MONASH University

AUSTRALIA

Department of Econometrics
and Business Statistics

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

Impact of Structural Change in Education, Industry and Infrastructure on Income Distribution in Sri Lanka

Ramani Gunatilaka, Duangkamon Chotikapanich and Brett Inder

NOVEMBER 2006

WORKING PAPER 21/06

Impact of Structural Change in Education, Industry and Infrastructure on Income Distribution in Sri Lanka

Ramani Gunatilaka, Duangkamon Chotikapanich *and* Brett Inder

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

December, 2006

Abstract

Income inequality increased in Sri Lanka following trade liberalization in 1977. This study applies a semi-parametric method to investigate whether structural changes in education, industry and infrastructure access underlay the change in the distribution. The study finds that while the concentration of people shifted towards higher income ranges at every stage in the distribution between 1985 and 2002, changes in access to infrastructure triggered much of the shift. Higher levels of educational attainment also had an impact. But the middle classes appear to have benefited disproportionately more from the provision of education and infrastructure services than did the poor. The analysis recommends that such services are targeted more effectively towards those in the poorest income deciles to enable them to move out of poverty to higher income ranges.

Keywords: income inequality; Sri Lanka; education; infrastructure; kernel density decomposition

JEL classification: D31

1. INTRODUCTION

Income inequality in Sri Lanka rose after economic liberalization in 1977. The Gini coefficient rose significantly from 0.31 in 1980 to 0.35 in 2002 (Gunatilaka and Chotikapanich 2006). Summary measures of the contributions of individual causal factors derived from regression-based decomposition methodology suggest that differential access to infrastructure, education and occupation drove the changes. Demographic and spatial factors contributed relatively little (ibid.). However, from the point of view of formulating policy and targeting interventions, it would be useful to find out where exactly on the distribution of income the more important policy-related variables exerted the greatest impact.

This investigation fills a key gap in the literature for two reasons. To our knowledge, this is the first to look at the impact of the changing structure of education, industry and infrastructure on the distribution of income in Sri Lanka. Earlier studies identified the determinants of inequality (Glewwe 1985; Gunatilaka and Chotikapanich 2006; Gunewardena 1996). None identified their impact in terms of the entire distribution.

Secondly, this study adds an insightful postscript to the growth-equity debate in the international literature of the 1980s which used Sri Lanka's experience with welfare policies as a test case (see Osmani 1994 for a comprehensive review). The proponents of welfare policy argued that Sri Lanka's favourable human development indicators in the 1950s, 1960s and the 1970s were due to heavy public investment in social welfare (Isenman 1980; 1987; Pyatt 1987; Sen 1981; 1988). Others argued that this development was obtained at the cost of economic growth and that if Sri Lanka had invested in growth-oriented policies it would have performed as well in the social development sphere (Bhalla 1988a; 1988b; Bhalla and Glewwe 1986). On the basis of household income and expenditure data collected just three years after economic liberalization, Bhalla (1988c) further argued that cutting expenditure on social welfare and liberalizing the economy actually improved Sri Lanka's social indicators and resulted in greater income equality.

The empirical bases of these contentions have been discussed exhaustively elsewhere and the fact established that growth alone could not have achieved Sri Lanka's levels of human development (see Anand and Kanbur 1991). Of greater interest is Sri Lanka's experience since the early 1980s. Hindsight shows that Bhalla (1988c) was premature in hailing the end of the welfare state with the opening up of the economy. Notwithstanding the policy turnaround of 1977, many of the political economy factors that first saw the rise of Sri Lanka's welfare state have continued to make successive governments highly responsive to issues of poverty, inequality and social welfare (Gunatilaka 2005). In fact, the government remains the single largest provider of health and education services - free of charge. Government investment in infrastructure continues to dominate the sector and is driven by equity considerations. For example, the fiscal straitjacket imposed by the secessionist conflict has seen drastic cuts in urgently needed, large infrastructure projects such as highways. But rural electrification schemes and rural access roads continue to be financed.

This study aims to look at the impact of these policies on the distribution of income. In particular, it looks at who benefited from greater access to policy-related income generating assets such as education and infrastructure. It also investigates the distributional impact of structural change in the economy after economic liberalization, when agriculture's dominance ended, industry claimed the largest share of the country's exports and services contributed most to total output.

The empirical analysis uses Sri Lankan household expenditure data for 1985 and 2002. It first looks at the actual distribution of income in the two periods, then uses semi-parametric techniques to try and isolate the effects of structural changes in education, industry and infrastructure access on any changes in the distribution of income.

The paper is organised as follows: Section 2 sets out the methodology and Section 3 discusses the data used and defines the variables examined. Section 4 presents the results of the analysis. Section 5 concludes by drawing out the policy implications of the study's findings.

2. THE METHOD

Since the main objective of this investigation is to derive more informed policy choices, we consider only those variables that government can influence through policy interventions: (i) changes in the structure of educational attainment of principal income earners, (ii) changes in the industrial structure in terms of the source of households' main income, and (iii) changes in household access to infrastructure. The two years we consider are 1985 and 2002 for reasons set out in the section on data and variables.

We use DiNardo et al.'s (1996) semi-parametric approach as it enables a visual and hence clearer representation of the impact of the determinants of inequality. The methodology has been extended in several recent papers to assess the impact of key variables on various functions related to income distribution. See for example, Cameron on Java (2000), D'Ambrosio (2001) on Italy, Johnson and Wilkins (2004) on Australia and Machado and Mata (2005) on Portugal.

The methodology involves constructing counterfactual densities to shed light on the factors that cause differences across income distributions. That is, the procedure simulates entire counterfactual distributions that are made to differ from the actual distribution in such a way that it enables one to identify the impact on the actual distribution, of a change in one of the variables that underlies it. It permits one to see the impact of a change in a single variable between two survey years on the entire distribution of income. Hence one has to be parsimonious when selecting explanatory factors.

Counterfactual densities are constructed by re-weighting kernel density estimates of actual income at the beginning of the reference period, by a system of weights carefully constructed to reflect the changes in household attributes that have actually taken place by the period's end. For example, if we were to look at the impact of the change in the structure of education of principal income earners on the distribution, the re-weighted counterfactual density shows us how the density of income in 1985 would have changed if the educational structure of principal income earners had been as it was in 2002. Hence, the methodology enables one to analyse where exactly on the distribution each factor exerts the greatest impact. The procedure may also be used to calculate changes in summary measures of inequality such as the Gini coefficient. In the following sections we first introduce the notation before setting out the decomposition methodology.

2.1 Notation

We begin by expressing each individual observation in terms of a vector $\left(y, z, t\right)$ of income *y*, a vector of household attributes *z*, and a date *t*. In this example and in the application to follow, the variable *t* takes only two values, 1985 and 2002. Let the joint cumulative distribution function (cdf) of these variables be denoted by $F\left(y, z, t\right)$. The probability density function (pdf) of income at one point in time, $f_t\left(y\right)$ is the integral of the pdf of income conditional on a set of household attributes and a date t_y, $f\left(y \middle| z, t_y\right)$, over the cdf of attributes $F\left(z \middle| t_z\right)$ at date t_z as follows:

$$\begin{aligned} f_t\left(y\right) &= \int_{z \in \Omega_z} f\left(y \middle| z, t_y = t\right) dF\left(z \middle| t_z = t\right) \\ &= f\left(y; t_y = t, t_z = t\right). \end{aligned} \tag{1}$$

Here, Ω_z is the domain of household attributes. In Equation (1) $f\left(y; t_y = t, t_z = t\right)$ is the pdf of income where both *y* and *z* are at time *t*. For example, $f\left(y; t_y = 85, t_z = 85\right)$ is the pdf of income in 1985 with the structure of the household attribute, *z* remaining as it is in 1985. As the expression in Equation (1) enables us to define different dates for *y* and *z*, the pdf can also be defined in terms of different dates for *y* and *z*. For example,

$f\left(y;t_y=85,\ t_z=02\right)$ is the density that would have prevailed in 1985 if the distribution of household attributes was the same as in 2002.

We now define the vector z to consist of three attributes, $z=\left(e,i,a\right)$ where e is the education level of the household's main income earner, i is the industrial sector of the household's main source of income and a is the household's access to infrastructure amenities. Equation (1) can be extended to reflect this. For example, $f\left(y;t_y=85,\ t_e=85,\ t_i=85,\ t_a=85\right)$ is the pdf of income in 1985 for households with attributes of 1985 and $f\left(y;t_y=85,\ t_e=02,\ t_i=85,\ t_a=85\right)$ is the counterfactual pdf of income which is the pdf of income that would have prevailed in 1985 if the education structure of households was the same as in 2002.

If $f\left(y;t_y=85,\ t_e=85,\ t_i=85,\ t_a=85\right)$ and $f\left(y;t_y=85,\ t_e=02,\ t_i=85,\ t_a=85\right)$ can be estimated, then the difference between the two will be the effect on the income distribution of a change in the structure of education of principal income earners between 1985 and 2002.

Given the exposition in the previous paragraph, the total change between the 1985 and 2002 distributions of income can be decomposed into the following four components using sequential decomposition as in Equation (2) below. Note how each component isolates the impact of one factor. Thus, the first isolates the impact on the distribution of the change in the education structure, the second isolates the impact of the change in the industrial structure, the third isolates the impact of change in access to infrastructure and the fourth is the residual distributional change.

The decomposition in Equation (2) can be carried out using a different ordering of variables, and a key drawback of sequential decomposition is that a different order may give different results. Hence, DiNardo et al. (1996) suggest that the decomposition should be performed again in reverse order to ensure that the impact of the change in one factor is not overstated by considering it first.

$$\begin{aligned} f_{85}(y) - f_{02}(y) = & \left[f\left(y; t_y = 85,\ t_e = 85,\ t_i = 85,\ t_a = 85\right) \right. \\ & \left. - f\left(y; t_y = 85,\ t_e = 02,\ t_i = 85,\ t_a = 85\right) \right] \\ & + \left[f\left(y; t_y = 85,\ t_e = 02,\ t_i = 85,\ t_a = 85\right) \right. \\ & \left. - f\left(y; t_y = 85,\ t_e = 02,\ t_i = 02,\ t_a = 85\right) \right] \\ & + \left[f\left(y; t_y = 85,\ t_e = 02,\ t_i = 02,\ t_a = 85\right) \right. \\ & \left. - f\left(y; t_y = 85,\ t_e = 02,\ t_i = 02,\ t_a = 02\right) \right] \\ & + \left[f\left(y; t_y = 85,\ t_e = 02,\ t_o = 02,\ t_a = 02\right) \right. \\ & \left. - f\left(y; t_y = 02,\ t_e = 02,\ t_o = 02,\ t_a = 02\right) \right]. \end{aligned} \tag{2}$$

Armed with the above notation, we can now set out the decomposition methodology. Note that Equation (2) involves the construction of three counterfactual pdfs. They are, $f\left(y; t_y = 85,\ t_e = 02,\ t_i = 85,\ t_a = 85\right)$, $f\left(y; t_y = 85,\ t_e = 02,\ t_i = 02,\ t_a = 85\right)$ and

$f\left(y;t_y=85,\ t_e=02,\ t_i=02,\ t_a=02\right)$. Following DiNardo et al. (1996), these counterfactual pdfs are obtained by re-weighting the original pdf $f\left(y;t_y=85,\ t_e=85,\ t_i=85,\ t_a=85\right)$ using re-weighting functions.

For example, consider $f\left(y;t_y=85,\ t_e=02,\ t_i=85,\ t_a=85\right)$ the first of the three counterfactual pdfs, which is the pdf of income that would have prevailed in 1985 if the education structure of households was the same as in 2002.

$$\begin{aligned} & f\left(y;t_y=85,t_e=02,t_i=85,t_a=85\right) \\ & =\iint f\left(y\,|\,e,i,a,t_y=85\right)dF\left(i,a\,|\,e,t_i=85,t_a=85\right)dF\left(e\,|\,t_e=02\right). \end{aligned} \tag{3}$$

This can be rewritten as:

$$f\left(y;t_y=85,t_e=02,t_i=85,t_a=85\right)$$

$$=\iint f\left(y\,|\,e,i,a,t_y=85\right)\,dF\left(i,a\,|\,e,t_i=85,t_a=85\right)\times\ \psi_e\,dF\left(e\,|\,t_e=85\right), \tag{4}$$

where the re-weighting function ψ_e is defined as,

$$\psi_e=\frac{dF\left(e\,|\,t_e=02\right)}{dF\left(e\,|\,t_e=85\right)}=\ \frac{\Pr\left(t_e=85\,|\,e\right)}{\Pr\left(t_e=02\,|\,e\right)}\cdot\frac{\Pr\left(t_e=02\right)}{\Pr\left(t_e=85\right)}\,. \tag{5}$$

It can be shown that the function ψ_e in Equation (5) is defined as the share of individuals in households whose principal income earner has education level *e* in 2002 divided by the corresponding proportion for 1985. The term $\Pr\left(t_e = 85 | e\right)$ is the probability of a household in 1985 whose main income earner is educated up to level *e*. The term $\Pr\left(t_e = 85\right)$ is defined as the probability of a household in 1985 regardless of education level. The corresponding terms for 2002 are defined accordingly.

It is important to note that in order to obtain Equation (4), we assume that the 1985 structure of income which is represented by $f\left(y | e, i, a, t_y = 85\right)$ does not depend on the distribution of the attributes *e*, *i* and *a*.

2.2 Estimation of Counterfactual Density

Counterfactual density is estimated by re-weighting the income density function for which we need to know the income distribution function. This can be estimated by using a kernel density function.

Let y_i, $i = 1, ..., N$ denote observed sample values of income and *h* is the bandwidth. The w_i 's denote weights associated with each y_i and *K* denote the kernel function. Then the kernel density estimator at income value *y* is given by,

$$\hat{f}\left(y\right) = \frac{1}{N}\sum_{i=1}^{N}\frac{w_i}{h\lambda_i}K\left(\frac{y - y_i}{h\lambda_i}\right). \qquad (6)$$

The kernel function used here is the Epanechnikov kernel. The term w_i denotes the population share associated with each income unit, y_i.

Local bandwidth factors λ_i are proportional to the square root of the underlying density functions at the sample points as follows:

$$\lambda_i = \lambda(y_i) = \left(\frac{g}{\tilde{f}(y_i)} \right)^{0.5}. \tag{7}$$

The term *g* is the geometric mean over all *i* of the pilot density estimate, $f(\tilde{y})$. The pilot density estimate is obtained with *h* as bandwidth (see Van Kerm 2003 for details).

To obtain the counterfactual density, we re-weight the adaptive kernel density function expressed by Equation (6) as follows:

$$\hat{f}(y; t_y = 85, t_e = 02, t_i = 85, t_a = 85) = \frac{1}{N} \sum_{i=1}^{N} \frac{w_i^*}{h\lambda_i} K\left(\frac{y - y_i}{h\lambda_i} \right). \tag{8}$$

In this equation, w_i^* is the weight associated with y_i and is defined as:

$$w_i^* = \frac{n_i \, \psi_e}{\sum_{i=1}^{N} n_i \, \psi_e}, \tag{9}$$

where n_i is the number of individuals associated with income y_i.

Thus, the difference between the actual 1985 density and the hypothetical density represented by Equation (4) represents the impact on the distribution of changes in the composition of the education level of principal income earners between 1985 and 2002. This completes the first part of the sequential decomposition set out in Equation (2).

The same procedure is repeated for the other compositional changes. For example, to isolate the impact of the changing industrial structure in Equation (2), we would need to estimate $f\left(y; t_y = 85,\ t_e = 02,\ t_i = 02,\ t_a = 85\right)$. This is the density that would have prevailed in 1985 if the education composition of principal income earners and the industrial sector of households' main source of income were as they were in 2002. To do this, the weighting function $\psi_{e|i}$ is calculated as the ratio of the shares of individuals in a certain education/industry category in 2002 to those in 1985. We apply this new set of adjusted weights to estimate the counterfactual density associated with the changed education and industrial structure as in Equation (8). The difference between this density and the counterfactual density represented by Equation (3) is the impact of the changing industrial structure.

Similarly, to isolate the impact of the changing pattern of access to infrastructure as in the third part of Equation (2), we calculate the weighting function $\psi_{e|i|a}$ as the ratio of shares of individuals in the education/industry/infrastructure categories in 2002 to those in 1985. We use the new set of weights to estimate $f\left(y; t_y = 85,\ t_e = 02,\ t_i = 02,\ t_a = 02\right)$, the counterfactual density that would have prevailed in 1985 if the education, industrial and infrastructure composition of the population were as they were in 2002. The difference between this counterfactual density and $f\left(y;t_y=85,\, t_e=02,\, t_i=02,\, t_a=85\right)$, the counterfactual density that would have prevailed in 1985 if only the education and industrial composition of the population were the same in 1985 as in 2002, represents the impact of the change in infrastructure access on the income distribution.

Finally, to obtain that part of inequality change between 1985 and 2002 that is not explained by the three factors, we isolate the impact of the residual as in the final part of Equation (2). That is, the difference between the actual income density of 2002, $f\left(y;t_y=02,\, t_e=02,\, t_o=02,\, t_a=02\right)$, and the counterfactual density that would have prevailed in 1985 if the education, industrial and infrastructure composition of the population were as they were in 2002, $f\left(y;t_y=85,\, t_e=02,\, t_i=02,\, t_a=02\right)$, represents the impact of the change in unexplained factors on the income distribution.

It is also possible to quantify the impact of each of the variables on total inequality by estimating any inequality measure such as the Gini coefficient based the original and counterfactual density functions. To do this we use the Lerman and Yitzhaki (1989) procedure to calculate the Gini coefficient based on the following formula,

$$G=\sum_{i=1}^{N}\eta_{i+1}\pi_i-\sum_{i=1}^{N}\eta_i\pi_{i+1}\,, \tag{10}$$

where π is the cumulative proportion of income units and η is the cumulative proportion of income.

In the following section we discuss data issues and define the variables.

3. DATA AND VARIABLES

The analysis uses consumption expenditure as proxy for income because it 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 less developed countries, while means-tested income support programmes can encourage under-reporting of income. Also, consumption expenditure is a direct measure of individual and household welfare whereas income streams exhibit transitory fluctuations (Barrett et al. 2000).

Data for the analysis were drawn from the 1985/86 Labour Force and Socio-Economic Survey (LFSS) and the 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. Since the 2002 survey does not include data from the Northern and Eastern Provinces due to the conflict situation, the present analysis relates only to the seven provinces outside the Northern and Eastern Provinces. This population accounts for about 85 per cent of Sri Lanka's population of roughly 19 million people.

Although an earlier data set closer to the time of liberalization exists, we decided to use 1985 rather than 1980 as one end point of the reference period because even though consumption expenditure of the surveys are broadly comparable, the income definitions of the 1980 survey are different from those of subsequent surveys. And since the construction of the explanatory variables we consider in the present analysis requires us to identify the principal income earner and principal source of income by using the income rather than the consumption expenditure data, we use data from the 1985 survey as being more comparable with the end point of the 2002 survey than the 1980 survey data.

The analysis takes into account only households with positive expenditure. We also excluded households that box plot analyses of consumption data revealed as outliers. Individual expenditure was adjusted to take into account equivalence scales, economies of scale, and temporal and spatial differences in the cost of living as follows.

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. One way to estimate m_i is as follows:

$$m_i = \left(\phi_1 n_{a,i} + \phi_2 n_{c,i} \right)^{\theta}. \qquad (11)$$

In the formulation above, the number of adults in household i is $n_{a,i}$ and the number of children is $n_{c,i}$. The parameter 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. The term ϕ_2 is the cost of a child relative to an adult and can take any value $0 \le \phi_2 \le 1$. The setting of these parameters for developing countries tends to be arbitrary. For example, based on analyses of household survey data for Sri Lanka and Indonesia of the late 1960s and 1970s, Deaton and Zaidi (2002) recommend that ϕ_1 be set at unity and ϕ_2

at between 0.25 or 0.33 for developing countries. However, in this paper we set ϕ_2 and θ as 0.6 and 0.9, respectively. This is because the costs of children relative to adults have increased in Sri Lanka over the last two decades and the scope for economies of scale has also increased.

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).

We consider the impact of three variables on the distribution of income. They are changes in the education composition of households' principal income earners, changes in the composition of the industrial sector of households' main source of income and household access to infrastructure. The methodology requires us to define these variables in terms of mutually exclusive and mutually exhaustive categories as set out below.

Five education categories are defined for the education level of principal income earners, *Primary education or less, Secondary education, Education up to GCE Ordinary Level (10th year certificate), Education up to GCE Advanced Level (12th year certificate),* and *University education and more***.**

To assess the impact of changing industrial structure on inequality, we defined four categories to denote the industrial sector which is the household's main source of income. The category A*griculture, fishing and mining* includes all cultivation activities and livestock production, hunting, fishing, forestry and logging, and mining and quarrying.

Manufacturing includes the manufacture of food, beverages and tobacco, textiles, wearing apparel and leather industries, manufacture of wood and wood products, manufacture of paper and paper products, printing and publishing, manufacture of chemicals, petroleum, rubber and plastic products, basic metal industries, and manufacture of fabricated metal products, machinery and equipment. *Wholesale and retail trade, hotels, transport, finance and real estate* is defined as a separate category to differentiate the sectors in this category from other services. This is because these sectors are likely to have grown faster than other service sectors following economic liberalisation. *Other services* includes electricity, gas and steam, water works and supply, construction, public administration and defence, sanitary and similar services, social and related community services, recreational and cultural services, personal and household services, and services not adequately defined.

Eight infrastructure categories represent households' access to infrastructure facilities. They are, *No access to either a vehicle, electricity or telephone, Access to a vehicle only, Access to electricity only, Access to telephone only, Access to vehicle and electricity only, Access to vehicle and telephone only, Access to electricity and telephone only, Access to all three infrastructure amenities.*

One practical problem which arose in defining these categories is that certain population categories that did not have any observations in 1985 had observations in 2002. This issue surfaced in at least a dozen education/industry/infrastructure categories considered here. For example, we found that households where the main income earner had secondary level education, whose main source of income was from the manufacturing sector, and which did not have access to electricity in 1985, had acquired electricity by 2002. This caused the total number of adult equivalents denoted by the weights adjusted for education/industrial/infrastructure categories, to exceed the total number of adult equivalents denoted by the original weights. Since this made the comparison of Gini coefficients based on the original and re-weighted data difficult, we adjusted the ψ 's associated with the infrastructure categories so that the total number of adult equivalents denoted by the adjusted weights would equal the total denoted by the original weights.

The percentage shares of the sample of households in the two survey years by education, industry and infrastructure categories are set out in Table 1.

As far as the change in the composition of the education level of the principal income earner goes, Table 1 reveals considerable improvement in educational attainments over the years. The proportion of principal income earners with only primary education or less has dropped markedly from 37 per cent to 31 per cent. At the same time, the proportion of principal income earners with GCE A' Level education has risen noticeably from 3.3 per cent to nearly 10 per cent.

In contrast, changes in the industrial composition of households' main source of income have been rather negligible. Agriculture, mining and fishing still claims the major share, though down from 46 per cent to 44 per cent. The share of manufacturing has hardly changed, and that of trade, hotels, transport, finance and real estate has, contrary to expectations, actually fallen from 19 per cent to 15 per cent. The share of other services, however, has increased markedly from 24 per cent to 30 per cent.

From the variables considered here, the most notable changes have occurred in the composition of households' access to infrastructure. The percentage of households without access to any of the infrastructure facilities has reduced drastically from 72 per cent to 29 per cent. Significant structural changes in infrastructure access that underlie this movement but are not immediately apparent from Table 1 are as follows. The proportion of households owning vehicles has doubled from 8 per cent to 16 per cent. The proportion owning telephones has increased many times over from a paltry 1 per cent to a more respectable 21 per cent, clearly facilitated by the privatisation of the state-owned telecommunications enterprise and the entry of mobile telecommunications service providers. Finally, there has been a spectacular increase in the proportion of households with access to electricity from 24 per cent in 1985 to 67 per cent by 2002, driven almost entirely by the state-owned service provider.

In terms of the mutually exclusive and mutually exhaustive categories considered in this chapter and illustrated in Table 1, these changes translate to the following transformations. The proportion of households with access to electricity only has doubled, and the share of households with access to electricity and telephones has increased from half a percentage point to 12.5 per cent. Thus, while many households which did not have electricity in 1985 had hooked up to the national grid by 2002, many households which already had electricity in 1985 had probably acquired telephones by 2002. The proportion of households with access to all three amenities has risen dramatically from half a percentage point to seven per cent.

What was the impact of these structural changes on the distribution of income? In the following section we address this question with the results of the decomposition analysis.

4. RESULTS

We begin this section by using adaptive kernel methods to estimate the empirical income distribution functions for the years 1985 and 2002 (see Figure 1). We do this to illustrate the magnitude of the change in distribution that occurred between these two years. It is this change that we decompose using the semi-parametric methodology discussed above.

In the top panel of Figure 1, we see a general shift to the right and the squashing down of the distribution. The middle mass shrank by a simultaneous process of levelling down density at lower income levels (see the left 'tail' of the distribution), an even more rapid levelling off of the middle and a squeezing out to the right, indicating a shift in the concentration of people towards higher income ranges at every stage in the distribution. These trends are reflected in the bottom panel of Figure 1 which plots the difference in density between the two years. In Burkhauser et al.'s (1999) words used in a different country context, 'Increased prosperity rather than 'immiserisation' best describes this movement' (p. 266).

In the following sections we analyse the role played by structural change in three factors in bringing about this change in the density of income distribution. The factors are, changes in the education composition of principal income earners, changes in the sector from which households derive their main income and, changes in households' access to infrastructure. We present the decompositions in terms of movements in counterfactual density functions associated with each stage of decomposition. This is done by means of two graphs for each stage showing the impact of the variable of interest. The first graph shows the actual density alongside the counterfactual density. The second shows the difference in the two densities.

4.1 Changes in Education Composition

In Table 1 we saw that educational attainment across the population of principal income earners increased between 1985 and 2002. What was the impact of this change on the distribution?

The top panel in Figure 2 shows the actual kernel density function of 1985 and how it looks when re-weighted to reflect the education structure of principal income earners in 2002. The bottom panel shows the difference in the two distributions. It can be seen that the re-weighting has resulted in noticeable, but small distributional consequences. The squashing down of the distribution is apparent, but there is hardly any evidence of a rightward shift brought about by changes in education composition.

It is more interesting to see where in the distribution the impact of the change in education levels has been greatest. Note that the impact has been minimal at levels of income less than Rs. 100 per adult equivalent per month. The most noticeable impact has been around the mode, at income levels within the Rs. 200-300 range. We also see an expansion of income receivers in the Rs. 400-1000 range. Hence, it appears that changes in the structure of education increased the density of income receivers in the lower to middle income range. It has had little impact on the proportion of income receivers in the higher income ranges. Thus, the impact of the change in educational composition between the two survey years has been concentrated around the lower to middle income ranges rather than at either the lower or the upper end of the distribution.

4.2 Changes in Industrial Composition

In this section we analyse the impact on the distribution of income in 1985 of changes between 1985 and 2002 in the composition of the industrial sectors from which households' main income derives. To do this we estimate the 1985 distribution adjusted for changes in education composition as well as the 1985 distribution re-weighted to reflect changes in composition of both education and industry of main source of income (see the top panel of Figure 3). The difference between the two densities which shows the impact of changing industrial composition on the income distribution is plotted in the bottom panel of Figure 3.

It is clearly apparent that the change in industrial composition has had only a marginal impact on the distribution of income. While the two re-weighted distributions almost perfectly map each other as in the top panel of Figure 3, the bottom panel shows that changes in industrial composition caused inequality to change only infinitesimally. At its maximum, the difference in density between the two was only 0.00002. Note in particular the scale of the graph in the bottom panel. Although the graph is as large as the bottom panel of the graph showing the impact of educational changes (Figure 2), so that the changes may be seen more clearly, the scales are very different. Note further that the slight impact on the distribution deriving from a change in the composition of industrial sectors is concentrated at the low end of the income scale. Thus, the results show that had the industrial composition of the main source of household income in 1985 had been as it turned out to be in 2002, there would have been little noticeable difference in the distribution.

This is not surprising. As is apparent from Table 1, there have only been very slight changes between the two years in the composition of the industrial sectors from which households derive their principal source of income.

4.3 Changes in Access to Infrastructure

Table 1 shows that between 1985 and 2002, the greatest structural changes in households' characteristics have occurred in the infrastructure categories. Figure 4 shows their impact on the distribution of income. The top panel of Figure 4 shows the density function for 1985 adjusted for educational and industry changes, as well as the same function adjusted for educational, industrial and infrastructure changes. The difference between these two functions is the impact of infrastructure changes on the distribution appears in the bottom panel of Figure 4.

It is apparent that of the three variables we have analysed in this chapter, the change in infrastructure access has had the biggest impact on the distribution. It has flattened the left 'tail', squashed the modal peak from a density of 0.003 to about 0.00225, as well as shifted it out to the right. The movement signifies a much longer range of income levels along which households moved to enjoy higher levels of income. Notice in particular that unlike education changes, infrastructure changes had a significant impact on the distribution of income at the lower end of the scale, beginning with those with income of about Rs. 75 per month. The flattening and squashing of the modal peak has caused an expansion in the right tail of the distribution, particularly between the Rs. 400-1100 income range. The biggest concentration has occurred around the Rs. 750 per month income level, but it can be seen that the increase in density in the right tail of the distribution has carried over even to the highest income levels.

Thus, it can be concluded that the change in the structure of access to infrastructure between 1985 and 2002 has had a marked impact on the distribution. In particular, it has resulted in greater prosperity with far greater numbers enjoying higher income levels than before.

4.4 Impact of Unexplained Factors

What remains to be seen is the impact on the distribution of structural changes in factors which we have not considered above. In order to analyse this, we compare the adaptive kernel density of 1985 re-weighted to reflect the structural composition of education, industrial and infrastructure variables in 2002, with the actual kernel density of 2002. The results can be seen in Figure 5.

Unexplained factors appear to have played a bigger role in bringing about changes in inequality than have changes in either industrial composition or education composition. Note the shift in the left tail of the distribution, almost evenly along its entire length, signifying the movement of these income receivers towards higher levels of income within the ranges of Rs. 300-750 per month and Rs. 850-1300 per month.

4.5 Impact on Inequality

We can also calculate the total change in the Gini coefficient that can be attributed to each stage of the decomposition. However, since the results may be dependent on the order in which variables are considered, we repeated the decomposition for the reverse sequence: that is, we adjusted the density first for infrastructure changes, secondly for infrastructure and industrial changes and finally for changes in infrastructure, industry and education. Changes in the Gini at each stage of decomposition for both sequences are presented in Table 2. The contributions of each factor to total change in inequality are presented in Table 3.

In Table 2 the Gini coefficients for 1985 and 2002 are those calculated with the historical weights. These are the number of adult equivalents associated with each income observation. The Gini coefficients for the three stages of decomposition are those calculated with the relevant adjusted weights associated with each income observation. For example, the Gini coefficient for the second stage of decomposition when the order of decomposition is education, industry and infrastructure (Sequence A), is calculated with weights that reflect the education and industrial composition of 2002. In contrast, the Gini coefficient of the second stage of decomposition when the order reversed (Sequence B) is calculated with weights that reflect the infrastructure and industry composition of 2002.

The change in Gini at each stage of decomposition represents the change in total inequality attributable to that factor whose impact is isolated at that stage. So for example, the impact of education is isolated at the first stage of decomposition in

Sequence A. Thus, the Gini coefficient of 0.3365 represents total inequality in 1985, if the composition of education of main income earners in 1985 were to reflect the changes that had taken place by 2002 (Table 2).

The difference between this figure and the Gini coefficient calculated with the original, unadjusted weights – that is 0.3232 – is the contribution of this factor to the total change in inequality. This is set out in Table 3 under Sequence A as +0.0133. Thus, education changes increased inequality. When expressed as a share of the original Gini coefficient, the change in the structure of education increased total inequality by 43.32 per cent.

Similarly, decomposition Sequence B shows that if the composition of infrastructure access alone of households in 1985 had been the way it was in 2002, inequality as measured by the Gini coefficient would have been 0.3905 in 1985 rather than 0.3232 (Table 2). Thus, infrastructure increased inequality by +0.0674, or by 219.54 per cent (Table 3).

It can be seen that changing the decomposition sequence did not affect the broad conclusions of the analysis. The change in the composition of access to infrastructure has by far the largest impact on the distribution in both sequences. The impact of the change in education composition was small, yet noticeable. The change in industrial structure had little impact in both runs, though its contribution in Sequence A was negative and in Sequence B it was positive. The results also show that to some extent, the inequality-increasing impact of infrastructure changes on total inequality were mitigated by the inequality-reducing impact of unexplained variables represented by the residual.

5. CONCLUSIONS

In this study we applied a semi-parametric method to investigate the impact of changes in three key variables on the entire distribution of income, in order to target policy interventions more effectively. The policy-related variables considered were, (i) changes in the structure of educational attainment of principal income earners, (ii) changes in the industrial structure in terms of the source of households' main income, and (iii) changes in household access to infrastructure.

The analysis found that the rise in inequality between 1985 and 2002 was accompanied by a shift in the concentration of people towards higher income ranges at every stage in the distribution. There was a noticeable reduction of density around the modal peak and a corresponding increase in the concentration of people in the middle to upper income ranges. Higher levels of educational attainment of principal income earners during this period had a noticeable impact on income receivers around the lower to middle income ranges. As a result, income density along the upper middle income range increased. Structural change in the composition of sectors from which households derived their main source of income showed little change, and as expected, there was little discernible impact on the distribution of income. In contrast, much of the rightward shift in the concentration of people at higher income ranges between 1985 and 2002 appears to have been caused by the change in access to infrastructure.

The analysis shows that Sri Lanka's complementary mix of growth-oriented and welfare policies equipped large sections of the population with income-generating assets that enabled them to take advantage of the open economic policy framework. Their incomes rose and they moved up along the distribution. This is a notable achievement, especially given that an intractable secessionist conflict has claimed the lion's share of the government's budget and inhibited the flow of foreign investment over the period.

The empirical findings of the present investigation rejects key tenets of the doctrine that claimed that cutting social welfare expenditure and liberalising the economy actually improved Sri Lanka's social indicators and resulted in greater income equality. Instead, a key conclusion to emerge from the analysis is the strong role that the government needs to play to ensure access to education and infrastructure services in such a way that ensures economic growth and greater social welfare.

Nevertheless, inequality has also risen as the lower middle classes and middle classes benefited disproportionately more from state provision of education and infrastructure services. Moreover, other studies have shown that those in the bottom two deciles of the income distribution remain below the poverty line (see Narayan and Yoshida 2004). The results suggest that policies related to education, infrastructure, and possibly health services provision should be targeted more effectively towards those in the poorest deciles.

Since many of these determinants of inequality are tied to space and operate within the spatial dimension, reducing spatial and regional inequalities in access to these services suggests itself as the best way to address the problem. For example, access to electricity is determined by the spread of the distribution system in terms of the laying out of power lines. Access to education and the level of educational attainment is determined by the distance from one’s home to the nearest school and the quality of education it provides. Similarly, the local transport system which again operates in the spatial dimension, determines whether one can get one’s products to the market before they perish or whether one can get a better paying job in the next village or in the next town. In particular, in rural areas that have few comparative advantages to attract business investment, enhancing the quality of education and skills development would enable workers to find jobs in industrialised urban centres that have location-based advantages for industrial development. The government may need to develop such urban centres and facilitate the movement of workers and goods to them from backward regions and so pursue a policy of spatial integration in terms commodity and factor markets. This should be easier in a small country like Sri Lanka where the terrain is more conducive to establishing more efficient transport and communications systems.

REFERENCES

Anand, S., & Kanbur, R. Public Policy and Basic Needs Provision: Intervention and Achievement in Sri Lanka. In J. Dreze & A. K. Sen (Eds.), *The Political Economy of Hunger: Endemic Hunger* (pp. 59-92). Oxford: Clarendon Press 1991.

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

Bhalla, S.S. Is Sri Lanka an Exception? A Comparative Study of Living Standards. In T.N. Srinivasan & P.K. Bardhan (Eds.), *Rural Poverty in South Asia* (pp. 89-117). New York: Columbia University Press 1988a.

Bhalla, S.S. Sri Lanka's Achievements: Fact and Fancy. In T.N. Srinivasan & P.K. Bardhan (Eds.), *Rural Poverty in South Asia* (pp. 557-65). New York: Columbia University Press 1988b.

Bhalla, S.S. (1988c). The Politics and Economics of Agricultural Price Policies in Sri Lanka. Unpublished mimeograph, The Policy Group, New Delhi.

Bhalla, S.S., & Glewwe, P. (1986). Growth and Equity in Developing Countries: A Reinterpretation of the Sri Lankan Experience. *World Bank Economic Review*, 1.

Burkhauser, R.V., Crews, A.D., Daly, M.C., & Jenkins, S.P. (1999). Testing the significance of income distribution changes over the 1980s business cycle: A cross-national comparison. *Journal of Applied Econometrics*, 14, 253-72.

Cameron, L. (2000). Poverty and inequality in Java: examining the impact of the changing age, education and industrial structure. *Journal of Development Economics*, 62, 149-180.

D'Ambrosio, Conchita (2001). Household Characteristics and the Distribution of Income In Italy: An Application of Social Distance Measures. *Rev Income Wealth*, 47, 43-64.

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

DiNardo, J., Fortin, N.M., & Lemieux, T. (1996). Labour market institutions and the distribution of wages, 1973-1992: a semi-parametric approach. *Econometrica*, 64, 1001-44.

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

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

Gunatilaka, R., & Chotikapanich, D. *Inequality Trends and Determinants in Sri Lanka 1980-2002: A Shapley Approach to Decomposition*. Business and Economics Working Paper No. 6/06, Melbourne: Monash University, 2006.

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

Isenman, P. (1980). Basic Needs: The Case of Sri Lanka. *World Development*, 8, 237-58.

Isenman, P. (1987). A Comment on "Growth and Equity in Sri Lanka: A Reinterpretation of the Sri Lankan Experience by Bhalla and Glewwe. *World Bank Economic Review*, 1, 521-531.

Johnson, D., & Wilkins, R. (2004). Effects of Changes in Family Composition and Employment Patterns on the Distribution of Income in Australia: 1981–1982 to 1997–1998. *Economic Record*, 80, 219-238.

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

Machado, J.A.F., & Mata, J. (2005). Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression. *Journal of Applied Econometrics*, 20, 445-465.

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

Osmani, S. R. (1994). Is There a Conflict Between Growth and Welfarism? The Significance of the Sri Lanka Debate. *Development and Change*, 25, 387-421.

Pyatt, G. (1987). A Comment on "Growth and Equity in Sri Lanka: A Reinterpretation of the Sri Lankan Experience by Bhalla and Glewwe. *World Bank Economic Review*, 1, 515-20.

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

Sen, A. K. (1981). Public Action and the Quality of Life in Developing Countries. *Oxford Bulletin of Economics and Statistics*, 43, 287-319.

Sen, A. K. Sri Lanka's Achievements: How and When? In T.N. Srinivasan & P.K. Bardhan (Eds.), *Rural Poverty in South Asia* (pp. 549-56). New York: Columbia University Press 1988.

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

Van Kerm, P. (2003). Adaptive kernel density estimation. *Stata Journal*, 3, 148-156.

Table 1

Population Shares by Education, Industry and Infrastructure Categories

	1985	2002
Education Category		
Primary education or less	37.11	31.72
Secondary education	41.62	39.77
GCE O' Levels	15.72	15.76
GCE A' Levels	3.26	9.74
University and above	2.29	3.02
Industry Category		
Agriculture, fishing, mining	46.22	44.02
Manufacturing	11.09	10.63
Trade, hotels, transport, finance, real estate	19.07	15.43
Other services	23.63	29.92
Infrastructure Category		
No access to any infrastructure amenity	72.02	29.35
Access to vehicle only	4.36	2.26
Access to electricity only	19.28	40.75
Access to telephone only	0.07	1.42
Access to vehicle and electricity only	3.25	6.21
Access to vehicle and telephone only	0.04	0.34
Access to electricity and telephone only	0.48	12.45
Access to all three amenities	0.50	7.23
Total number of households	18420	16318

Table 2

Two Sequences of Decomposing the Gini Coefficient

Decomposition Sequence		Gini Coefficient at Each Stage				
		Gini 1985	1st Stage	2nd Stage	3rd Stage	Gini 2002
A	Education, Industry, Infrastructure	0.3232	0.3365	0.3362	0.4052	0.3539
B	Infrastructure, Industry, Education		0.3906	0.3941	0.4052	

Notes:

The Gini coefficient at each stage represents the Gini calculated with the weights for each stage of the decomposition. For example, the first stage of the first ordering of the decomposition denotes the Gini calculated with the weights adjusted to reflect changes in education composition between 1985 and 2002. The Gini at the second stage for the same ordering of variables denotes that calculated with weights adjusted to reflect changes in education and industry composition in 2002.

Table 3

Contributions to Total Change in Inequality

Total change 0.0307

Contribution from	Sequence A	% Contribution	Sequence B	% Contribution
Education	+0.0133	+43.32	+0.0111	+36.16
Industry	-0.0003	-0.98	+0.0035	+11.40
Infrastructure	+0.0690	+224.76	+0.0674	+219.54
Residual	-0.0513	-167.10	-0.0513	-167.10

Figure 1: Adaptive Kernel Density Estimation of Income Distribution and Change in Sri Lanka, 1985, 2002

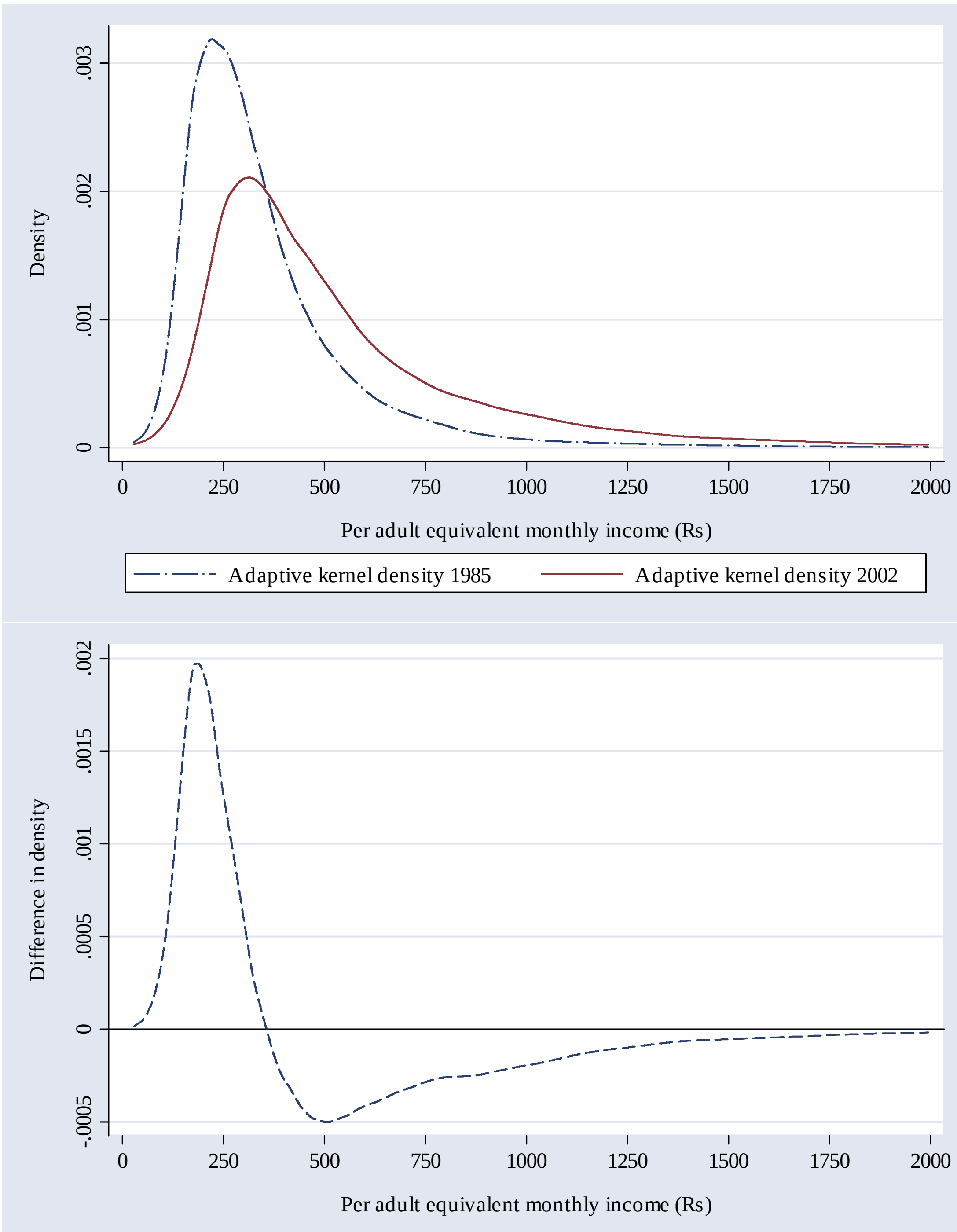

Figure 2: Distribution Adjusted for Change in Education Composition

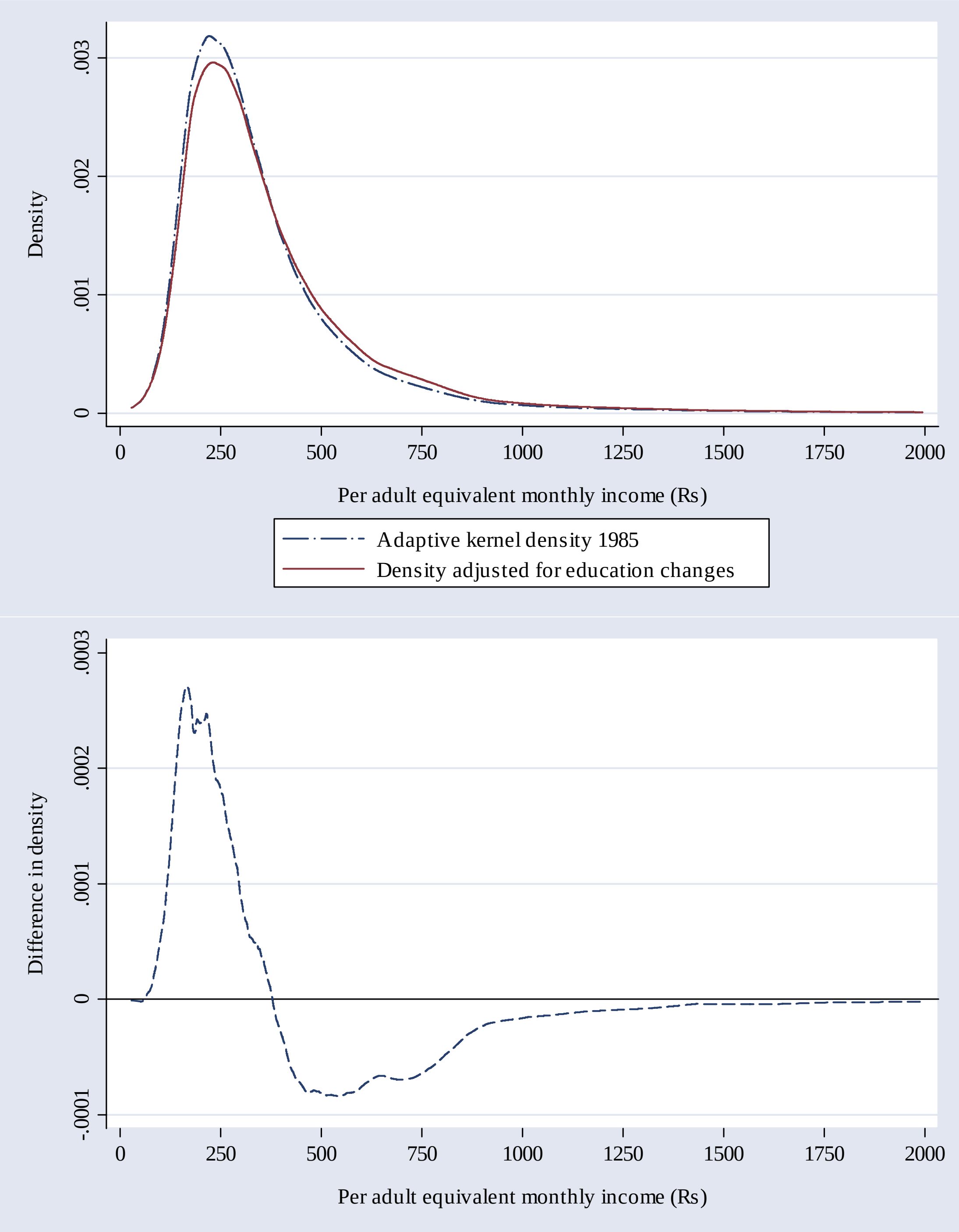

Figure 3: Distribution Adjusted for Change in Education and Industrial Composition

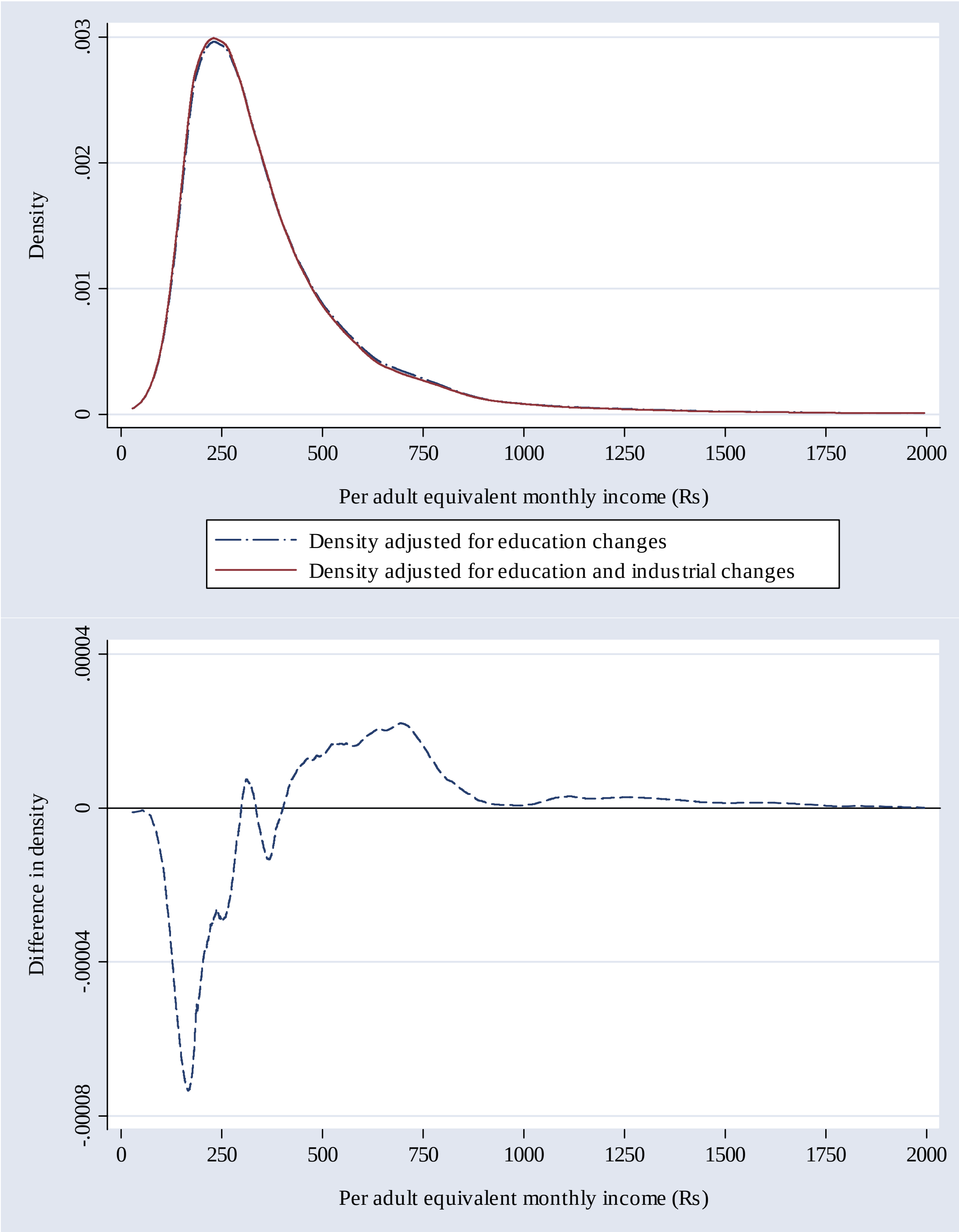

Figure 4: Distribution Adjusted for Change in Education, Industrial and Access to Infrastructure Composition

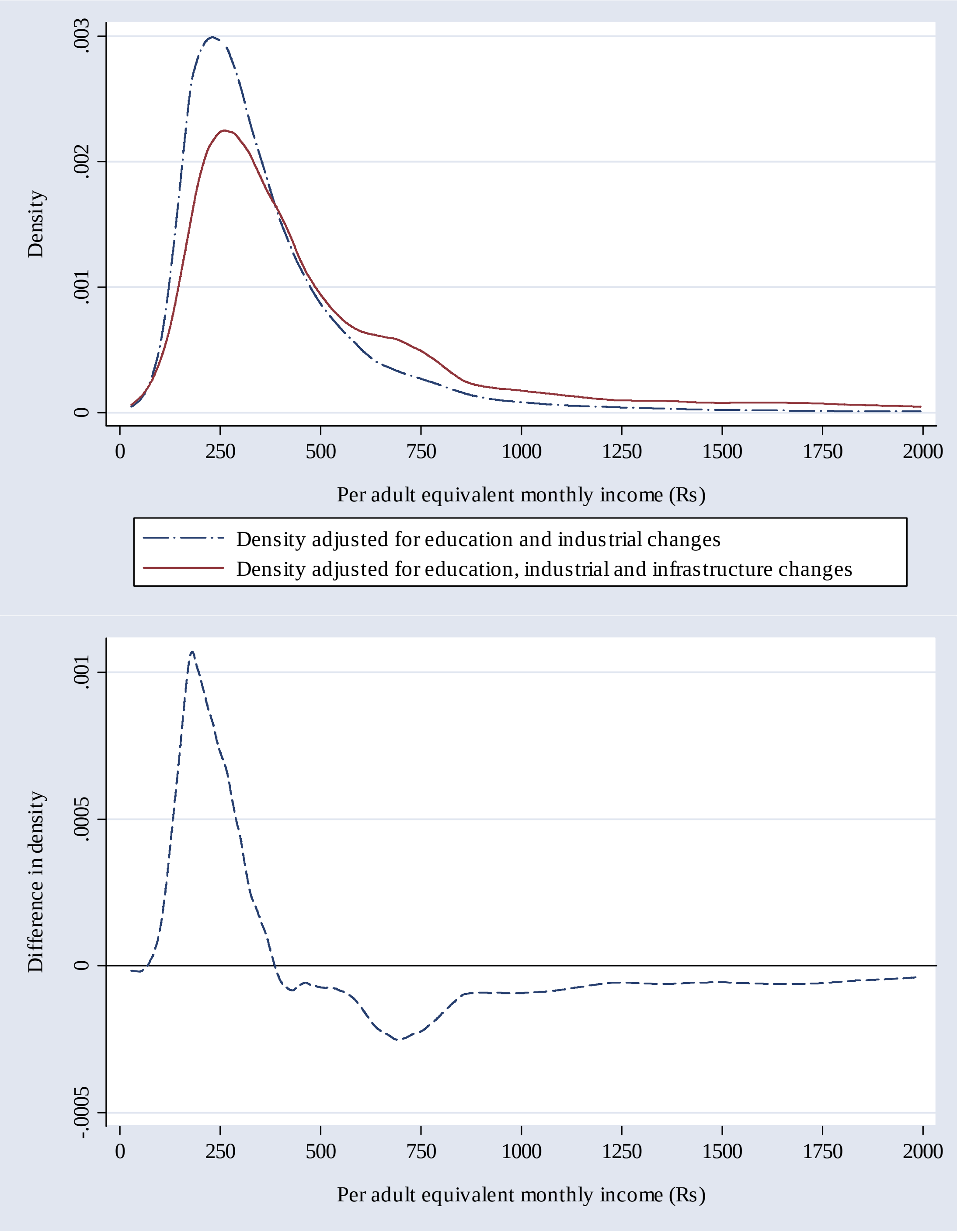

Figure 5: Impact of Unexplained Factors