes in the dispersion of income in Sri Lanka? In the sections to follow we use the regression-based, Shapley value decomposition methodology to decompose inequality into its contributory factors.

3. REGRESSION-BASED DECOMPOSITION AND THE SHAPLEY APPROACH

Unlike traditional methods of decomposition by population subgroups and decomposition by income source, regression-based approaches have the advantage of enabling analysts to include any mix of explanatory factors including economic, social, demographic and policy variables. They also enable researchers to include continuous variables.

There have been many recent innovations in such methodologies, for example, Shorrocks (1999), Bourguignon et al. (2001), Morduch and Sicular (2002), and Fields (2003). All approaches begin with an income generating function, which, in linear form can be written as:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + \varepsilon_i , \qquad (3)$$

where y_i is the income variable for the *i*-th income unit. The variables x_{ki} represent exogenous household endowments of resources that determine its income. If y_i is per capita income, the explanatory variables could be land per capital, household size, workers per household and average education of adults (Morduch and Sicular 2002; Ravallion and Chen 1999). If y_i denotes wages, the explanatory variables could be gender, race, occupation, industry, or region (Fields 2003). Provided the income generating function expressed by Equation (4) gives accurate and meaningful estimates, $\hat{\beta}_k x_{ki}$ can be regarded as the share of household i's income that flows from its endowment of x_k. The term ε_i is a random error.

The results of the estimation of the income generating function are then used to quantify the contribution of any number of factors to total inequality. The Fields method, for example, manipulates the equation so that it can be written in terms of covariances. The contribution of the independent variables to distributional change is then expressed as a function of the size of the coefficients of the income equation and the magnitude of the change in the variable relative to the variation in income. In the Morduch and Sicular method, the resulting coefficients are regarded as estimates of the income flows attributed to household variables. This permits the application of decomposition by income source or factor income to apportion inequality to any number of explanatory variables. This is in contrast to the method proposed by Bourguignon et al. (2001), which can be used to decompose differences in income distribution into just three broad components: price effects, participation effects and population effects.

Nevertheless, the Morduch and Sicular method has been criticised on the basis that although the methodology requires the inclusion of an error term into the original income generating equation it does not make any contribution towards overall inequality (see Wan 2004, p. 352). In contrast, Fields' decomposition methodology accounts for the contribution of the regression error to total inequality, but this tends to be large, leaving unexplained the major proportion of inequality. Neither method accounts for the contribution of the constant term to total inequality.

3.1 Shapley Approach to Decomposition

In contrast to other regression-based methods, the Shapley value decomposition methodology circumvents the problem of a large residual and decomposes inequality completely into its contributory factors as it accounts for all parts of the income-generating equation (Shorrocks 1999). Starting with an income generating function, the method can be applied to decompose any inequality index using an income generating model of any functional form.

Shorrocks' (1999) general application of the Shapley value method to decompose income inequality derives from Shapley's (1953) solution to the problem of calculating the real power of any given voter in a coalition voting game with transferable utility, when all orders of coalition formation are equally probable. As Shorrocks (1999) puts it, the Shapley value decomposition procedure is a solution to the 'general decomposition problem' which yields 'an exact additive decomposition' of the inequality index into the contributory factors (Shorrocks 1999, p. 3). All factors are treated even-handedly and therefore the Shapley decomposition is symmetric in all variables.

While the Shapley value has been used in a number of cost allocation models since 1953 (see Albrecht et al. 2002), its application to decompose income is relatively recent. To our knowledge, Wan's (2004) paper on income inequality in rural China is the only published study which applies the Shapley value decomposition procedure to decompose income inequality. Almost all analyses using the methodology are still at the working paper stage and as many of them apply it to the decomposition of poverty as to the decomposition of income inequality. For example, Kolenikov and Shorrocks (2003), and D'Ambrosio et al. (2004) use the procedure to decompose poverty. Rongve (1995) and Chantreuil and Trannoy (1997) apply it to the decomposition of income. Devicienti (2003) also applies it to decompose wage income.

Shorrocks (1999) gives an extensive mathematical exposition of the Shapley decomposition procedure, particularly its application to poverty decomposition. D'Ambrosio et al. (2004) give a more succinct exposition. In what follows we take a pragmatic approach and set out the decomposition procedure step by step using a simplified model. In this example and in the application to follow we use the Gini coefficient to measure inequality. However, the procedure may be applied to decompose any inequality index.

To demonstrate the Shapley value decomposition, we begin by estimating a simple income generating model with only three explanatory variables.

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon \,. \qquad (4)$$

The dependent variable y_i represents the income variable. Both the income variable and the functional form can take any form but in this example we use a linear functional form. We use the results of the regression to obtain predicted income as follows,

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} \,. \qquad (5)$$

$\hat{y}_i$ is then used to calculate $\hat{G}_{TOT}$ which is total income inequality as calculated by the Gini index, in turn determined by the distribution of incomes attributable to x_1, x_2 and x_3.

The Shapley value decomposition aims to measure the extent by which $\hat{G}_{TOT}$ would change if income flows from x_{ki}, that is $\beta_k x_{ki}$, was removed from total income y_i. The change represents the contribution of income from x_{ki} to total inequality. In practice there are many possible sequences in which x_{ki} may be eliminated. Hence the contribution to inequality of income from x_{ki} is the average of all contributions of x_{ki} in all possible eliminating sequences. This makes sure that the estimated contribution of each variable to total inequality does not depend on the order in which it is eliminated.

Note that the regression error is excluded from calculating predicted income. The constant term β_0 does not contribute to inequality either. This is because if $y_i = \beta_0$, then it means that everybody is given the same income β_0, which cannot contribute anything towards inequality.

We now apply the concept of the Shapley decomposition to decompose inequality obtained from income defined in Equation (5). It should be noted that there are many rounds of variable elimination. The number of rounds is determined by the number of variables themselves. So if we are looking at a model with only three explanatory variables, then we have three rounds of elimination.

Round 1: Dropping one variable at a time

We remove the effect of one x at a time from Equation (5) as follows.

If we drop x_1 first, then predicted income is,

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} \,. \qquad (6)$$

We use $\hat{y}_i$ to calculate the Gini coefficient $G^1_{x_2x_3}$, where the superscript denotes round number (1) and the subscripts denote the variables included in the calculation of predicted income (x_2 and x_3).

Then the first round contribution of x_1 to inequality or its marginal effect is,

$$C^1_{x_1} = \hat{G}_{TOT} - G^1_{x_2x_3} . \tag{7}$$

Similarly, we drop x_2 and x_3 from the calculation of predicted income and calculate the first round contribution of x_2 to inequality as $C^1_{x_2} = \hat{G}_{TOT} - G^1_{x_1x_3}$ and the first round contribution of x_3 to inequality as $C^1_{x_3} = \hat{G}_{TOT} - G^1_{x_1x_2}$.

Round 2: Dropping two variables

Round two involves removing the effect of two $x's$ at a time from Equation (5). We then calculate the Gini coefficient and thereafter their contribution to total inequality. In the second round we get two values for x_1's contribution to total inequality. The combinations in which two $x's$ may be removed at a time are, for example, x_1 and x_k, $k = 2,3$. To illustrate, when x_1 and x_2 are dropped, predicted income is calculated with the remaining variable x_3 according to which the Gini coefficient $G^2_{x_3}$ is calculated. Then the first component of the second round contribution of x_1 to inequality is,

$$C^2_{x_1} = G^1_{x_1x_3} - G^2_{x_3} . \tag{8}$$

Similarly, we obtain the second component of the second round contribution of x_1 to inequality as $C_{x_1}^2 = G_{x_1x_2}^1 - G_{x_2}^2$. The average of the two different values of $C_{x_1}^2$ is x_1's contribution to total inequality in Round 2.

The same procedure is applied to the other variables. For example, x_2's contribution to total inequality in the second round is an average of two values of $C_{x_2}^2$ derived from $G_{x_2x_1}^1 - G_{x_1}^2$ and $G_{x_2x_3}^1 - G_{x_3}^2$.

The 'Final' Round

If we were to continue the logic of variable elimination as set out above, the final round would involve calculating the Gini coefficient from predicted income in Equation (5) from which all variables have been dropped. We would then need to subtract this value from the Gini coefficient calculated with predicted income from which all but the variable in question had been dropped, that is the Gini coefficients obtained in Round 2.

However, note that if all variables were to be dropped, predicted income would be $\hat{y}_i = \beta_0$ and the Gini coefficient would be zero. Hence, the contribution of each variable to total inequality in the final round would be equivalent to the Gini coefficient calculated with each in Round 2. So, for example, x_1's contribution to total inequality in the final round would be just one value, that is $C_{x_1}^3 = G_{x_1}^2$. Similarly, we derive the final round contributions to total inequality of x_2 and x_3 as $C_{x_2}^3 = G_{x_2}^2$ and $C_{x_3}^3 = G_{x_3}^2$.

Observe how each variable's marginal effects or contributions to inequality at each round of elimination are averaged across each round. For example, in our example, x_1's contribution to inequality in the first round would be just one value, but its contribution to inequality in the second round would be an average of two different values. Variable x_1's contribution to inequality in the final round would also be just one value. Finally, the average contributions of each round are averaged across all rounds to obtain the total marginal effect of each variable on total inequality.

The methodology also permits one to calculate the proportion of total inequality that is not explained by Equation (4), that is C_R as follows:

$$C_R = G - \hat{G}_{TOT}, \quad (9)$$

where G is the Gini calculated using actual income data and $\hat{G}_{TOT}$ is the Gini calculated using predicted income from all the explanatory variables, as defined in Equation (5).

In theory, the Shapley value decomposition may be used to decompose inequality into an unlimited number of contributory factors. But including a large number of variables in the model can be costly in terms of computing time, data processing and storage capacity. This is because the number of calculations increases exponentially with the number of variables. However, the methodology has the advantage that groups of factors may be treated as a single entity without affecting their total contribution (Shorrocks 1999).

The Shapley value decomposition results may also be used to calculate how much of the difference in income inequality between one period and the next is accounted for by the explanatory factors. The procedure only requires two comparable household surveys so that identical income generating equations can be estimated for each survey data set.

If the Gini coefficient, $G_{y_i t}$ is the inequality index to be decomposed and the contribution of the *k*'th endowment to total inequality in year *t* as averaged across all rounds is defined as C_{kt}, then the contribution of the k'th endowment to the change in total inequality between two years 1 and 2 is expressed as,

$$\gamma_k = \frac{C_{k2} G_{y_i 2} - C_{k1} G_{y_i 1}}{G_{y_i 2} - G_{y_i 1}}. \quad (10)$$

3.2 Variables

The dependent variable used in the regression-based decomposition analysis is $\ln(y_i)$ log of income. We use real per adult equivalent consumption expenditure as proxy for income. The independent variables $x_1, ..., x_k$, are factors that influence the level of household income. We categorise these factors into five groups. They are demographic, education, occupation, infrastructure and spatial characteristics of households. Although we include a large number of variables as defined below in the income generating equation, we decompose inequality into only seven components as we include or drop the ethnicity, education, occupation, infrastructure and spatial variables as groups rather than separately.

The regression model includes six demographic variables: *Adult males, share* is the proportion of adult male members of working age (more than 16 years of age) in a household; *Adult females, share* is the share of adult female members of working age in the household. These two variables are included to denote the household's potential income earning capacity. Demographic variables also include four ethnic dummies. They are, *Sri Lankan Tamil dummy, Indian Tamil dummy, Ethnic Moor dummy* and, *Ethnic Other dummy*. The omitted dummy variable is that for the majority Sinhalese ethnic group.

The impact of education attainment on income inequality is captured by the proportion of household members of working age who have completed four levels of education as follows: *Secondary education* denotes the proportion of household members of working age who have completed between 6 and 10 years of schooling; *GCE O'Level* denotes achievement of the General Certificate of Examination at Ordinary Level, which indicates success at the 10th year qualifying examination; *GCE A'Level* is the General Certificate of Examination at Advanced Level denoting success at the 12th year qualifying examination; *University, share* denotes the proportion of household workers who have achieved tertiary education and more. Since primary education is virtually universal in Sri Lanka we have not included it in the analysis.

We define five variables for occupation and employment status. They denote the proportion of household members of working age in the following five categories. *Unemployed* denotes the share of the household's labour force participants who have been unemployed and looking for work in the week prior to being enumerated for the survey. *Managers* includes professional and executive positions. The remaining variables are *Clerical, Service sector workers* and *Farmers.* The definitions for managers, clerks and service sector workers in the 2002 survey are specified as salaried employees whereas the earlier surveys had broader definitions which included self-employed workers. The excluded category refers to product workers and those not classified elsewhere.

Three dummy variables in the model capture households' access to infrastructure: *Vehicle* is included to denote mobility and access to roads that may be travelled on by means of bicycle, scooter, motorbike, car or van. The other two variables are *Electricity* and *Telephone.* Explicit information on access to all three facilities is available only from the 1980 survey, and for telephones only in the 1990 and 1995 surveys. Information about access to vehicles and electricity for all survey years other than for 1980 and for telephone connectivity for 1985 and 2002 has been inferred by means of data on expenditure: any expenditure on these items or ancillaries such as vehicle spare parts or licenses within the survey reference period has been adduced as denoting access or ownership.

Spatial variables included in the model are geographic dummy variables at provincial level. Provincial dummies are meaningful units of analysis at sub-national level and are used instead of rural-urban dummies because of the unsatisfactory nature of rural-urban definitions in the surveys. The provincial dummies are as follows: *Central Province, Southern Province, North Western Province, North Central Province, Uva Province, Sabaragamuwa province.* The omitted geographic dummy here is Western Province.

3.3 Results

Table 3 sets out the regression coefficients for all five survey years, estimated using *Stata*. The regressions appear to perform well. R^2 varies from 0.34 to 0.48, which is reasonable for cross-sectional regressions of this sort. Almost all of the included regressors are significant and have the expected signs. We discuss them under each group of variables in what follows.

As far as demographics are concerned, the results show that per adult equivalent income increases with the proportion of adults in the household, more so for males than for females other than in 1995. The coefficients are significant at the 1 per cent level.

Coefficients of the ethnic dummies are interesting. They show that Sri Lankan Tamils are better-off on average than the Sinhalese but the results are significant only for 1985 and 1995. It should be recalled that the data relates to the provinces outside the conflict areas of the north and the east where at least two thirds of the Sri Lankan Tamil population reside.

The fact that Sri Lankan Tamils in the south are better off on average than the Sinhalese could be due to a mix of factors: Sri Lankan Tamils living in the south are generally the more skilled among their community and foreign remittances to this group from the Tamil Diaspora living abroad have been substantial since 1983.

Indian Tamils, too, are on average better off than the majority Sinhalese community. The positive sign and significance of the coefficient on the 'Ethnic other' dummy other than for 1980 reveals that on average they, too, are better off than the majority Sinhalese. This is to be expected since they include small groups such as the Borahs and the Sindhis who are mercantile communities and well-represented in the business sector, as well as the Burgher community, descendants of Portuguese and Dutch colonialists and settlers. In contrast, ethnic Moors have been consistently worse off than the Sinhalese.

Education increases income and increasingly so for higher levels. Over the years, however, returns to the proportion of household workers with only secondary education have declined somewhat. Even so, returns to the proportion of workers educated up to GCE A' Levels and more have increased. Unemployment has a decreasing, significant effect on income other than in 1985, when, surprisingly, it is associated with higher income. More skilled occupation status is associated with higher income and the results are significant. For example, the share of managers in a household is associated with a larger coefficient than the share of service sector workers, who in turn are associated with higher income than households with a larger share of adult members engaged in farming.

All three infrastructure-related dummies have positive and significant coefficients, denoting that ownership of vehicles, access to electricity and telephones are associated with higher income.

The possible presence of endogeneity cautions against giving too much weight to the precise magnitudes of the contributions of these variables to consumption, particularly in the case of infrastructure. Even so, it is well known that efficient transport, electricity and telecommunications services raise productivity levels, integrate markets and enable factors of production to flow to areas and sectors with the best possible returns. Hence it is plausible that such factors raise the income-generating capacity of individuals who have access to them.

All regional dummies other than for North Western Province for 1985, North Central Province for 1985, 1990, 1995 and 2002, and Sabaragamuwa Province for 1990, are negative and significant, denoting that on average, residents living in these provinces are worse off than those living in Western Province. This is to be expected as Western Province with its metropolitan hub of Colombo, is the most favoured region both in terms of infrastructure development and economic activity. The positive coefficients on the North Central Province dummy for 1980, 1995 and 2002 are puzzling, for, as in all other provinces, one would expect income in this province for these years to be negative relative to average income in Western Province. Nevertheless, other than for 1980, the results are not significant.

The regression results are used next to calculate inequality weights using the Shapley procedure in order to decompose the level of inequality in each survey year by each factor. Predicted log of income was converted to real income before calculating the Gini coefficient at each stage. Also recall that inequality is decomposed into only seven components as we group the ethnicity, education, occupation, infrastructure and spatial variables and include or drop them as groups rather than separately.

The results of the Shapley decomposition of the Gini index in terms of the percentage share of total inequality explained by each factor are presented in Table 4.[1] The

[1] Estimation was carried out using *Stata*. Kolenikov's (Kolenikov 2000) Shapley.ado programme was

contribution of the Gini coefficient $\hat{G}_{TOT}$ of the saturated model (or the model that includes all explanatory variables in its calculation of predicted income) to total inequality as measured by the Gini coefficient calculated on original income data, is also set out. Note that the five values of G decomposed using the Shapley value method are those presented in Table 2. The residual or the proportion of total inequality that remains unexplained by the variables included in the model is $G - \hat{G}_{TOT}$. We present the contributions to total inequality of $\hat{G}_{TOT}$ and the residual in order to indicate the extent to which the explanatory variables together account for total inequality.

While income flows from all groups of endowments appear to have contributed positively to increase inequality, income flows from access to infrastructure accounted for the largest share of total inequality. Its contribution to total inequality increased progressively over the years from 19 per cent in 1980 to 37 per cent of inequality in 2002. This is followed by education which accounts for between 17 and 21 per cent. No clear trend is visible. Income flows from occupation endowments contributed between 7 and 10 per cent over the period, but its contribution to inequality appears to be declining slightly. The contribution to inequality from demographic characteristics has been roughly stable until 1995 at around 8 per cent, after which it has reduced drastically to less than 3 per cent in 2002. The contribution of spatial factors has decreased during the 1980s and increased thereafter. The contribution of unidentified factors (denoted by the residual) has declined from roughly 40 per cent in 1980 to 30 per cent in 2002.

---

extensively modified and adapted for this purpose.

Table 5 sets out the factor contributions to the change in inequality by substituting the Shapley values for the factor inequality weights C_{kt} in Equation (10). Recall that the change in inequality decomposed here is the change in the Gini coefficient calculated using original income and it is apparent that the Gini rose consistently during all sub-periods other than the period 1985-90. Also note that a positive contribution denotes that the factor or group of factors acted to intensify the change in inequality as measured by the Gini coefficient, in whichever direction the change in Gini took place.

In terms of groups of factors, infrastructure and education helped increase inequality in the three sub-periods when inequality rose. They also helped decrease inequality in the period 1985-90. Occupation helped increase inequality in the first half of each decade but helped reduce inequality in every other period. It is apparent that the main drivers of inequality increase over the period have been infrastructure and education: education consistently so, with a marked spurt towards the end of the period; infrastructure more spectacularly, again in the last sub-period. Note how the contribution to inequality of income from ethnic endowments has reversed gear – from pushing the increase in inequality in the 1980s, to mitigating it, however weakly, from then onwards. Other demographic variables mimic the pattern of occupation's contribution to inequality change. Note in particular how they have restrained the rise in inequality in 1995-2002. Income from spatial endowments reduced the rise in inequality only in the 1980s, but contributed positively towards its rise thereafter.

Thus, the Shapley value decomposition shows that income flows associated with access to infrastructure, education and occupation were the principal determinants of inequality and the main drivers of the change in income dispersion in Sri Lanka during the reference period. In fact, the contribution of income flows from education and infrastructure to total inequality has increased over the years. The contribution from income flows associated with occupation endowments has declined slightly. In contrast, demographic factors, including income flows from ethnicity, contributed relatively little. Spatial factors account for about as much of inequality as does income from demographic factors but their contribution has been increasing in the recent past whereas the contribution of demographic factors has been declining.

4. CONCLUSIONS AND POLICY IMPLICATIONS

This study showed that income inequality in Sri Lanka rose after economic liberalisation even as all income groups ended the period enjoying higher levels of income than they had at its beginning. The principal determinants of inequality change were access to education, occupation and infrastructure.

The findings suggest a growth-equity trade off: economic liberalisation and concomitant economic growth has caused inequality to rise in a stereotypical Kuznets-like phenomenon. Forces of economic growth unleashed by economic liberalisation appear to have increased returns to certain income-earning assets of households such as education and infrastructure. The ability to access such assets has influenced household income and in turn, the dispersion of household income.

Thus, the macroeconomic policy framework that favoured trade liberalisation and economic growth, and sectoral policies investing in education and infrastructure services, have caused incomes to rise across the board, though proportionately more for the higher income ranges. The Sri Lankan case appears to be a clear example of how these policies complement each other and lead to greater income prosperity.

However, inequality also increased. While the government got its basic policy mix right and everybody could reap the benefits of greater income growth, the rise in inequality in the medium term has been an unfortunate by-product of the process. Semi-parametric decomposition analysis concentrating on the impact of changes in these variables on the entire distribution suggests that this is because the middle classes benefited disproportionately more from the provision of education and infrastructure services than did the poor (see Gunatilaka 2005).

The analysis suggests that the government should better target services towards those at the bottom of the income distribution. The growth in income in the lowest quintile notwithstanding, income receivers in this segment remain unacceptably poor: 23 per cent of the population were found to be below the official poverty line in 2002 (Narayan and Yoshida 2004).

Hence, the government urgently needs to undertake pro-poor investments in education and infrastructure provision. But this needs a well-calibrated approach. For example, it is not enough to merely build roads in rural areas where many of the poor reside. The government will also need to provide a public transport service where it is unviable for private transport service providers to do so. Restoring the rural transport services which became the casualties of an ideologically-driven, ill-planned privatisation of the system, and enabling the integration of communities with markets would be a key policy initiative that would enable those at the bottom end of the income distribution to benefit from the economic liberalisation process.

While pro-poor investments would enable those in the lowest income strata to move up along the income distribution and out of poverty, they may not succeed in narrowing income gaps: trade liberalisation and technological change can increase returns to some factors and exacerbate income differentials. The more urgent issue appears to be to contain the social tensions that would inevitably arise with increasing inequality, a flashpoint for social conflict in diverse societies like Sri Lanka's. Policies need to be designed and targeted in such a way that the stresses generated by rising inequality do not occur along sectoral, regional and ethnic fault lines. In particular, the government needs to be vigilant and design and implement policies that mitigate the rise of inequality between sectors, between regions and between ethnic groups; that is, between any population groupings that can mobilise along any of these attributes, cite that attribute as being a cause for discrimination or neglect, and mount a violent protest against the rest of society and the state.

APPENDIX

Estimation of Adult Equivalence Scales

Equation (1) in the main text sets out the formula for the adult equivalent scale m_i. We reproduce it here as follows.

$$m_i = \left(\phi_1 n_{a,i} + \phi_2 n_{c,i}\right)^{\theta}. \qquad \text{(A1)}$$

The number of adults is $n_{a,i}$ and the number of children is $n_{c,i}$. The term θ is a measure of economies of scale within the household. The term ϕ_1 is the cost of an adult member and ϕ_2 is the cost of a child relative to an adult.

The setting of parameters for equivalence scales in the developing country context tends to be arbitrary. Deaton and Zaidi (2002) suggest setting ϕ_2 equivalent to 0.25 or 0.33 - that is, expenditure on a child amounting to either a fourth or a third of the cost of an adult. These recommendations have been based on the results of an application of Rothbarth's procedure for measuring child costs to Sri Lankan and Indonesian data of the late 1960s and early 1970s (Deaton and Muellbauer 1986). Deaton and Zaidi (2002) recommend setting θ high, perhaps at 0.9.

However, it is likely that a child cost ratio of 0.25 or 0.33 may be too low for Sri Lanka in the 1980s and 1990s. First, while children cost less in poor, agricultural economies, agriculture accounted for a smaller share of GDP in Sri Lanka in the 1980s and 1990s than it did twenty years earlier, while the contribution of the services sector has increased significantly. Secondly, certain child specific costs such as expenditure on milk food are likely to have increased significantly since almost all milk products are imported and the value of the Sri Lankan rupee has been declining steadily over the last twenty years. Thirdly, even though the government provides education free of charge, quality varies and parents' aspirations for their children's education have risen. As a result, parents probably spent an increasing proportion of income on private tuition, equipment and materials during the reference period than they did earlier. Fourth, family size has declined. For example, average household size has declined from 4.9 members in 1980/81 to 4.2 members in 2002, with the result that families probably have fewer children and spend more on each child than they did earlier. Scope for economies of scale may have also increased over the years as the budget share on food has declined. For example, while households spent an average of 0.73 of total expenditure on food in 1980, by 2002 this proportion had dropped to 0.53.

Given the arbitrary nature of setting equivalence scales we tested the sensitivity of the Gini coefficient to different values of ϕ_2 and θ, including those suggested by Deaton and Zaidi (2002). We found that for a given year, there was little significant difference between the estimates (see Gunatilaka 2005). Hence for the purpose of the current research we determined to set $\phi_2 = 0.60$ as being a more realistic estimation of child costs in Sri Lanka during the last two decades. This is at least twice as high as that recommended by Deaton et al. (2002), but we believe it is justified given the reasons advanced above. We set θ=0.9, recognising that scope for economies of scale may have been limited in the 1980s, and that a lower θ makes little difference to the Gini coefficient.

REFERENCES

Abeyratne, S. (2004). Economic Roots of Political Conflict: The Case of Sri Lanka. *World Economy*, 27, 1295-1314.

Albrecht, J., Francois, D., & Schoors, K. (2002). A Shapley decomposition of carbon emissions without residuals. *Energy Policy*, 30, 727-736.

Athukorala, P., & Jayasuriya, S. (2000). Trade Policy Reforms in Sri Lanka. *The World Economy*, 23.

Athukorala, P., & Rajapatirana, S. (2000). Liberalization and Industrial Transformation: Lessons from the Sri Lankan Experience. *Economic Development and Cultural Change*, 48, 543-572.

Barrett, G. F., Crossley, T. F., & Worswick, C. (2000). Consumption and Income Inequality in Australia. *The Economic Record*, 76, 116-138.

Biewen, M. (2000). Income Inequality in Germany During the 1980s and 1990s. *Review of Income and Wealth*, 46, 1-19.

Biewen, M. (2002). Bootstrap inference for inequality, mobility and poverty measurement. *Journal of Econometrics*, 108, 317-342.

Bourguignon, F., Fournier, M., & Gurgand, M. (2001). Fast Development With a Stable Income Distribution: Taiwan, 1979-94. *Rev Income Wealth*, 47, 139-163.

Chantreuil, F., & Trannoy, A. (1997). Inequality decomposition values. Unpublished mimeograph, Universite de Cergy-Pontoise, Cergy-Pontoise.

Cuthbertson, A. G., & Athukorala, P. (1991). Sri Lanka. In D. Papageorgiou, M. Michaely & A. M. Choksi (Eds.), *Liberalizing Foreign Trade: Indonesia, Pakistan and Sri Lanka* (pp. 287-414). Oxford: Basil Blackwell.

D'Ambrosio, C., Deutsch, J., & Silber, J. (2004). Multidimensional Approaches to Poverty Measurement: An Empirical Analysis of Poverty in Belgium, France, Germany, Italy and Spain, based on the European Panel. 28th General Conference of The International Association for Research in Income and Wealth, Cork, Ireland. August 22 – 28, 2004.

Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Washington D.C.: The World Bank.

Deaton, A., & Muellbauer, J. (1986). On Measuring Child Costs: With Applications to Poor Countries. *Journal of Political Economy*, 94, 720-744.

Deaton, A., & Zaidi, S. (2002). *Guidelines for Constructing Consumption Aggregates for Welfare Analysis*. LSMS Working Paper, 135. Washington D.C.: The World Bank.

Devicienti, F. (2003). *The Rise in Wage Inequality in Italy: Observable Factors or Unobservables?*2004. Torino: LABORatorio R. Revelli and ISER, University of Essex.

Dunham, D., & Jayasuriya, S. (2000). Equity, Growth and Insurrection: Liberalisation and the Welfare Debate in Contemporary Sri Lanka. *Oxford Development Studies*, 28, 97-110.

Dunham, D., & Kelegama, S. (1997). Does Leadership Matter in the Economic Reform Process? Liberalization and governance in Sri Lanka. *World Development*, 25, 179-190.

Fields, G. S. (2003). Accounting for Income Inequality and Its Changes: A New Method with Application to the Distribution of Earnings in the United States. *Research in Labor Economics*, 22, 1-38.

Glewwe, P. (1985). An analysis of income distribution and labor markets in Sri Lanka. Unpublished PhD, Stanford University.

Glewwe, P. (1986). The Distribution of Income in Sri Lanka in 1969-70 and 1980-81: A Decomposition Analysis. *Journal of Development Economics*, 24, 255-274.

Gunasinghe, N. (1986). Open Economic Policy and Peasant Production. *Upanathi*, 1, 37-67.

Gunatilaka, R. (2005). Income Distribution After Economic Liberalisation: Sri Lanka 1980-2002. Unpublished PhD, Monash University.

Gunewardena, D. (1996). An Analysis of Income Inequality in a Developing Country: Sri Lanka, 1985-1986 and 1990-1991. Unpublished PhD, American University.

Kolenikov, S. (2000). SHAPLEY: Stata module to perform additive decomposition of sample statistic. 24 April 2004. http://fmwww.bc.edu/repec/bocode/s/shapley.ado.

Kolenikov, S., & Shorrocks, A. F. (2003). *A Decomposition Analysis of Regional Poverty in Russia*. Discussion Paper, No. 2003/74. Helsinki: WIDER.

Lakshman, W. D. (1997). Income Distribution and Poverty. In W. D. Lakshman (Ed.), *Dilemmas of Development* (pp. 171-221). Colombo: Sri Lanka Association of Economists (SLAE).

Lerman, R. I., & Yitzhaki, S. (1989). Improving the Accuracy of Estimates of Gini Coefficients. *Journal of Econometrics*, 42, 43-47.

Morduch, J., & Sicular, T. (2002). Rethinking Inequality Decomposition, With Evidence From Rural China. *Economic Journal*, 112, 93-106.

Narayan, A., & Yoshida, N. (2004). Poverty in Sri Lanka: the Impact of Growth with Rising Inequality. Unpublished mimeograph, World Bank, Washington D.C.

Rao, D. S. P. (2003). PPPs for the Measurement of Global and Regional Poverty: Issues and Options. Unpublished mimeograph, Development Economics Data Group World Bank,

Ravallion, M., & Chen, S. (1999). When Economic Reform is Faster than Statistical Reform: Measuring and Explaining Income Inequality in Rural China. *Oxford Bulletin of Economics and Statistics*, 61, 33-56.

Rongve, I. (1995). *A Shapley decomposition of inequality indices by income source*. Discussion Paper, 59. Regina: Department of Economics, University of Regina.

Shapley, L. (1953). A value for n-person games. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the theory of games*. Princeton: Princeton University.

Shorrocks, A. F. (1983). Ranking Income Distributions. *Economica*, 50, 3-7.

Shorrocks, A. F. (1999). Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value. Unpublished mimeograph, Department of Economics, University of Essex.

Summers, R. (1973). International Price Comparisons Based Upon Incomplete Data. *Review of Income and Wealth*, 19, 1-16.

Wan, G. (2004). Accounting for Income Inequality in Rural China: A Regression Based Approach. *Journal of Comparative Economics*, 32, 348-363.

Table 1: Mean Income in Sri Lanka 1980-2002 (Rupees)

Quintile of Income Distribution	1980	1985	1990	1995	2002
1	147.13	154.98	185.69	199.99	220.53
	(1.651)	(0.880)	(1.025)	(1.086)	(1.357)
2	214.22	228.48	269.82	288.74	330.12
	(1.973)	(1.072)	(1.208)	(1.171)	(1.716)
3	271.40	296.01	344.46	371.69	441.80
	(2.196)	(1.344)	(1.531)	(1.686)	(2.489)
4	351.41	390.90	450.02	500.66	616.24
	(3.507)	(1.951)	(2.231)	(2.682)	(4.157)
5	666.17	752.49	841.44	1009.21	1252.16
	(14.463)	(8.372)	(7.856)	(10.185)	(13.103)
	Change in Mean Income				
	1985-1980 [z-stat]	1990-1985 [z-stat]	1995-1990 [z-stat]	2002-1995 [z-stat]	2002-1980 [z-stat]
1	7.80	30.70	14.30	20.51	73.31
	[4.170]	[22.731]	[9.576]	[11.799]	[34.304]
2	14.19	41.33	18.94	41.35	115.82
	[6.322]	[25.594]	[11.257]	[19.903]	[44.294]
3	24.44	48.57	27.16	70.10	170.26
	[9.492]	[23.835]	[11.923]	[23.314]	[51.293]
4	39.17	59.32	50.52	115.61	264.61
	[9.761]	[20.015]	[14.481]	[23.369]	[48.657]
5	86.31	88.75	167.79	242.82	585.67
	[5.165]	[7.730]	[13.045]	[14.631]	[30.010]
	Percentage Change in Mean Income				
	1985-1980	1990-1985	1995-1990	2002-1995	2002-1980
1	5.3	19.8	7.7	10.3	49.9
2	6.7	18.1	7.0	14.3	54.1
3	9.1	16.4	7.9	18.9	62.8
4	11.2	15.1	11.3	23.1	75.4
5	13.0	11.8	19.9	24.1	88.0

Notes: Standard errors in round parentheses, z-statistics in square parentheses, derived from 1000 bootstrap samples. The 5 per cent critical values for the z-statistic are as follows: 1.96 for a two-tailed test; 1.65 for a one-tailed test.

Table 2

Income Inequality and Changes in Sri Lanka 1980-2002

Gini Coefficient				
1980	1985	1990	1995	2002
0.310	0.323	0.309	0.336	0.355
(0.006)	(0.003)	(0.002)	(0.002	(0.003)
Change in Gini Coefficient				
1985-1980 [z-stat]	1990-1985 [z-stat]	1995-1990 [z-stat]	2002-1995 [z-stat]	2002-1980 [z-stat]
0.014	-0.015	0.027	0.019	0.045
[2.222]	[-4.004]	[7.889]	[5.446]	[7.366]
Percentage Change in Gini Coefficient				
1985-1980	1990-1985	1995-1990	2002-1995	2002-1980
4.5	-4.5	8.6	5.7	14.6

Notes: Gini coefficient calculated using as the unit of analysis adult equivalent income. Standard errors in round parentheses, z-statistics in square parentheses, derived from 1000 bootstrap samples. The 5 per cent critical values for the z-statistic are as follows: 1.96 for a two-tailed test; 1.65 for a one-tailed test.

Table 3: Regression Results

Dependent variable: log of income per adult equivalent

	1980	1985	1990	1995	2002
Demography					
Adult males, share	0.345**	0.3719**	0.2982**	0.3385**	0.1724**
	(0.0318)	(0.0160)	(0.0168)	(0.0171)	(0.0185)
Adult females, share	0.2231**	0.2939**	0.2857**	0.3385	0.1209**
	(0.0358)	(0.0171)	(0.0174)	(0.0178)	(0.0194)
Sri Lankan Tamil dummy	0.0032	0.0817**	0.0212	0.0695**	0.0279
	(0.0293)	(0.0158)	(0.0149)	(0.0137)	(0.0169)
Indian Tamil dummy	0.08**	0.2284**	0.23**	0.1427**	0.0669**
	(0.0277)	(0.0166)	(0.0176)	(0.0150)	(0.0171)
Ethnic Moor dummy	-0.0083	-0.0632**	-0.0886**	-0.1001**	-0.0333*
	(0.0331)	(0.0136)	(0.0135)	(0.0156)	(0.0148)
Ethnic other dummy	0.12	0.1555**	0.2089**	0.1992**	0.1659**
	(0.0696)	(0.0359)	(0.0346)	(0.0458)	(0.0468)
Education					
Secondary education, share	0.2134**	0.194**	0.1647**	0.1088**	0.1951**
	(0.0211)	(0.0104)	(0.0111)	(0.0114)	(0.0121)
GCE O'Level, share	0.4421**	0.4688**	0.3786**	0.3448**	0.4067**
	(0.0308)	(0.0148)	(0.0143)	(0.0145)	(0.0152)
GCE A'Level, share	0.5598**	0.5862**	0.5289**	0.5471**	0.6001**
	(0.0664)	(0.0296)	(0.0270)	(0.0226)	(0.0197)
University, share	0.7508**	0.7127**	0.8012**	0.7729**	0.8863**
	(0.1201)	(0.0460)	(0.0500)	(0.0470)	(0.0385)
Occupation					
Unemployed, share	-0.2318**	0.0489**	-0.383**	-0.3997**	-0.2986**
	(0.0322)	(0.0144)	(0.0223)	(0.0235)	(0.0199)
Managers, share	0.3384**	0.5254**	0.3692**	0.4232**	0.3165**
	(0.0651)	(0.0298)	(0.0303)	(0.0299)	(0.0242)
Clerical, share	0.4308**	0.4874**	0.3062**	0.3426**	0.2617**
	(0.0700)	(0.0295)	(0.0308)	(0.0314)	(0.0359)
Service workers, share	0.2847**	0.399**	0.2544**	0.2747**	0.0427
	(0.0428)	(0.0193)	(0.0194)	(0.0198)	(0.0369)
Farmers, share	0.0381	0.2137**	0.213**	0.14**	0.0708**
	(0.0298)	(0.0136)	(0.0128)	(0.0122)	(0.0135)

Table 3: Regression Results (contd.)

	1980	1985	1990	1995	2002
Infrastructure					
Vehicle, dummy	0.3698**	0.3306**	0.3699**	0.3587**	0.2818**
	(0.0282)	(0.0122)	(0.0107)	(0.0105)	(0.0103)
Electricity dummy	0.2507**	0.2588**	0.2251**	0.2165**	0.1964**
	(0.0213)	(0.0090)	(0.0084)	(0.0081)	(0.0083)
Telephone dummy	0.3851**	0.474**	0.2979**	0.3761**	0.3983**
	(0.0701)	(0.0316)	(0.0222)	(0.0815)	(0.0097)
Spatial					
Central Province	-0.1407**	-0.0843**	-0.0551**	-0.085**	-0.0929**
	(0.0204)	(0.0105)	(0.0105)	(0.0108)	(0.0108)
Southern Province	-0.1138**	-0.0952**	-0.026*	-0.0635**	-0.1189**
	(0.0209)	(0.0103)	(0.0104)	(0.0111)	(0.0116)
North Western Province	-0.0437*	-0.0206	-0.0598**	-0.0753**	-0.1129**
	(0.0213)	(0.0115)	(0.0119)	(0.0122)	(0.0123)
North Central Province	0.1005**	-0.0052	-0.0082	0.0283	0.0099
	(0.0322)	(0.0133)	(0.0135)	(0.0147)	(0.0151)
Uva Province	-0.2339**	-0.0573**	-0.0474**	-0.1651**	-0.1296**
	(0.0282)	(0.0128)	(0.0132)	(0.0135)	(0.0147)
Sabaragamuwa Province	-0.0409	-0.114**	-0.0209	-0.0801**	-0.1472**
	(0.0225)	(0.0115)	(0.0117)	(0.0119)	(0.0124)
Constant	5.3539**	5.1237**	5.3565**	5.4404**	5.6339**
	(0.0351)	(0.0171)	(0.0160)	(0.0172)	(0.0191)
Observations	4514	19470	18459	19747	16922
R-squared	0.37	0.38	0.34	0.36	0.48

Notes: Standard errors in parentheses. * significant at 5%; ** significant at 1%

Table 4

Factor Contribution to Level of Inequality Using Shapley Value Decomposition (%)

	1980	1985	1990	1995	2002
Adult males	4.90	4.72	3.44	3.39	1.34
Adult females	2.06	3.08	3.03	3.65	1.07
Ethnicity	0.31	1.37	1.83	0.88	0.28
Education	17.96	18.49	16.88	17.33	20.99
Occupation	9.89	10.13	8.73	8.73	6.70
Infrastructure	19.44	22.68	23.50	22.72	36.59
Spatial	6.38	2.76	1.27	3.73	4.38
$\hat{G}_{TOT}$	60.94	63.23	58.68	60.43	71.36
Residual	39.06	36.77	41.32	39.57	28.64
G	100.00	100.00	100.00	100.00	100.00

Table 5

Factor Contribution to Changes in Inequality Using Shapley Value Decomposition (%)

	1985 - 1980	1990-1985	1995-1990	2002-1995
Adult males	1.82	12.41	2.05	-11.83
Adult females	10.58	2.42	4.76	-15.33
Ethnicity	9.71	-2.40	-2.98	-3.52
Education	17.19	23.95	13.15	37.74
Occupation	9.05	16.91	5.80	-7.88
Infrastructure	42.56	9.04	12.04	112.38
Spatial	-28.06	12.63	11.75	7.20
Residual	6.50	-8.64	19.54	-47.62
Change in G	3.74	-4.08	8.17	5.39

Figure 1: Lorenz Curves 1980-2002

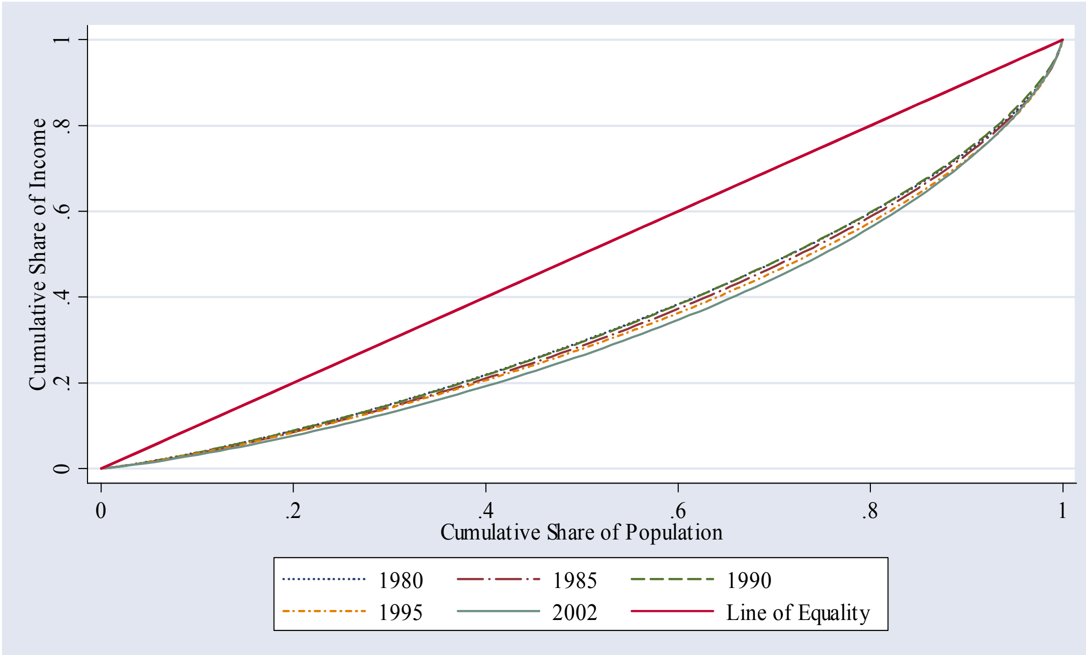

Figure 2: Transformed Lorenz Curves 1980-2002

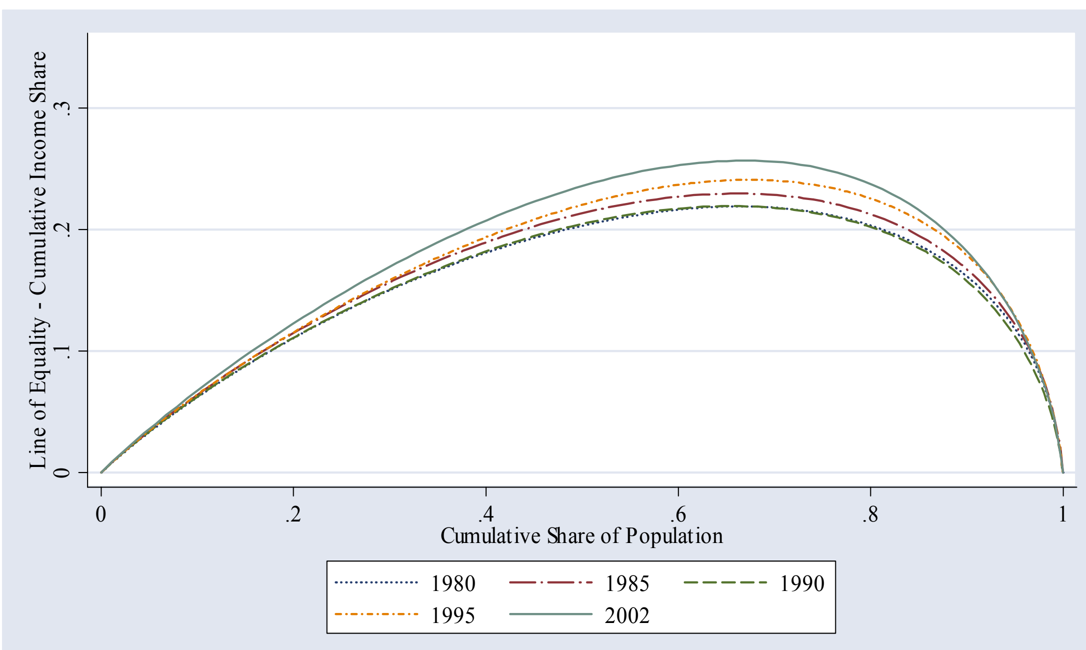

Figure 3: Generalized Lorenz Curves 1980-2002