Working Papers in Trade and Development

Explaining Changes in Inequality:
Myanmar, 2005 to 2010

Lwin Lwin Aung

lwin.aung@alumni.anu.edu.au

and

Peter Warr

Arndt-Corden Department of Economics

Crawford School of Public Policy

Australian National University

Peter.Warr@anu.edu.au

May 2020
Working Paper No. 2020/06

Arndt-Corden Department of Economics
Crawford School of Public Policy
ANU College of Asia and the Pacific

Explaining Changes in Inequality: Myanmar, 2005 to 2010

Lwin Lwin Aung and Peter Warr
Australian National University

We study the determinants of changes in the distribution of household expenditures in Myanmar, using a large, nationwide panel household survey data set covering 2005 and 2010. These data, compiled by the United Nations Development Program and the Myanmar government, constitute the only panel household survey data currently available for Myanmar. They show that although there was no statistically significant change in mean real consumption expenditures between these years, the distribution of that consumption across households became significantly less unequal. The sources of this large decline in measured inequality have not been systematically investigated. The paper attempts that, using the above panel data set and applying the regression-based inequality decomposition developed by Fields and its subsequent extension by Yun. The results indicate that region-specific variables, occupational changes and changes in education were the main contributors to the narrowing of expenditure inequality, but that road development expenditures had the opposite effect. On the methodology of decomposition, we argue first that the Yun extension of Fields' approach promises enhanced policy relevance, by distinguishing between changes in household characteristics and changes in their impact on expenditures, as reflected in the estimated regression coefficients. Second, we show that the decomposition described by Yun entails arbitrary sequencing of the distributions being compared, influencing the decomposition obtained. We explain the underlying reason for this problem and illustrate a simple solution.

Key words: Expenditure inequality, panel data, regression-based decompositions, Myanmar.

JEL codes: C51; D31; D63.

Explaining Changes in Inequality: Myanmar, 2005 to 2010

1. Introduction

Following decades of political repression and economic stagnation, Myanmar is gradually embarking on a program of economic and political reform. In 2011 a nominally civilian government replaced the previous military regime, and in the national elections of November 2015 a new government was elected with a mandate for wide-ranging social and economic policy reform. In this ongoing reform process, economic inequality, and changes in it are sure to be issues of continuing controversy. A key concern is that growth-promoting economic reforms may increase disparities within Myanmar. Historically, Myanmar has suffered more than one civil war related to distributional matters (South 2012). Examining the factors influencing economic inequality and, more particularly, changes in it, is therefore important if policies that promote economic growth are to be combined with supplementary measures that restrain possible increases in inequality.

A large, national household survey, conducted by the United Nations Development Programme and the Myanmar government's Ministry of National Planning and Economic Development, provides an opportunity for such an analysis (IDEA & IHLCA, 2007a, b and c) and (IHLCA, 2011a, b and c). The survey was conducted in 2005 and again in 2010, covering roughly 18,600 households in each of these years.[1] The two surveys are summarized in Table 1.[2] Uniquely for Myanmar, these surveys each contained a large panel component, covering 9,102 households, around half of those surveyed, within which the households included were the same in both years. This feature lends itself to the study of changes in economic inequality because differences in outcomes are not confounded by changes in the household samples, facilitating a focus on the factors driving change. Within this panel data set, the computed nationwide Gini coefficient of inequality of expenditure per adult equivalent declined from

[1] The raw household data were collected between November-December 2004 and May 2005 and then between December 2009 and May 2010, respectively. For brevity, we refer to these as the 2005 and 2010 surveys, respectively.
[2] A fuller summary of the panel household data is provided in Appendix Table A1.

0.248 in 2005 to 0.221 in 2010.[3] At the same time, measured poverty incidence declined from 31.7 to 24.9 per cent of the total population. The decline in both measures was statistically significant. But mean real expenditures rose by only 2.3 per cent and that sample-based estimate of the change was not significantly different from zero.[4] That is, although mean expenditures barely changed, their distribution became significantly more equal. The question is what caused this change.

[Table 1 here]

The answer may have important policy implications. Analysis of the factors determining (a) the *level* of consumption inequality in each year, and more particularly (b) the *changes* between years, described by Fields (2003) as the *levels question* and the *differences question*, respectively, may indicate "whether existing inequalities are due to intrinsic unchangeable characteristics, such as location or ethnicity, or due to variables whose distribution can be changed through policy" (Naschold, 2009, p.747). A significant decline in measured inequality within a particular country is unusual.[5] Analysis of this event has intrinsic scientific interest and is of particular policy interest for Myanmar because the years 2005 to 2010 immediately preceded the reform period, beginning in 2011. These data sets provide a baseline for the subsequent reforms.

Despite the political importance of distributional issues within Myanmar, no systematic empirical study of the determinants of the level of, and changes in, consumption inequality has yet been possible because of the unavailability of survey data like those analysed in this paper. The present paper augments Warr and Aung (2019), which explored whether a particular event, the 2008 Tropical Cyclone Nargis, could explain the above-mentioned reduction in inequality, concluding that it explained it only partially. This paper broadens the list of possible explanatory variables and uses the Myanmar data to review critically the methodological issues involved in decomposing changes in inequality.

[3] The surveys collected detailed data on household expenditures, but not incomes.

[4] It is important that survey estimates of mean expenditures, poverty and inequality are sample-based estimates of population parameters where the sample sizes are often not particularly large. The estimates have a sampling variance and it is therefore appropriate to consider whether measured changes in these estimates are significantly different from zero.

[5] Bourguinon (2015) shows that the global experience of recent decades has been declining inequality between countries but increasing inequality within most countries. The experience of Myanmar over the 2010 to 2015 period is therefore an exception to these within-country outcomes.

Our primary focus in this study is on explaining the observed decline in inequality between 2005 and 2010, using the panel data set. We apply the Fields (2003) regression-based methodology, which draws upon Shorrocks' (1982) inequality decomposition techniques, using the results of micro-econometric estimation of the determinants of expenditures at the household level. An expenditure-generating function is estimated for each of the two years 2005 and 2010 by regressing real household expenditures per adult equivalent on economically meaningful explanatory variables such as levels of education, occupations of household members, local infrastructure and geographic dummies. The list of explanatory variables draws upon standard human capital and production theory, along with past empirical findings. The estimated expenditure-generating equations are then used to calculate Fields' relative factor inequality weights, allowing the expenditure function to be decomposed into its exogenous causal factors. We also explore critically the Yun (2006) extension of Fields' analysis, drawing attention to a serious flaw in Yun's methodology and suggesting a method for correcting it.

The paper is organized as follows. Section 2 provides a brief literature review, with a focus on studies of inequality in developing countries. Section 3 introduces the regression-based inequality decomposition methodologies designed to identify the determinants of inequality and changes in it. Section 4 describes the data for Myanmar. Section 5 discusses the regression results, and explains the determinants of the level of, and changes in, consumption expenditure inequality using the Fields (2003) decomposition and the Yun (2006) extension of it. We draw attention to a serious flaw in Yun's methodology and suggest a method for correcting it. Section 6 draws conclusions on the implications of these findings.

2. Literature review

The literature on the decomposition of inequality measures focuses primarily on income inequality. Much but not all of this literature is also applicable to inequality in other relevant economic variables, such as consumption expenditures, the focus of this study. We shall review the main approaches that have been used and then briefly summarize the main findings of applications in developing countries.

2.1 Empirical approaches to decomposing inequality

Three approaches have dominated the literature. The first follows Pyatt, Chen, and Fei (1980) and Shorrocks (1982) in separating the incomes of individuals or households into its components, such as labor income, investment income, and transfer income. It then estimates the contributions these components make to total income inequality. A second approach, following the earlier studies of Bourguignon (1979), Shorrocks (1980 and 1984), and Cowell (1980, 2000), separates the population into sub-groups and estimates the contributions that within-group and between-group differences make to the observed inequality.

Both of these approaches are of value, but from the policy-oriented perspective of the present study, neither satisfactorily explains inequality or its change over time. The first describes the contribution that the various components of income make to total inequality or changes in it, but does not relate these findings to the underlying contributing factors determining these components of income, such as education, experience and gender. When the objective is to draw policy implications from the analysis, this is a crucial deficiency. Moreover, this approach is not applicable when the purpose is to explain inequality in variables other than incomes, as in the present study, where expenditure data are collected within household surveys but income data are not. Since consumption expenditure is generally considered a better measure of household welfare than income, expenditure data may be preferred in any case.

In the decomposition of population subgroups, it is possible to separate the 'between-group' attribute, due to differences in mean incomes across subgroups, and the 'within-group' attribute. However, these approaches cannot control for other factors when identifying and measuring the contribution of a particular variable. The decomposition by subgroups describes how 'between' and 'within' group components of income differ, but it does not explain why they differ. It can thus be considered a descriptive, rather than analytical approach (Cowell & Fiorio, 2011).

A third, regression-based approach, uses plausible explanatory variables, both discrete and continuous, based on economic and social theory. The main attractiveness of this method is that it can assess the contribution of fundamental variables to total inequality, controlling for other variables, and that endogeneity problems such as reverse causality can potentially be addressed (Wan 2002). Oaxaca (1973) and Blinder (1973) pioneered this approach. They focused on the difference in mean income between two groups, which has been used mainly to analyze discrimination in the labor market. Juhn, Murphy and Pierce (1993)—subsequently JMP—extended this approach so that the decomposition depends on the difference in the entire income distribution between two groups rather than just the difference in mean income.

Bourguignon, Fournier, and Gurgand (2001) relaxed the requirement of a linear income-generating function imposed by JMP.

Morduch and Sicular (2002) applied Theorem 1 of Shorrocks (1982). Wan (2002) criticized this approach on several grounds, including the use of the logarithm of income as the dependent variable. Wan states that only additively decomposable measures of inequality can be used correctly with the Morduch and Sicular approach. This means the Generalized Entropy class, a category that does not include the widely-used Gini coefficient. Wan (2004) also argued that the scaling of variables may distort the results of inequality decomposition though it may not affect the rankings of income.

Fields (2003) proposed a simple regression-based decomposition to analyse labour earnings inequality in the United States according to its determinants, using variables such as experience, schooling, occupation, industry, gender, race and region of residence. Fields points out that the resulting method can be applied to the decomposition of inequality in any continuous dependent variable and not just wages (or incomes). He argued that exogenous variables in the regression function can be regarded in the same way as 'factors' in the earlier inequality literature. The resulting decomposition is exact (adds to one) and allows analytical interpretations. Fields' method provides a connection between statistical analysis of inequality and economic theories of its causes. Fields' derivation uses Theorem 3 of Shorrocks (1982) and his expression for factor shares applies to any inequality indices which satisfy Shorrocks' six assumptions, thereby including Gini, variance, coefficient of variation, and the Generalized Entropy (GE) class, such as Theil's T index GE(1).

Naschold (2009) points out that Fields' regression-based decomposition approach has a number of advantages. First, it is not limited to predetermined income sources and may be applied to the decomposition of inequality in any economic variable into the factors explained by a regression model. Second, the method facilitates combining the relative factor inequality weights of a subset of variables into a single group factor and is flexible in the way these groups are defined. Third, the regression constant does not affect inequality as its relative factor inequality weight is zero by definition. Most usefully for the present study, Fields' method facilitates a distinction between explaining the level of inequality, on the one hand, and explaining differences in inequality between regions, countries or time periods, on the other.

A sophisticated extension of the regression-based approach is the Shapley value decomposition advocated by Shorrocks (1999), which decomposes inequality totally, accounting for all parts of the income-generating equation, including the error term. It also derives the marginal effects that the contributing factors have on inequality. Sastre and Trannoy (2002) show that the Shapley value decomposition is symmetric in all variables and is sensitive to the inequality index applied. On the other hand, Cowell and Fiorio (2011, p. 511) contend that "despite its internal consistency and attractive interpretation, the Shapley value decomposition in empirical applications raises some dilemmas that cannot be solved on purely theoretical grounds". Charpentier and Mussard (2011, p. 531) argue that "the results derived from the Shapley value are either different or identical to traditional decomposition techniques. They cannot be better." However, Chantreuil and Trannoy (2013) propose methods for decomposing inequality measures by income sources based on the Shapley value, which seemingly deal with these objections.

Finally, Yun (2006) proposes an integration of the approaches of Fields (2003) and JMP (1993) and argues that the result combines the advantages of both. The Fields (2003) approach has the advantage of decomposing the contribution of a factor to the change in inequality. The approach of JMP (1993) has the advantage of decomposing the contribution into coefficients (price), characteristics (quantity) and non-observable effects (residuals) by using an auxiliary equation. But it does so only at the aggregate level. Yun proposes a method for decomposing each of the individual factor components identified by Fields into coefficients and characteristics components, a distinction with potentially important policy applications. He applies the resulting decomposition to the changes in earnings inequality in the United States, 1969-99. A limitation of Yun's method, as he acknowledges, is that it is limited to the variance of logarithms (log-variance) as the inequality index.[6] His method cannot be applied to the wider range of inequality measures used in Fields (2003), or to the percentile differences in log-earnings (e.g. 90–10, 90–50, and 50–10) used in JMP (1993).

2.2 Findings of empirical studies on developing countries

Several empirical applications of the regression-based inequality decomposition approach are found in the literature, attempting to explain either the level of inequality at a particular time or changes in it, or both. The findings are diverse. Most of these studies are country-specific and several have focused on inequality in rural China since that country's agricultural reforms

[6] Some authors have criticised the use of the log-variance measure. Barrett, Crossley, and Worswick (2000) argue that "the variance of log earnings is scale dependent and therefore sensitive to the choice of reference year prices" (p.118). Jolliffe and Krushelnytskyy (2000) contend that the variance of logarithms does not satisfy the 'Principle of transfers' when the transfer is between two particularly rich observations. Foster and Ok (1999) show that the variance of the logarithms is not always consistent with the Lorenz ordering.

in the late 1970s, because the reforms led to both rapid economic growth and rising inequality. Morduch and Sicular (2002) decompose the inequality in the incomes of 259 farm households of 16 villages in China from 1990 to 1993. Their study decomposes the Theil-T index, squared Coefficient of Variation / Variance, alternative Coefficient of Variation, and the Gini coefficient to quantify the sources of the inequality based on the variables which are grouped into regional segmentations, human capital accumulation, and political variables. Their findings show that in all decompositions the contributions of spatial characteristics are large whereas the contribution of political variables is relatively small.

Wan and Zhou (2005) combine both the Shapley value framework of Shorrocks (1999) and the regression-based decomposition proposed by Morduch and Sicular (2002) to examine the determinants of changes in income inequality in rural China, using household-level datasets for 1995–2002. They find that while geography was the prominent factor in explaining the level of income inequality, it was less important in contributing to the changes of total inequality. Variation in capital input was a major factor in explaining the changes in income inequality.

Wan (2007) finds that location and location-related factors were the key contributors to the level of regional inequality in rural China, but the percentage contribution declined over time. Molini and Wan (2008) report that between 1993 and 1998 in rural Vietnam, the contributions of land, credit access, and ethnicity to total inequality decreased while those of education, physical capital, labour and community infrastructure increased. Using panel data from rural Pakistan, Naschold (2009) finds that "land ownership is key to explaining the level of inequality, but not its changes. In contrast, higher education drives changes, but not the level of inequality. Household location affects both, reflecting growing differences in market access across regions" (p.746).

Gunatilaka and Chotikapanich (2009) apply three regression-based approaches to decomposition—the Fields method, the Shapley value decomposition and the Yun approach—when exploring Sri Lanka's expenditure inequality. Their study finds that the rich enjoyed faster expenditure growth, resulting in increased inequality. They conclude that the change in inequality occurred mainly because of differential access to infrastructure, education, and occupation status. Demographic factors such as ethnicity and spatial factors contributed very little.

Cain, Hasan, Magsombol, and Tandon (2010) apply the regression-based decomposition developed by Fields (2003), to consumer expenditure surveys from India to investigate poverty and inequality within 17 major states in 1983, 1993, and 2004. Their explanatory variables are age, gender, social group, production sector, occupation, level of education, and state of residence. Educational attainment of the household head was the single most important factor driving inequality increases, as measured by the Gini coefficient. Finally, in a study on India using a regression-based inequality decomposition using Atkinson's index, Bigotta, Krishnakumar, and Rani (2015) conclude that education, household size, employment status and regional differences are the largest contributing factors to income inequality in both rural and urban areas.

The diversity of experience in determining the driving forces behind rising or declining income (expenditure) inequality across countries suggests that there is no single satisfactory explanation for these trends. The analysis must be conducted on a country-specific basis. The frequent finding that 'location' is an important explanatory variable may reflect deficiencies in the underlying data. Location presumably serves as a proxy for fundamental variables that are relevant for explaining inequality and which vary between regions but which are not explicitly captured in the survey data. In the case of Myanmar, despite great public interest in the subject of inequality, no study has yet attempted systematic analysis of either its level or its changes over time.

2.3 The method chosen for this study

Although the literature does not reach complete agreement on the best way to identify the determinants of changes in inequality, based on the above review, Fields' regression-based decomposition approach seems the most suitable for the present study. The Fields approach has appealing analytical features and it is relatively straightforward to program computationally. From a policy perspective, Fields' approach has a significant weakness. Suppose it is found that, say, 'education' is a factor that drives changes in inequality. We would like to know whether this effect arises because of changes in years of schooling in particular categories (household characteristics) or because of changes in the impact that years of schooling has on consumption outcomes (coefficients). The 'synthesis approach' developed by Yun (2006) promises to address that question. We shall therefore apply the Fields method, along with its proposed extension by Yun.

The Shapley regression-based decomposition approach is attractive but computationally intensive. The number of variables to be included in the regression model needs to be limited, as each variable must have its marginal impacts estimated. In most applications, the number of explanatory variables is limited no more than around 30, for example, Gunatilaka and Chotikapanich (2009) used around 25-30 variables, far fewer than the number of interest in the present study. While the capacity of the Shapley approach to derive the marginal inequality effects of explanatory variables is of analytical interest, from the perspective of the present study, this feature is not relevant.

The regression-based approach, including Fields' decomposition, has important limitations. It draws upon the data contained in cross-sectional surveys and its capacity to explain the level and changes in inequality is constrained by the explanatory power of the independent variables that happen to be included in the survey data. If relevant explanatory variables are not included, their contribution will be captured in the residual term. Regression-based approaches typically explain at most around half of the variation in the dependent variable, as in other studies (Gunatilaka & Chotikapanich, 2009, for example). This is troublesome in itself and there is scope for the omission of important explanatory variables to bias the estimated effects of explanatory variables that are included.

3. Regression-based decomposition of differences in inequality

We wish to decompose the observed difference in measured inequality between two distributions of household real expenditures, called *A* and *B*, for the purpose of understanding better the causes of these differences. For convenience, we shall assume that these two distributions represent the same households in two time periods. They may also represent two groups, two regions or two countries, though in those cases the households included could not be the same.[7] We consider a measure of inequality I defined on the distribution of household real expenditures per person or per adult equivalent,

$$I^t = I(E_1^t, E_2^t, \dots, E_H^t), \quad i = 1,2,\dots,H; \quad t = A,B, \qquad (1)$$

where E_i^t denotes the real expenditure of household *i* at time *t*. We wish to decompose the difference between I^A and I^B.

[7] Most applications compare the observed distributions of income or expenditures in two time periods, but both Fields (2003, p. 2 and note 18, p. 31) and Yun (2006, p. 128) stress that the two distributions may also represent different groups or countries. In his exposition of the JMP and Yun methods Fields (2003, pp. 15-17), citing a 2002 working paper version of Yun's study, uses the notation 2 and 1 where Yun uses *A* and *B*, respectively. Fields describes 1 as the 'base distribution' and 2 as the 'comparison distribution'. He decomposes the difference in log-variance $V^{k2} - V^1$, whereas Yun decomposes $V^A - V^B$. In the case of two time periods, neither Fields nor Yun state which period is first.

3.1 The Fields decomposition

Following Fields (2003), we define the natural logarithm of the level of real expenditure per adult equivalent within household i at period t (or for group or country t), $\ln E_i^t$. We assume that the number of households, H, is the same for each t. Now consider a vector of J available explanatory variables for each household i, $x_i^t = (x_{i1}^t, x_{i2}^t, \ldots, x_{iJ}^t)$, which might be used to account for the variation in the level of $\ln E_i^t$, or for changes in that variation. Suppose we have estimated the regression model, subsequently called an expenditure-generating function,

$$\ln E_i^A = \beta_0^A + \sum_{j=1}^{J} \beta_j^A x_{ij}^A + \varepsilon_i^A, \quad i = 1,2, \ldots, H \tag{2}$$

and

$$\ln E_i^B = \beta_0^B + \sum_{j=1}^{J} \beta_j^B x_{ij}^B + \varepsilon_i^B, \quad i = 1,2, \ldots, H, \tag{3}$$

where β_0^t denotes the estimated intercept term and β_j^t, where j = (1, 2, …, J) are the estimated coefficients corresponding to the J explanatory variables and ε_i^t are the estimated residuals, (t = (A,B)).

For notational convenience, for each t, we shall write X^t for the matrix containing the explanatory variables for all households, $\beta^t = (\beta_0^t, \beta_1^t, \beta_2^t, \ldots, \beta_J^t)$ for the vector of J+1 coefficients (including the intercept term), and $\varepsilon^t = (\varepsilon_1^t, \varepsilon_2^t, \ldots, \varepsilon_H^t)$ for the vector of residuals for each household estimated in (2) and (3). The RHS of (2) and (3) are now each fully described by ($\beta^A, X^A, \varepsilon^A$) and ($\beta^B, X^B, \varepsilon^B$), respectively.[8]

Fields shows that the share of the log-variance of E_i^t that is attributable to the j^{th} explanatory variable is what he calls its 'relative factor inequality weight', given by s_j^t, (j = 1, 2, …, J+1), where, again for convenience, we define $Z_{ij}^t = \beta_j^t x_{ij}^t$, where

$$s_j^t = \frac{cov[Z_{ij}^t, \ln E_i^t]}{\sigma^2[\ln E_i^t]} = \frac{\beta_j^t\, \sigma[x_{ij}^t]\, cor[\ln E_i^t, x_{ij}^t]}{\sigma[\ln E_i^t]}, \quad j = (1, 2, \ldots, J+1); \; t = (A, B), \tag{4}$$

[8] In matrix notation the expenditure-generating equations are $\ln E^t = \beta^t X^t + \varepsilon^t$, ($t$ = A,B), where $\ln E^t$ and ε^t are column vectors of H household values of $\ln E_i^t$ and estimated residuals ε_i^t, respectively, X^t is an $H \times (J+1)$ matrix of household characteristics, (with the first column consisting of all 1s, corresponding to the intercept term) and β^t is a column vector of J+1 estimated coefficients, including the intercept term.

and where the J+1th element relates to the residual, and can be denoted s^t_ε. The estimated contributions of the explanatory variables depend not only on the magnitudes of the estimated regression coefficients, β^t_j, but also on their explanatory power, captured by the terms cor[lnE^t_i, x^t_{ij}] in equation (4).

Fields then shows that the estimated s^t_j terms can be used to decompose the estimated *level* of inequality using any inequality index defined across H households $I^t = I(\ln E^t_1, \ln E^t_2, \ldots, \ln E^t_H)$, which is continuous and symmetric and for which, $I(\mu, \mu, \ldots, \mu) = 0$, where μ is the arithmetic mean of the lnE^t_i. This set of inequality measures includes most of those commonly used, including the Gini coefficient and the class of Generalized Entropy measures. Remarkably, within this broad class of inequality measures, the s^t_j terms are independent of the choice of inequality measure. That is, it is not necessary to agree on which particular inequality measure to decompose, because the method produces the same percentage decomposition into 'relative factor inequality weights' for all.

Fields adds that the estimated values of s^t_j have some additional useful properties, including first, that

$$\sum_{j=1}^{J} s^t_j = R^2(t), \tag{5}$$

where $R^2(t)$ denotes the familiar coefficient of determination of the estimated regression for year t. Second, the estimated relative factor inequality weights s^t_j can be grouped for any subset of variables g, where $j \in g$, simply by adding the s^t_j terms belonging to subset g. These aggregation properties apply not only to continuous explanatory variables, but also to non-continuous variables such as dummy variables and categorical variables (Naschold, 2009, pp. 766-767).

On the other hand, Fields points out that the inclusion of interaction terms, (for example, gender) with other variables such as experience, education, and so forth, raises a problem because the factor inequality weights given by the model cannot decompose into gender, experience and education components neatly. Ssewanyana, Okidi, Angemi, and Barungi (2004) argue that this imposes an important limitation. However, Heltberg (2002) suggests that interaction terms can likewise be included in relevant groups for constructing a single group factor inequality weight that includes interaction terms, while stressing that subgroups must be exogenous.

The difference in inequality between two periods (or groups or countries) can now be decomposed by

$$I^A - I^B = \sum_{j=1}^{J+1}(s_j^A I^A - s_j^B I^B) \quad (6)$$

$$= \sum_{j=1}^{J}(s_j^A I^A - s_j^B I^B) + (s_\varepsilon^A I^A - s_\varepsilon^B I^B),$$

where, in the final bracketed term, s_ε^A and s_ε^B denote Fields' relative factor inequality weights associated with the residuals of (2) and (3). Unfortunately, this decomposition of the difference in inequality between years (or between regions or countries) is not independent of the inequality measure being used.

3.2 The Yun synthesis

Yun (2006) proposes an approach to decomposing differences in inequality that synthesizes Fields' analysis, above, with the earlier approach due to JMP (1993). Both Fields and JMP begin with a pair of expenditure-generating equations like (2) and (3). The JMP method decomposes the total difference in inequality into three aggregate components: differences in coefficients, including the intercept (β^A and β^B); differences in household characteristics (X^A and X^B); and differences in residuals (ε^A and ε^B).

Yun's description of the JMP decomposition begins with construction of what he calls an auxiliary equation:[9]

> "Start with the earnings equation of time period *A*. First, replace the coefficients of the earnings equation of time period *A* with those of time period *B*, while keeping the individual characteristics and residuals unchanged." (Yun 2006, p. 129).
>
> This defines the auxiliary equation:

$$\ln E_i^*(\beta^B, X^A, \varepsilon^A) = \beta_0^B + \sum_{j=1}^{J}\beta_j^B x_{ij}^A + \varepsilon_i^A. \quad (7)$$

[9] Our discussion of the JMP method, especially the construction of the auxiliary equation, closely follows Yun (2006, p. 129) because Yun describes it more fully than JMP (2003, pp. 426, 428 and 429). The latter does not mention an 'auxiliary equation' or an 'auxiliary distribution'.

In our notation for the auxiliary equation $\ln E_i^*(\beta^B, X^A, \varepsilon^A)$, the asterisk signifies that this equation is counterfactual, constructed by the researcher from the regression results and raw data, rather than being directly estimated econometrically, and $(\beta^B, X^A, \varepsilon^A)$ indicates that it is constructed using the coefficients estimated from period *B*, and the household characteristics and residuals from period *A*.

The difference in inequality that we wish to decompose can now be rewritten as the identity

$$I^A - I^B = [\, I^A - I^*(\beta^B, X^A, \varepsilon^A)] + [I^*(\beta^B, X^A, \varepsilon^A) - I^B], \tag{8}$$

where $I^*(\beta^B, X^A, \varepsilon^A)$ denotes the application of inequality measure *I* to the auxiliary distribution defined by (7). Yun now combines (8) with the Fields decomposition, given by (6) above, and shows that in the specific case where the measure of inequality *I* is the log-variance, which we denote *V*, equation (8) becomes

$$\begin{aligned} V^A - V^B = & \left[\textstyle\sum_{j=1}^{J} (s_j^A V^A - \; s_j^*(\beta^B, X^A, \varepsilon^A) V^*(\beta^B, X^A, \varepsilon^A)\right] \\ & +[\textstyle\sum_{j=1}^{J} s_j^*(\beta^B, X^A, \varepsilon^A) V^*(\beta^B, X^A, \varepsilon^A) - s_j^B V^B] \\ & +[V_\varepsilon^A - V_\varepsilon^B], \end{aligned} \tag{9}$$

where V^t and V_ε^t, $t = (A, B)$, denote the variances of $\ln E_i^t$ and ε_i^t, respectively, in equations (2) and (3). Yun then states that the three square-bracketed terms of (9) are, respectively, the coefficients effect, the characteristics effect and the residuals effect, with the first two disaggregated by factors.[10]

Because the residual component of (9) does not involve the auxiliary equation and, as Yun notes, cannot be disaggregated by explanatory factors, it is intuitively helpful to subtract this term from both sides.

[10] Yun (2006) shows that in the case of the log-variance measure the first bracketed term of (8) is equal to the first bracketed term of (9) and that the second bracketed term of (8) divides into the second and third bracketed terms of (9).

$$(V^A - V^B) - (V_\varepsilon^A - V_\varepsilon^B) = \left[\sum_{j=1}^{J}(s_j^A V^A - s_j^*(\beta^B, X^A, \varepsilon^A)V^*(\beta^B, X^A, \varepsilon^A)\right] + [\sum_{j=1}^{J} s_j^*(\beta^B, X^A, \varepsilon^A)V^*(\beta^B, X^A, \varepsilon^A) - s_j^B V^B]. \quad (10)$$

The resulting equation (10) states that the *explained part* of the difference in log-variance, the total difference in inequality minus the unexplained difference (the difference in residuals), can be decomposed fully into two components: a difference-in-coefficients component, and a difference-in-characteristics component (the JMP decomposition), each of which is then disaggregated by the *J* explanatory factors (the Fields decomposition). Equivalently, because the decomposition is linear, after rearrangement this same equation states that the explained part of the difference in inequality is decomposable into the individual contributions of the *J* factors (the Fields decomposition), each of which is then decomposable into a difference-in-coefficients component and a difference-in-characteristics component (the JMP decomposition).

But something arbitrary seems to have occurred. Why was the auxiliary equation constructed in the above sequence, and does it matter? Suppose we had started with distribution *B* instead of *A*, reversing *A* and *B* throughout Yun's account, following exactly the same steps. The auxiliary equation would then be constructed from the coefficients of *A* and the individual characteristics and residuals of *B*. The auxiliary equation would then be

$$\ln E_i^*(\beta^A, X^B, \varepsilon^B) = \beta_0^A + \sum_{j=1}^{J} \beta_j^A x_{ij}^B + \varepsilon_i^B. \quad (11)$$

Equation (10) would now read

$$(V^B - V^A) - (V_\varepsilon^B - V_\varepsilon^A) = \left[\sum_{j=1}^{J}(s_j^B V^B - s_j^*(\beta^A, X^B, \varepsilon^B)V^*(\beta^A, X^B, \varepsilon^B)\right] + \left[\sum_{j=1}^{J} s_j^*(\beta^A, X^B, \varepsilon^B)V^*(\beta^A, X^B, \varepsilon^B) - s_j^A V^A\right]. \quad (12)$$

The LHS of (12) is exactly the negative of the LHS of (10), so the RHS terms of (10) and (12) must also add to the same total, with opposite signs. But after adjusting for this sign difference do they necessarily imply the same division of the explained difference in log-variance into difference-in-coefficients and difference-in-characteristics components?[11]

[11] Yun (2006) does not raise this question and in his exposition of Yun's synthesis, Fields (2003) similarly does not consider it.

Simply from inspection of these equations, the answer is not obvious. In Section 5 we use empirical data for Myanmar to show that the answer is "no". Using Yun's sequence (*A*, *B)*, rather than (*B, A)*, or *vice versa*, is obviously arbitrary and should not affect the decomposition. Unfortunately, it does.

4. Data for Myanmar

4.1 Data sources

The data from the 2004/05 and 2009/10 Integrated Household Living Conditions Surveys, subsequently referred to as the 2005 and 2010 surveys, are described in detail in IDEA and IHLCA (2007a, b and c) and IHLCA (2011a, b and c). The full surveys covered 18,634 and 18,609 households, respectively, using the same survey designs and methodology. The 2010 survey retains a panel of 9,102 households contained within the 2005 sample, roughly half of the total samples in each of these years. Descriptive statistics on the panel sample and the full sample are summarized for each year in Table 1 and a fuller summary is provided in Appendix Table A1. Because our subsequent analysis will draw upon the panel component of the data set, we focus the discussion on this part of the full sample.

4.2 The dependent variable: household real expenditure per adult equivalent

The natural logarithm of household consumption expenditure per adult equivalent is the dependent variable in the regression-based decomposition analysis. The 'per adult equivalent' part of this definition followed the procedures recommended in Deaton and Zaidi (2002). In the IHLCA calculations of household expenditures, two important omissions were expenditures on health-related items and consumer durables. The proportion of expenditure allocated to these items may be a function of the level of household income, so measured inequality and changes in it could be affected by their omission. The raw household survey data included actual expenditures on both of these two items and quantities of consumer durables owned by the household.

The present study amends the data used by the IHLCA team to include these expenditures.[12] Health expenditures and the user costs of durable goods per year per adult

[12] This amendment to the data and a comparison with the unamended IHLCA data are summarised in Appendix Table A2.

equivalent are deflated by a Paasche index of purchaser prices within the survey periods estimated by the IHLCA team and adds them to non-food consumption expenditures for 2005 and 2010. In the case of consumer durables, our calculations convert the value of items owned or purchased by the household into an annual rental equivalent[13]. The amended expenditure data are compared with the IHLCA published data in Appendix Table A2. The amended expenditures exceed the IHLCA expenditures in all cases, but the differences are not particularly large.

4.3 Explanatory variables

Explanatory (independent) variables include characteristics that seem likely to influence household expenditure and which are not themselves a function of expenditure. They are listed in Table 2.

Characteristics of household heads: These are age, gender, and ethnicity of the household head. The relationship between age and income (as proxy by expenditure) is expected to have an inverted-U shape, consistent with theories of life-cycle earnings and with empirical findings elsewhere (Knight & Song 1999; Gustafsson & Shi 2001). Several empirical studies show that well-being is U-shaped for age (Blanchflower & Oswald, 2004, 2008). These non-linear possibilities are captured by using the age of the household head and its square.

Ethnic differences: Bamar (ethnic Burmese) is the major ethnic group, comprising about 84% of all Myanmar nationals. Its effect could be large as most ethnic minorities (non-Bamar) live in rural and mountainous areas. Human capital theory (Schultz, 1963) indicates that income is a function of education and experience, but years of work experience is not included in the IHLCA survey questionnaires. In this study, years of non-agricultural business in operation of household heads is included as a proxy for work experience.

Characteristics of household members: These are household size, the proportion of household members within different age groups, occupation status, and the economic sector of the household head's occupation. To account for possible non-linearities in the impact of household size a squared household size term is also included. The age structure of the

[13] The average interest rate used for this purpose was just under 4% per annum.

household members can also matter as expenditures per adult equivalent may vary, depending on the stage of the life cycle. Age structure is divided into the six subgroups: the proportion of members below 6 years of age; between 6 and 10; 11 and 15; 16 and 65; and older than 65 years of age. The reference group is the proportion of household members above aged 65.

Occupation: Occupation is a categorical variable and the proportion of members with the main occupation in the last 7 days is divided into nine subgroups listed in Appendix Table A1. The proportion of working-age adults (aged from 15 to 64) in the household working in various sectors in the last six months is also a categorical variable and the sector is reclassified into primary, secondary and tertiary sectors.

Education of household members: Human capital is captured by six variables: the proportion of household members with tertiary education, upper secondary education, lower secondary education, primary education and illiterates (the reference group is unclassified level of education), since the whole household is likely to benefit from the education levels of individual household members.

Health Indicator: The independent variables include the proportion of household members who were sick in the last 30 days, to control for ability to work.

Land ownership and access along with the cultivation of crops: The level of expenditure and inequality in it may be affected by access to physical assets, particularly land. The study thus includes the area of land owned, access to irrigated and unirrigated land. Because land varies across the country and cultivated crops also vary, dummy variables of cultivated crops are included to capture the effect of crop diversification on expenditure, using six major commodity categories.

Location and regional variables: The Union of Myanmar is divided into 17 states/regions (divisions). Regional effects are potentially important for the determination of consumption expenditures, reflecting a relationship with fixed natural resources (such as jade mining versus teak forest), market access, and infrastructure. To account for these factors, four regional dummies are used: Dry Zone, Delta region, Coastal areas, and the Hills region (the reference group). In 2010 a severe cyclone, known as Cyclone Nargis, negatively affected the Delta areas of Myanmar, and its impact is incorporated in its dummy variable**.**

Community characteristics: The IHLCA surveys in 2005 and 2010 also provide information on community level characteristics at the village/ward level. These include: topography of the village/ward such as inland plains, hills, mountains, deltas and valleys; distance to the nearest town/township; bank or financial services; hospital or rural health center, maternity hearth care; primary and monastic school, and lower and upper secondary schools to reflect differences in community services available to the village tracts and wards.

Infrastructure variables: Infrastructure variables represent productive economic assets provided by the government. Road density is the ratio of the length of the state and region's total road network to the state and region's land area. The road network includes all roads in the states and regions: bituminous, gravel surfaced, and earth roads reported. Road density is calculated as miles of total road length (in states and regions) per 100 square miles of land area. Proxy variables to assess the level of infrastructure in the community are number of months that cars/four-wheel drives can be driven in the community in a year, common modes of transportation, and electricity and water supply.[14]

As Naschold (2009) points out, some of the independent variables described above may be considered endogenous, especially over long time periods. For example, household size may be affected by migration decisions, and the level of household assets and education depends on the household's decisions. Nevertheless, it is not clear what instruments could be used. In this study, all household characteristics are treated as exogenous.

[Table 2 here]

5. Empirical findings for Myanmar

5.1 Fields' decomposition

The regression results for panel households are reported in Appendix Table A3. In Table 2 the first two columns use the estimated regression results to compute Fields' 'relative factor inequality weights', the s_j^t terms in equation (4), for 2005 and 2010, respectively.[15] The next

[14] All of the data described in this paragraph were obtained for 2005 and 2010 from the *Statistical Yearbook 2011*, Central Statistical Organization, Yangon, 2012.

[15] The full regression results for panel households are reported in Appendix Table A3.

four columns similarly use equation (6) to apply these weights to the decomposition of the differences between years in the Gini coefficient and three commonly used members of the Generalized Entropy class (α = 0, 1 and 2).

Regarding levels of inequality in each year (columns (1) and (2) of Table 2), the proportion of total variation explained by the regressions, as reflected in the R^2 coefficient, is around 35 per cent. This is in line with earlier applications of the regression-based approach (Gunatilaka & Chotikapanich, 2009; Deng & Shi, 2009), but it means that almost two-thirds of the variation is unexplained. The results indicate that in both years higher inequality is associated with higher levels of: education; employment in white-collar, services, and blue-collar manufacturing jobs, but not agricultural jobs; land ownership; residence in the Delta region; larger household sizes; and local provision of water and electricity infrastructure**.**

Our primary interest is in explaining the overall decline in inequality (the differences question (columns (3) to (6) of Table 2). The proportion of total variation explained is between 45 and 70 per cent. Because each of the four inequality measures declined, a positive contribution to that outcome, reflected in a positive number in these four columns, means that the explanatory variable concerned is estimated to have contributed to the observed *decline* in inequality. A negative number means that the explanatory variable is estimated to have *increased* inequality.

In the case of the Gini coefficient, 53 per cent of the observed decline was explained by the contributing factors. Major contributors were improved education standards, improved employment levels within better-paid employment categories, and improved road access and water supply within low-income regions. These factors together explained 39 per cent of the total reduction in the coefficient, or 73 per cent of the explained portion of that decline. Results were qualitatively similar to this for each of the GE measures except GE(2), where the explained proportion was 70 per cent.

Previous studies have often found that location is an important contributor to changes in inequality. Heltberg (2002) on Vietnam, Wan (2007) on rural China, and Naschold (2009) on rural Pakistan are examples. In our findings, residence in the Delta region was a large contributor to reduced national inequality, but for an unfortunate reason. In May 2008 Tropical Cyclone Nargis impacted primarily the Delta region, with devastating effects, including the deaths of around 138,000 people. In 2005, before the cyclone, the Delta was the best-off region of Myanmar, and remained the best-off afterwards, in 2010. The cyclone reduced real expenditures in that region, but had no direct impact on other regions, thereby reducing inequality between Myanmar's major regions.[16] The cyclone's impacts within the directly affected regions were much less significant. Treating the dummy variable associated with the Delta region as a measure of the cyclone's impact, it reduced measured inequality in Myanmar by 43 per cent of the total observed decline.[17] This impact is analysed in greater detail in Warr and Aung (2019).

Not all explanatory variables contributed to the decline in inequality. Overall, infrastructure impacted in the opposite direction, primarily because improved infrastructure, such as upgraded (bituminous) roads and improved electricity supply went to areas that were already better off, thereby increasing inequality.[18] The effect of household size is more complex. Average household sizes declined between 2005 and 2010. This contributed to a rise in measured inequality because the estimated s_j^t term associated with household size increased between these years, from 5.6 to 7.9. Household size became a more significant contributor to a high level of inequality despite the reduction in mean household size.

The estimated inequality impact of a particular policy variable is not necessarily intrinsic to that policy measure, but also reflects the way the policy was implemented. For example, the impact of bituminous road upgrading was found to be inequality-increasing. This reflects the fact that the areas benefiting from the road upgrades implemented between 2005 and 2010 tended to be better-off areas to begin with. This investment strategy may or may not have made logistical sense, in terms of the overall road network—upgraded roads have to begin somewhere. Although these particular investments were inequality-increasing during the period of our data, the findings do not necessarily mean that public investment of this kind is inherently inequality-increasing. Different allocative decisions, more or less pro-poor than those that actually occurred, could have been made and may be made in the future with this and other forms of public investment.

5.2 Yun's synthesis

We now apply Yun's synthesis to the Myanmar data and test its robustness with respect to arbitrary changes in the sequence of years. Throughout this discussion and in Tables 3 and 4, periods *A* and *B* denote 2005 and 2010, respectively. The first three columns of Table 3 show

---

[16] In our panel sample, between 2005 and 2010 mean real expenditures per adult equivalent declined by 10.7 per cent in the region affected by Cyclone Nargis but increased by 6.9 per cent in the rest of the country. The national mean increased by 2.3 per cent.

[17] See Table 2, factor 10, Delta region, column 2.

[18] This is not to deny that improved roads and electricity supplies raise mean incomes by raising productivity. The discussion here is about the impact on inequality, not the impact on mean incomes.

information required for the decomposition described by Yun (2006, pp. 129-132), as in equation (9). The first two columns relate to the two observed distributions of log expenditures, as in equations (2) and (3). The third column, labeled (*A*, *B*) relates to Yun's auxiliary distribution of log expenditures, constructed as in equation (7). The first 11 rows of these three columns use Fields' equation (4) to calculate the three sets of relative factor inequality weights, calculated from each of these three distributions.

[Table 3 here]

The fourth column of Table 3, labeled (*B*, *A*), shows comparable information to the third column, but reversing the sequence of the two years in the Yun decomposition from (*A*, *B*) to (*B*, *A*), where *A* and *B* still denote 2005 and 2010, respectively. The auxiliary distribution is now generated by equation (11). The relative factor inequality weights are calculated from this auxiliary distribution, again using equation (4). A key finding is that the properties of all four distributions, including the two auxiliary distributions, are different. This includes the 11 relative factor inequality weights, the totals of these weights (third last row), their means (second last row) and their variances (last row).

Table 4 now uses this information to show the findings of the Yun decomposition and to test its robustness. The left side (first six columns) shows Yun's decomposition of $V^A - V^B$, described by equation (9), using the information contained in the first three columns of Table 3. The first and third columns show the coefficients and characteristics decomposition for each factor and the 12th row 'Aggregate' shows the sum of these components. The coefficients component explains about 35 per cent of the total difference in log-variance and the characteristics components explains about 9 per cent. Combined, these two 'explained' components account for 44 per cent of the total difference in log-variance.

The right side of Table 4 (last six columns) shows the comparable information, but reversing the sequence of the two years, as in the fourth column of Table 3. Because $(V^A - V^B) = -(V^B - V^A)$, the total change in log-variance is of course the same on the two sides of the table, but with opposite signs. The size of the 'unexplained' residual component is also the same, but with opposite sign.[19] Therefore, the aggregate size of the remaining 'explained' component—coefficients plus characteristics—must also be the same, but with opposite sign. But the findings show that its division into coefficients and characteristics components is

[19] As equation (9) shows, the size of the residuals component does not involve the auxiliary equations.

substantially different on the two sides of Table 4, in aggregate and for every individual factor. The aggregate contribution of the characteristics component now exceeds the coefficients component. The findings show that the arbitrary choice between the sequences (2005, 2010) and (2010, 2005) can substantially affect the results of the Yun decomposition.

[Table 4 here]

Our empirical findings for Myanmar show conclusively that the two sequences do not necessarily produce the same decomposition. When would the two versions of Yun's decomposition give the same answers? Consider the coefficients and characteristics effects for factor *j*. Using sequence (*A*, *B*), the sum of these two components would be, from equation (10), $\left(s_j^A V^A - s_j^B V^B\right)$. Under sequence (*B*, *A*), from equation (12) their sum must be the same as that, but with the opposite sign. Therefore, if the coefficients effects are the same, after adjusting for the sign reversal, the characteristics effects must also be the same, and *vice versa*. It is sufficient to focus on either one. The coefficients effects will be the same if and only if

$$(s_j^A V^A - s_j^*(\beta^B, X^A, \varepsilon^A) V^*(\beta^B, X^A, \varepsilon^A)) = -(s_j^B V^B - s_j^*(\beta^A, X^B, \varepsilon^B) V^*(\beta^A, X^B, \varepsilon^B)),$$

or

$$\left(s_j^A V^A + s_j^B V^B\right) = (s_j^*(\beta^B, X^A, \varepsilon^A) V^*(\beta^B, X^A, \varepsilon^A) + s_j^*(\beta^A, X^B, \varepsilon^B) V^*(\beta^A, X^B, \varepsilon^B)). \quad (13)$$

By adding across factors *j*, a comparable condition is obtained for the aggregate of the coefficients effects to be the same.[20] In the case of our Myanmar data, this condition is not satisfied for any of the 11 factors or for the aggregate.

Why does the sequence matter? The simple answer lies in the structure of the expenditure-generating functions. In equations (2) and (3), the coefficients (β_j^t) and household characteristics (x_{ij}^t) are multiplied together in the determination of expenditure. Therefore, the effect of changing one depends on the level of the other. This would not be true if they entered the determination of expenditure additively, as with the residual.[21] Now consider the size of the aggregate coefficients effect. By inspection of equation (10), the Yun (2006) decomposition measures it by evaluating the change in log-variance resulting from the observed change in the

[20] Focusing instead on the characteristics effect yields the same necessary and sufficient condition (13).
[21] Recall that if $z = x + y$, then $dz = dx + dy$. But if $z = xy$, then $dz = ydx + xdy$.

coefficients, but holding the household characteristics constant at their period *A* values. From equation (12), our reversal of the sequence of years evaluates the same changes in coefficients, but instead holds the household characteristics constant at their period *B* values.[22] It is hardly surprising that they can give different results. Intuitively, the greater the difference between the period *A* and period *B* values of the household characteristics, the greater the scope for the two sequences to give different values for the coefficients effect and therefore different decompositions. The two sequences are equally arbitrary and potentially misleading. Neither is satisfactory.

Our proposed solution is to derive the size of the coefficients effect holding the household characteristics constant at the arithmetic means of their period *A* and period *B* values—their arithmetic mid-point. We do the same thing for the characteristics effect. Fortunately, the linearity of the system permits a computationally convenient and equivalent short-cut. It is to compute both decompositions, as described above, and then to calculate the arithmetic mean of the estimated coefficients and characteristics effects, after adjusting for the sign reversal between them, producing their mid-point. This is done in Table 5. The coefficients effects shown in the first and seventh columns of Table 4 are added and divided by 2, after adjusting for the sign difference.[23] The same is done for the characteristics effects (third and ninth columns, respectively). This produces a decomposition of the difference in inequality that is independent of sequencing and free of the distorting arbitrariness that we have demonstrated.

[Table 5 here]

6. Conclusions

This paper analyzes changes in household consumption inequality in Myanmar, drawing on a large and unique household-level panel data set, covering the years 2005 and 2010. A significant reduction in measured inequality occurred between these years. The Fields (2003)

[22] Similarly, in measuring the size of the characteristics effect the Yun (2006) decomposition holds the coefficients constant at their period *B* values, whereas our alternative sequencing substitutes their period *A* values.

[23] In calculating Table 5 we did this by reversing all the signs of the coefficients and characteristics effects shown in Table 4 for the sequence (*B*, *A*). These were then added to the corresponding effects shown for sequence (*A*, *B*) and the result was divided by 2. If no adjustment is made for the sign reversal, the coefficients effects for one sequence must be subtracted from the other and the result divided by 2, and the same with the characteristics effects.

regression-based approach is used to calculate Fields' relative factor inequality weights and these are then used to decompose changes in standard inequality measures—the Gini coefficient and Generalized Entropy measures—into explanatory factors such as geographical location, occupation and education. The findings show that spatial variables, including regional location dummies and distance to public services, were the largest single contributor to declining inequality. Other important contributors were educational attainment and the occupations of household members. Road development operated in the opposite direction, at least in the short-run, because it favoured regions that were already better-off.

Using Fields' decomposition, the explained component of the decline in inequality accounts for 53% of the total observed decline in the Gini coefficient. Region-specific variables account for 40% of the total decline and the share of household members with different occupations contributes a further 16%. The educational attainment of working-age adults contributes about 18%. These findings are roughly in line with earlier studies covering other countries, including Vietnam (Heltberg 2002), rural China (Wan, 2007), rural Pakistan (Naschold, 2009) and Sri Lanka (Gunatilaka and Chotikapanich 2009).

Locational effects may partly reflect deficiencies in the underlying data. They can capture explanatory variables not recorded in the household level data but associated with location. For example, a devastating cyclone negatively impacted the best-off region of the county in 2008. Its impact reduced inequality at the national level by reducing consumption differences between regions. All researchers studying Myanmar are aware of this particular 'location effect', although looking only at the household level data sets would not have revealed it. Other, apparent 'location effects' may not be so readily identified. Examples may include access to local markets, local differences in financial and health services, and schools. Fixed natural resources (such as jade mines or teak forests) can also be correlated with region dummies, along with local topography. But because these location-correlated variables may not be captured fully within the household data set, their impact is captured only via 'location'.

We also apply the Yun (2006) synthesis of the decomposition methods developed earlier by Fields (2003) and Juhn, Murphy and Pierce (1993). Yun's synthesis is important. In principle, it separates the explained part of the contribution of each factor, as described above, into a change in 'characteristics' and a change in the 'coefficients' component. This could help answer policy-relevant questions such as: to what extent does the contribution to the observed change in inequality deriving from a 'factor' like education, as obtained from Fields' decomposition, occur because of changes in the numbers of years of schooling (the characteristics component) or because of changes in the degree to which years of schooling affects expenditures (the coefficients component)? However, we demonstrate that the Yun decomposition suffers from a previously unrecognized problem of arbitrary sequencing that can substantially affect its findings. If the decomposition is applied with the sequence of years, groups or countries arbitrarily reversed, a qualitatively different decomposition can be obtained. A simple solution is proposed and demonstrated.

There are important limitations to this form of empirical study. In particular, the value of the decomposition methods we have discussed rests on the relevance for inequality of the data collected in the underlying household surveys. This is most evident from the fact that less than half of the observed difference in inequality is typically explained and also from the large explanatory role played by 'location' in this and in previous studies. If unobserved variables, correlated with location, partly explain the findings, what are these unobserved variables, does their omission significantly bias the estimated effects of the included variables, and are they potentially changeable through policy interventions?

References

Barrett, G. F., & Pendakur, K. (1995). The asymptotic distribution of the generalized Gini indices of inequality. *The Canadian Journal of Economics/Revue Canadienne D'Economique*, *28*(4b), 1042–1055.

Barrett, G. F., Crossley, T. F., & Worswick, C. (2000). Consumption and income inequality in Australia. *Economic Record*, *76*(233), 116–138. http://dx.doi.org/10.1111/j.1475-4932. 2000.tb00011.x.

Bigotta, M., Krishnakumar, J., & Rani, U. (2015). Further results on the regression-based approach to inequality decomposition with evidence from India. *Empirical Economics*, *48*(3), 1233–1266. http://dx.doi.org/10.1007/s00181-014-0819-5.

Blanchflower, D. G., & Oswald, A. J. (2004). Well-being over time in Britain and the USA. *Journal of public economics*, *88*(7), 1359–1386. http://dx.doi.org/10.1016/S0047-2727 (02)00168-8

Blanchflower, D. G., & Oswald, A. J. (2008). Is well-being U-shaped over the life cycle? *Social science & medicine*, *66*(8), 1733–1749. http://dx.doi.org/10.1016/j.socscimed.2008. 01.030.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *The Journal of Human Resources*, *8*(4), 436–455. http://dx.doi.org/10.2307/144855.

Bourguignon, F. (1979). Decomposable income inequality measures. *Econometrica*, *47*(4), 901–920. http://dx.doi.org/10.2307/1914138.

Bourguignon, F. (2015). *The globalization of inequality*. Princeton, NJ: Princeton University Press.

Bourguignon, F., Fournier, M., & Gurgand, M. (2001). Fast development with a stable income distribution: Taiwan, 1979–94. *Review of Income and Wealth*, *47*(2), 139–163. http://dx.doi.org/10.1111/1475-4991.00009.

Cain, J. S., Hasan, R., Magsombol, R., & Tandon, A. (2010). Accounting for inequality in India: Evidence from household expenditures. *World Development*, *38*(3), 282–297. http://dx.doi.org/10.1016/j.worlddev.2009.11.014.

Central Statistical Organization. (2005). *Selected monthly economic indicators*. Nay Pyi Taw: Author, Ministry of National Planning and Economic Development.

Central Statistical Organization. (2010). *Selected monthly economic indicators*. Nay Pyi Taw: Author, Ministry of National Planning and Economic Development.

Central Statistical Organization. (2012). *Statistical yearbook 2011*. Nay Pyi Taw: Author, Ministry of National Planning and Economic Development.

Chantreuil, F., & Trannoy, A. (2013). Inequality decomposition values: The trade-off between marginality and efficiency. *The Journal of Economic Inequality*, *11*(1), 83–98. http://dx.doi.org/10.1007/s10888-011-9207-y.

Charpentier, A., & Mussard, S. (2011). Income inequality games. *The Journal of Economic Inequality*, *9*(4), 529–554. http://dx.doi.org/10.1007/s10888-011-9184-1.

Cowell, F. A. (1980). On the structure of additive inequality measures. *The Review of Economic Studies*, *47*(3), 521–531.

Cowell, F. A. (2000). Measurement of inequality. In A.B. Atkinson & F. Bourguignon (Eds.), *Handbook of income distribution* (Vol.1, pp. 87–166). Amsterdam: North-Holland.

Cowell, F. A., & Fiorio, C. V. (2011). Inequality decompositions-a reconciliation. *The Journal of Economic Inequality*, *9*(4), 509–528. http://dx.doi.org/10.1007/s10888-011-9176-1.

Davidson, R., & Duclos, J. Y. (2000). Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica*, *68*(6), 1435–1464. http://dx.doi.org/10.1111/1468-0262.00167.

Deaton, A., & Zaidi, S. (2002). *Guidelines for constructing consumption aggregates for welfare analysis* (LSMS Working Paper. 135). Washington, DC: World Bank Publications.

Deng, Q., & Shi, L. (2009). What lies behind rising earnings inequality in urban China? Regression-based decompositions. *CESifo Economic Studies*, *55*(3–4), 598–623. http://dx.doi.org/10.1093/cesifo/ifp015.

Fields, G. S. (2003). Accounting for income inequality and its change: A new method, with application to the distribution of earnings in the United States. *Research in labour economics*, *22*(1.38). http://dx.doi.org/10.1016/S0147-9121(03)22001-X.

Foster, J. E., & Ok, E. A. (1999). Lorenz dominance and the variance of logarithms, *Econometrica*, *67*(4), 901–907. http://dx.doi.org/10.1111/1468-0262.00057.

Gunatilaka, R., & Chotikapanich, D. (2009). Accounting for Sri Lanka'S expenditure inequality 1980–2002: Regression-based decomposition approaches. *Review of Income and Wealth*, *55*(4), 882–906. http://dx.doi.org/10.1111/j.1475-4991.2009.00351.x.

Gustafsson, B., & Shi, L. (2001). A more unequal China? Aspects of inequality in the distribution of equivalent income. In C. Riskin, R. Zhao, & L. Shi, (Eds.), *China's retreat from equality: Income distribution and economic transition* (pp. 44–83). New York, NY: ME Sharpe.

Heltberg, R. (2002). *Causes of inequality in Vietnam* (Mimeo). Copenhagen: Institute of Economics, University of Copenhagen.

IDEA International Institute & IHLCA Project Technical Unit. (2007a). *Poverty profile: Integrated Household Living Conditions survey in Myanmar (2009–2010)*. Nay Pyi Taw: Ministry of National Planning and Economic Development.

IDEA International Institute & IHLCA Project Technical Unit. (2007b). *Quality report: Integrated Household Living Conditions survey in Myanmar (2009–2010)*. Nay Pyi Taw: Ministry of National Planning and Economic Development.

IDEA International Institute & IHLCA Project Technical Unit. (2007c). *Technical report: Integrated Household Living Conditions survey in Myanmar (2009–2010)*. Nay Pyi Taw: Ministry of National Planning and Economic Development.

IHLCA Project Technical Unit. (2011a). *Poverty profile: Integrated Household Living Conditions survey in Myanmar (2009–2010)*. Nay Pyi Taw: Ministry of National Planning and Economic Development.

IHLCA Project Technical Unit. (2011b). *Quality report: Integrated Household Living Conditions survey in Myanmar (2009–2010)*. Nay Pyi Taw: Ministry of National Planning and Economic Development.

IHLCA Project Technical Unit. (2011c). *Technical report: Integrated Household Living Conditions survey in Myanmar (2009–2010)*. Nay Pyi Taw: Ministry of National Planning and Economic Development.

Jenkins, S. P. (2008). *Estimation and interpretation of measures of inequality, poverty, and social welfare using Stata* [PowerPoint slides]. PowerPoint presentation at North American Stata Users' Group meetings 2006, Boston MA. Retrieved from <http://repec.org/nasug2006/nasug2006_jenkins.zip>.

Jolliffe, D., & Krushelnytskyy, B. (2000). Bootstrap standard errors for indices of inequality. *Stata Technical Bulletin*, *9*(51), 28–32.

Juhn, C., Murphy, K. M., & Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, *101*(3), 410–442. http://dx.doi.org/10.1086/261881.

Knight, J. & Song, L. (1999). *The Rural-Urban divide: Economic disparities and interactions in China*. New York: Oxford University Press.

Kovacevic, M. S., & Binder, D. A. (1997). Variance estimation for measures of income inequality and polarization—the estimating equations approach. *Journal of Official Statistics*, *13*(1), 41–58.

Molini, V., & Wan, G. (2008). Discovering sources of inequality in transition economies: A case study of rural Vietnam. *Economic Change and Restructuring*, *41*(1), 75–96. http://dx.doi.org/10.1007/s10644-008-9042-8.

Morduch, J., & Sicular, T. (2002). Rethinking inequality decomposition, with evidence from rural China. *The Economic Journal*, *112*(476), 93–106. http://dx.doi.org/10.1111/1468-0297.0j674.

Naschold, F. (2009). Microeconomic determinants of income inequality in rural Pakistan. *Journal of Development Studies*, *45*(5), 746–768. http://dx.doi.org/10.1080/00220380902753193.

Oaxaca, R. (1973). Male-female wage differentials in urban labour markets. *International economic review*, *14*(3), 693–709. http://dx.doi.org/10.2307/2525981.

Pyatt, G., Chen, C. N., & Fei, J. (1980). The distribution of income by factor components. *The quarterly journal of economics*, *95*(3), 451–473. http://dx.doi.org/10.2307/1885088.

Sastre, M., & Trannoy, A. (2002). Shapley inequality decomposition by factor components: Some methodological issues. *Journal of Economics*, *77*(1), 51–89. http://dx.doi.org/10.1007/BF03052500.

Schultz, T. W. (1963). *The economic value of education* (Vol. 63). New York, NY: Columbia University Press.

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. *Econometrica*, *48*(3), 613–625. http://dx.doi.org/10.2307/1913126.

Shorrocks, A. F. (1982). Inequality decomposition by factor components. *Econometrica*, *50*(1), 193–211. http://dx.doi.org/10.2307/1912537.

Shorrocks, A. F. (1984). Inequality decomposition by population subgroups. *Econometrica*, *52*(6), 1369–1385. http://dx.doi.org/10.2307/1913511.

Shorrocks, A. F. (1999). *Decomposition procedures for distributional analysis: A unified framework based on the Shapley value* (Mimeo). Colchester: Department of Economics, University of Essex.

South, A. (2012). *Prospects for peace in Myanmar: Opportunities and threats* (Peace Research Institute Oslo Paper). Oslo: PRIO.

Ssewanyana, N. S., Okidi, A. J., Angemi, D., & Barungi, V. (2004). *Understanding the determinants of income inequality in Uganda* (No. WPS/2004–29). Oxford: Centre for the Study of African Economies, Department of Economics, University of Oxford.

Wan, G. (2002). *Regression-based inequality decomposition: Pitfalls and a solution procedure* (UNU-WIDER Discussion Paper No. 2002/101). Helsinki: World Institute for

Development Economics (UNU-WIDER). Retrieved from <https://www.wider.unu.edu/sites/default/files/dp2002-101.pdf>.

Wan, G. (2007). *Regional income inequality in rural China, 1985–2002: Trends, causes and policy implications* (UNU-WIDER Research Paper No. 2007/05). Helsinki: World Institute for Development Economics (UNU-WIDER). Retrieved from <https://www.wider.unu.edu/sites/default/files/rp2007-05.pdf>.

Wan, G., & Zhou, Z. (2005). Income inequality in rural China: Regression-based decomposition using household data. *Review of development economics, 9*(1), 107–120. http://dx.doi.org/10.1111/j.1467-9361.2005.00266.x.

Warr, P., & Aung, L. L. (2019). Poverty and inequality impact of a natural disaster: Myanmar's 2008 Cyclone Nargis. *World Development*, *122*, 446–461. https://doi.org/10.1016/j.worlddev.2019.05.016.

Yun, M. S. (2006). Earnings inequality in USA, 1969–99: Comparing inequality using earnings equations. *Review of Income and Wealth*, *52*(1), 127–144. http://dx.doi.org/10.1111/j.1475- 4991.2006.00179.x.

Table 1. Data summary: Mean expenditures, inequality and poverty, 2005 and 2010, panel sample and full sample

Variable	Panel sample 2005	Panel sample 2010	Panel sample Absolute Change	Panel sample Percent change	Full sample 2005	Full sample 2010	Full sample Absolute change	Full sample Percent change
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean nominal expenditures	228,254	542,505	314,251	137.7	230,308	542,971	312,663	135.8
Mean real expenditures	530,058 (13,081)	542,505 (8,987)	12,447 [13,375]	2.3	534,826 (13,110)	542,971 (8,854)	8144 [15,266]	1.5
Gini coefficient of inequality	0.248 (0.0106)	0.221 (0.0076)	-0.027** [0.0109]	-11.0	0.256 (0.0100)	0.220 (0.0074)	-0.0360*** [0.0120]	-14.0
GE(0) measure of inequality	0.1024 (0.0090)	0.0814 (0.0060)	-0.0210** [0.0091]	-20.5	0.1072 (0.0087)	0.0814 (0.0057)	-0.0258** [0.0101]	-24.1
GE(1) measure of inequality	0.1200 (0.0142)	0.0975 (0.0100)	-0.0225 [0.0146]	-18.8	0.1277 (0.0126)	0.0968 (0.0090)	-0.0308** [0.0149]	-24.2
GE(2) measure of inequality	0.1916 (0.0402)	0.1792 (0.0416)	-0.0124 [0.0480]	-6.5	0.2163 (0.0360)	0.1736 (0.0352)	-0.0427 [0.0484]	-19.7
Poverty incidence (headcount %)	31.74 (0.0165)	24.93 (0.0148)	-6.81*** [0.0184]	-21.5	32.14 (0.0164)	24.99 (0.0135)	-7.15*** [0.0205]	-22.2
Observations	9,102	9,102	n.a.	n.a.	18,634	18,609	n.a.	n.a.
Memo: Consumer price index	100	232.2	132.2	132. 2	100	232.2	132.2	132. 2

Notes:

1. '2005' refers to the IHLCA survey data collected between November-December 2004 and May 2005. '2010' refers to IHLCA survey data collected between December 2009 and May 2010.
2. Real expenditures mean household expenditures per adult equivalent per year, calculated at December 2009 prices using the consumer price index as deflator. The calculation of 'adult equivalent' uses the weights recommended in Deaton and Zaidi (2002). Poverty incidence means headcount measure, using the Myanmar government's official poverty line.
3. Deflation for different months of data collected within each of the two survey periods used a household-specific Paasche index of consumer prices reflecting price variation across states, assembled by the IHLCA survey team for the months of the survey but not for the inter-survey years. This deflator was used by the IHLCA survey team to produce data for the two survey periods in December 2004 and December 2009 prices, respectively. These data were then converted to December 2009 prices using the nation-wide consumer price index published by the Central Statistical Office, Yangon (December 2004 = 428.55; December 2009 = 995.19, an inflation rate of 132.22 per cent).
4. Standard errors (round brackets) for poverty incidence and Gini coefficients are based on the STATA code of Jenkins (2008), based on Kovacevic and Binder (1997). Significance tests were applied to absolute changes in mean real expenditures, changes in the four measures of inequality and changes in poverty incidence. *** and ** indicate statistical significance at the 99% and 95% confidence levels, respectively.
5. Consumption expenditures comprise: (i) food; (ii) non-food, including clothing and other apparel, home appliances, house rent and repair, education, travel and other household worker services; (iii) housing expenditures are the yearly user costs, approximated by rental value, measured by actual monthly rental value or estimated monthly rental value, as in IDEA and IHLCA (2007c, pp. 11-16); IHLCA (2011c, pp. 45-48), which describe the detailed steps of construction of the consumption aggregate, but also incorporating the authors' estimates of health expenditures and expenditures on durable goods, based on the IHLCA data.
6. All calculations are weighted by (survey weights X household size). Survey weights are calculated as the inverse of the sampling fraction.
7. Z-statistics [square brackets] are calculated using the method of Barrett and Pendakur (1995) and Davidson and Duclos (2000).
8. n.a. means not applicable.

Source: Authors' calculations, using data from IDEA and IHLCA (2007a, b and c) and IHLCA (2011a, b and c). Consumer price index from Central Statistical Organization (2005 and 2010).

Table 2. Fields decomposition: Level and difference of expenditure inequality, 2005 and 2010, panel sample

(Units: per cent)

Variables	Level		Difference 2005 to 2010			
	S_j^{2005}	S_j^{2010}	Gini	GE(0)	GE(1)	GE(2)
	(1)	(2)	(3)	(4)	(5)	(6)
1. Characteristics of household head	**1.85**	**1.34**	**6.01**	**3.86**	**4.11**	**10.35**
Age (years)	-0.32	-0.18	-1.40	-0.84	-0.90	-2.53
Age squared (years)	0.40	0.16	2.37	1.35	1.47	4.41
Gender (Dummy)	0.01	-0.04	0.41	0.20	0.23	0.84
Ethnicity (Myanmar =1, other = 0)	0.37	0.38	0.33	0.35	0.35	0.29
Years of non-agricultural business in operation	1.39	1.03	4.30	2.79	2.97	7.34
2. Household size	**5.61**	**7.93**	**-13.03**	**-3.38**	**-4.51**	**-32.48**
Number of persons	11.26	16.30	-29.30	-8.29	-10.77	-71.64
Number of persons squared	-5.65	-8.37	16.27	4.92	6.26	39.15
3. Proportion of hh. members by age group	**-0.70**	**-0.94**	**1.23**	**0.23**	**0.34**	**3.24**
Proportion of members aged under 6	-0.26	-0.30	0.14	-0.07	-0.04	0.55
Proportion of members aged 6-10	-0.66	-0.21	-4.29	-2.41	-2.63	-8.07
Proportion of members aged 11-15	-0.38	-0.30	-0.98	-0.67	-0.71	-1.61
Proportion of members aged 16- 65	0.59	-0.12	6.36	3.37	3.72	12.37
4. Proportion of hh. members by level of education (last 6 months)	**9.78**	**8.73**	**18.22**	**13.85**	**14.36**	**27.02**
Proportion of members with tertiary education	6.24	8.13	-8.99	-1.10	-2.03	-24.89
Proportion of members with upper secondary	1.69	2.77	-7.08	-2.54	-3.08	-16.23
Proportion of members with lower secondary	0.17	-0.05	1.98	1.04	1.15	3.86
Proportion of members with primary education	0.72	-2.20	24.25	12.07	13.50	48.81
Proportion of members illiterate	0.96	0.08	8.06	4.38	4.82	15.48
5. Proportion of hh. members sick/ ill/ injured (last 30 days)	**0.46**	**0.86**	**-2.79**	**-1.11**	**-1.30**	**-6.17**
6. Proportion of hh. members by occupation (last 7 days)	**5.91**	**4.66**	**15.98**	**10.77**	**11.38**	**26.49**
Legislators, senior officials and managers	4.75	2.22	25.09	14.55	15.79	46.31
Professionals	0.92	0.41	5.00	2.89	3.13	9.26
Technicians and associate professionals	0.97	1.33	-1.93	-0.43	-0.60	-4.96
Service workers and shop and market sales workers	4.12	2.34	18.48	11.04	11.92	33.45
Skilled agricultural and fishery workers	-2.59	-0.55	-19.04	-10.52	-11.52	-36.21
Craft and related trades workers	-0.41	-0.24	-1.79	-1.07	-1.16	-3.23
Plant and machine operators and assemblers	0.95	0.27	6.46	3.61	3.95	12.21
Elementary occupations	-2.80	-1.13	-16.28	-9.30	-10.12	-30.34
7. Land ownership and access / crop type	**2.63**	**2.83**	**0.97**	**1.83**	**1.73**	**-0.76**
Owned and accessed irrigated land area per capita	1.04	1.14	0.25	0.66	0.61	-0.59
Owned and accessed unirrigated land area per capita	1.12	0.93	2.64	1.85	1.94	4.23
Landless (Dummy)	0.47	0.76	-1.91	-0.68	-0.82	-4.39
Cultivation of cereal crops (Dummy)	-0.07	-0.10	0.14	0.03	0.05	0.36
Cultivation of pulses (Dummy)	-0.04	0.02	-0.49	-0.25	-0.28	-0.97
Cultivation of oilseed crops (Dummy)	0.13	0.01	1.11	0.60	0.66	2.13
Cultivation of tuber/root crops, spices/medicinal plants and vegetables (Dummy)	-0.05	0.00	-0.44	-0.24	-0.26	-0.85
Cultivation of fruit crops (Dummy)	0.03	0.02	0.10	0.06	0.07	0.19
Cultivation of industrial crops (Dummy)	0.00	0.05	-0.43	-0.21	-0.23	-0.87

Table 2. (cont'd) Fields decomposition: Level and difference of expenditure inequality, 2005 and 2010, panel sample

Variable	Level		Difference			
	S_j^{2005}	S_j^{2010}	Gini	GE(0)	GE(1)	GE(2)
	(1)	(2)	(3)	(4)	(5)	(6)
8. Proportion of hh. members openly unemployed (last 6 months)	**0.08**	**0.37**	**-2.27**	**-1.05**	**-1.20**	**-4.71**
9. Proportion of hh. members by sector of work (last 6 months)	**1.21**	**0.10**	**10.16**	**5.52**	**6.07**	**19.50**
Agriculture, forestry, fishing and mining	-2.26	-0.72	-14.62	-8.22	-8.97	-27.52
Manufacturing and construction	0.12	0.09	0.32	0.21	0.23	0.53
Services	3.35	0.73	24.46	13.53	14.81	46.49
10. Location and regional effects	**7.28**	**3.17**	**40.41**	**23.25**	**25.27**	**74.99**
Dry Zone region (Dummy)	0.01	0.33	-2.63	-1.27	-1.43	-5.38
Coastal region (Dummy)	-0.04	0.23	-2.18	-1.07	-1.20	-4.41
Delta region (Dummy)	4.97	0.30	42.55	23.09	25.38	81.77
Village Tract/Wards: Inland plains (Dummy)	0.25	0.42	-1.09	-0.39	-0.47	-2.48
Village Tract/Wards: Hills (Dummy)	-0.02	0.01	-0.26	-0.13	-0.15	-0.50
Village Tract/Wards: Mountains (Dummy)	1.08	0.04	9.44	5.11	5.62	18.16
Village Tract/Wards: Delta (Dummy)	-0.20	0.38	-4.88	-2.46	-2.74	-9.76
Village Tract/Wards: Valley (Dummy)	-0.07	0.02	-0.83	-0.44	-0.48	-1.63
Distance to nearest market (Miles)	-0.10	0.58	-5.61	-2.76	-3.09	-11.35
Distance to nearest financial services (Miles)	0.77	0.13	5.94	3.26	3.58	11.32
Distance to nearest health services (Miles)	0.17	0.30	-0.88	-0.34	-0.40	-1.97
Distance to primary and monastic school (Miles)	-0.02	-0.01	-0.09	-0.05	-0.06	-0.17
Distance to lower secondary school (Miles)	0.34	0.53	-1.12	-0.36	-0.45	-2.66
Distance to upper secondary school (Miles)	0.14	-0.10	2.05	1.06	1.18	4.05
11. Infrastructure and transport	**2.38**	**5.39**	**-21.85**	**-9.30**	**-10.78**	**-47.12**
Road Density by states and regions	-2.10	-0.65	-13.83	-7.76	-8.47	-26.08
Bituminous (Dummy)	0.09	2.54	-19.63	-9.42	-10.62	-40.21
Gravel roads (Dummy)	0.00	-0.11	0.92	0.45	0.50	1.88
Laterite roads (Dummy)	-0.19	-0.05	-1.33	-0.74	-0.81	-2.51
Dirt roads (Dummy)	0.60	1.40	-5.89	-2.53	-2.92	-12.66
Months on road (vehicle) and on water (boat)	0.76	-0.19	8.35	4.42	4.88	16.27
Water supply (Dummy)	0.90	0.56	3.69	2.25	2.42	6.60
Electricity supply (Dummy)	0.38	0.96	-4.25	-1.85	-2.13	-9.07
Normal transport taxi/bus (Dummy)	1.04	1.00	1.36	1.19	1.21	1.69
Normal transport taxi/bus ship/boat (Dummy)	-0.07	-0.20	0.92	0.41	0.47	1.96
Normal transport taxi/bus bullock cart (Dummy)	0.85	0.16	6.41	3.53	3.87	12.21
Normal transport horse (Dummy)	0.13	-0.04	1.43	0.75	0.83	2.79
Explained	**36.48**	**34.42**	**53.05**	**44.47**	**45.48**	**70.34**
Residual	**63.52**	**65.58**	**46.95**	**55.53**	**54.52**	**29.66**
Total	**100**	**100**	**100**	**100**	**100**	**100**

Note: All columns are presented as percentages.
Source: Authors' calculations.

Table 3. Relative factor inequality weights and log-expenditure distribution properties: *A* = 2005, *B* = 2010

Distributions	Observed distributions		Auxiliary distributions Sequence	
	A	*B*	*(A, B)*	*(B, A)*
	(1)	(2)	(3)	(4)
Relative factor inequality weights				
1. Characteristics of household head	0.01852	0.01335	0.01750	0.01944
2. Household size	0.05612	0.07927	0.06037	0.06707
3. Proportion of hh. members by age group	-0.00704	-0.00944	-0.00756	-0.00840
4. Proportion of hh. members by level of education	0.09781	0.08734	0.09588	0.10651
5. Proportion of hh. members sick/ ill/ injured	0.00457	0.00859	0.00599	0.00665
6. Proportion of hh. members by occupation	0.05913	0.04662	0.05447	0.06051
7. Land ownership and access / crop type	0.02626	0.02832	0.03479	0.03864
8. Proportion of hh. members openly unemployed	0.00076	0.00367	0.00125	0.00139
9. Proportion of hh. members by sector of work	0.01208	0.00097	0.01033	0.01147
10. Location and regional effects	0.07280	0.03165	0.05785	0.06427
11. Infrastructure and transport	0.02378	0.05386	0.02002	0.02224
Total	**0.36477**	**0.34420**	**0.35088**	**0.38980**
Properties of overall distribution of log-expenditures				
Mean	**13.07836**	**13.12253**	**13.09417**	**13.15057**
Variance	**0.18094**	**0.14251**	**0.18057**	**0.14631**

Note: Observed distributions A and B mean $\ln E_i^A$ and $\ln E_i^B$, as in equations (2) and (3), respectively. Auxiliary distributions sequence *(A, B)* and (*B*, *A*) mean $\ln E_i^*(\beta^B, X^A, \varepsilon^A)$ and $\ln E_i^*(\beta^A, X^B, \varepsilon^B)$, as in equations (7) and (11), respectively. Yun (2006) describes the construction of distribution with sequence *(A, B)*, but not (*B*, *A*).

Source: Authors' calculations.

Table 4. Yun's synthesized decomposition: Auxiliary equation sequence (*A, B)* = (2005, 2010) and (*B, A)* = (2010, 2005)

Variable Group	**Auxiliary equation sequence of periods**											
	(*A*, *B*) = (2005, 2010)						**(*B*, *A*) = (2010, 2005)**					
	Coefficients effect	**%**	**Characteristics effect**	**%**	**Total**	**Total (%)**	**Coefficients effect**	**%**	**Characteristics effect**	**%**	**Total**	**Total (%)**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
1. Characteristics of household head	0.00110	2.86	0.00035	0.91	0.00145	3.77	-0.00094	2.45	-0.00051	1.32	-0.00145	3.77
2. Household size	-0.00189	-4.92	0.00075	1.95	-0.00114	-2.97	0.00148	-3.86	-0.00034	0.89	0.00114	-2.97
3. Proportion of hh. members by age group	0.00076	1.99	-0.00069	-1.80	0.00007	0.18	-0.00012	0.30	0.00004	-0.12	-0.00007	0.18
4. Proportion of hh. members by level of education	0.00400	10.40	0.00125	3.26	0.00525	13.67	-0.00314	8.16	-0.00211	5.50	-0.00525	13.67
5. Proportion of hh. members sick/ ill/ injured	-0.00028	-0.73	-0.00012	-0.31	-0.00040	-1.04	0.00025	-0.65	0.00015	-0.38	0.00040	-1.04
6. Proportion of hh. members by occupation	0.00254	6.62	0.00151	3.93	0.00405	10.55	-0.00221	5.75	-0.00184	4.80	-0.00405	10.55
7. Land ownership and access / crop type	0.00138	3.58	-0.00066	-1.72	0.00072	1.86	-0.00162	4.21	0.00090	-2.35	-0.00072	1.86
8. Proportion of hh. members openly unemployed	-0.00034	-0.89	-0.00004	-0.11	-0.00039	-1.00	0.00032	-0.83	0.00007	-0.17	0.00039	-1.00
9. Proportion of hh. members by sector of work	0.00227	5.90	-0.00022	-0.57	0.00205	5.33	-0.00154	4.01	-0.00051	1.32	-0.00205	5.33
10. Location and regional effects	0.00797	20.74	0.00069	1.80	0.00866	22.54	-0.00489	12.73	-0.00377	9.81	-0.00866	22.54
11. Infrastructure and transport	-0.00395	- 10.28	0.00058	1.50	-0.00337	-8.78	0.00442	-11.51	-0.00105	2.73	0.00337	-8.78
Aggregate	**0.01355**	**35.26**	**0.00340**	**8.84**	**0.01695**	**44.11**	**-0.00798**	**20.77**	**-0.00897**	**23.34**	**-0.01695**	**44.11**
Explained (Coefficients + Characteristics) (%)					**0.01695**	**(44.11%)**					**-0.01695**	**(44.11%)**
Residual (%)					**0.02148**	**(55.89%)**					**-0.02148**	**(55.89%)**
Total (%)					**0.03843**	**(100%)**					**-0.03843**	**(100%)**

Notes: Auxiliary equation sequence (*A*, *B*) means the sequence of periods described in Yun (2006, p.129), as in equation (7) and column 3 of Table 3 above. Sequence (*B, A*) means the reverse sequence, as in equation (11) and column 4 of Table 3.
Source: Authors' calculations.

Table 5. Decomposition of change in inequality independent of sequence: Panel sample, 2005 to 2010

Variable Group	Coefficients effect	%	Characteristics effect	%
	(1)	(2)	(3)	(4)
1. Characteristics of household head	0.00102	2.65	0.0004	1.11
2. Household size	-0.00169	-4.39	0.0005	1.42
3. Proportion of hh. members by age group	0.00044	1.14	-0.0004	-0.96
4. Proportion of hh. members by level of education	0.00357	9.28	0.0017	4.38
5. Proportion of hh. members sick/ ill/ injured	-0.00027	-0.69	-0.0001	-0.35
6. Proportion of hh. members by occupation	0.00238	6.19	0.0017	4.37
7. Land ownership and access / crop type	0.00150	3.90	-0.0008	-2.04
8. Proportion of hh. members openly unemployed	-0.00033	-0.86	-0.0001	-0.14
9. Proportion of hh. members by sector of work	0.00190	4.95	0.0001	0.38
10. Location and regional effects	0.00643	16.73	0.0022	5.80
11. Infrastructure	-0.00419	-10.89	0.0008	2.11
Aggregate	**0.01077**	**28.02**	**0.0062**	**16.09**
Explained (Coefficients + Characteristics) (%)			**0.01695**	**(44.11%)**
Residual (%)			**0.02148**	**(55.89%)**
Total (%)			**0.03843**	**(100%)**

Source: Authors' calculations.

Appendix Table A1. Sample statistics: Panel sample

Variables	2005			2010		
	Obs.	Mean	Std. dev.	Obs.	Mean	Std. dev.
Ln real household expenditure per adult equivalent	9,102	13.08	0.43	9,102	13.12	0.38
1. Characteristics of household head						
Age	9,102	51.36	13.16	9,102	54.54	12.94
Age squared	9,102	2811.47	1416.65	9,102	3141.71	1469.85
Gender (Dummy)	9,102	0.84	0.36	9,102	0.82	0.39
Ethnicity (Myanmar =1, other = 0)	9,102	0.77	0.42	9,102	0.75	0.43
Years of non-agricultural business in operation	9,102	2.18	6.21	9,102	2.07	6.25
2. Household size						
Number of persons	9,102	6.30	2.54	9,102	6.05	2.45
Number of persons squared	9,102	46.08	39.27	9,102	42.65	38.23
3. Proportion of hh. members by age group						
Proportion of members under 6	9,102	0.10	0.12	9,102	0.07	0.11
Proportion of members 6-10	9,102	0.10	0.12	9,102	0.09	0.12
Proportion of members 11-15	9,102	0.10	0.12	9,102	0.10	0.12
Proportion of members 16- 65	9,102	0.65	0.21	9,102	0.68	0.22
Proportion of members above 65	9,102	0.05	0.12	9,102	0.06	0.14
4. Proportion of hh. members by level of education (last 6 months)						
Proportion of members with tertiary education	8,783	0.09	0.24	8,743	0.11	0.26
Proportion of members with upper secondary	8,783	0.12	0.25	8,743	0.16	0.27
Proportion of members with lower secondary	8,783	0.24	0.33	8,743	0.25	0.32
Proportion of members with primary education	8,783	0.48	0.41	8,743	0.42	0.39
Proportion of members illiterate	8,783	0.06	0.20	8,743	0.04	0.16
Proportion of members unclassified education	8,783	0.01	0.09	8,743	0.02	0.12
5. Proportion of hh. members sick/ ill/ injured (last 30 days)						
Proportion of members sick/ ill/ injured	9,102	0.07	0.13	9,102	0.08	0.15
Proportion of members who were not sick/ ill/ injured	9,102	0.93	0.13	9,102	0.92	0.15
6. Proportion of hh. members by occupation (last 7 days)						
Legislators, senior officials and managers	8,337	0.05	0.19	8,670	0.03	0.16
Professionals	8,337	0.03	0.13	8,670	0.03	0.14
Technicians and associate professionals	8,337	0.03	0.15	8,670	0.04	0.15
Clerks	8,337	0.02	0.12	8,670	0.02	0.11
Service workers and market sales workers	8,337	0.09	0.24	8,670	0.13	0.28
Skilled agricultural and fishery workers	8,337	0.37	0.44	8,670	0.35	0.43
Craft and related trades workers	8,337	0.11	0.27	8,670	0.12	0.26
Plant and machine operators and assemblers	8,337	0.04	0.16	8,670	0.04	0.15
Elementary occupations	8,337	0.27	0.39	8,670	0.25	0.38
7. Land ownership and access / crop type						
Owned and accessed irrigated land per capita	9,102	0.20	0.69	9,102	0.27	1.09
Owned and accessed unirrigated. land per capita	9,102	0.40	1.08	9,102	0.39	1.22
Landless (Dummy)	9,102	0.15	0.36	9,102	0.12	0.32
Cultivation of cereal crops (Dummy)	9,102	0.36	0.48	9,102	0.34	0.47
Cultivation of pulses (Dummy)	9,102	0.13	0.33	9,102	0.18	0.38
Cultivation of oilseed crops (Dummy)	9,102	0.17	0.38	9,102	0.16	0.37
Cultivation of tuber/root crops, spices/medicinal plants and vegetables (Dummy)	9,102	0.06	0.24	9,102	0.07	0.26
Cultivation of fruit crops (Dummy)	9,102	0.03	0.17	9,102	0.01	0.08
Cultivation of industrial crops (Dummy)	9,102	0.07	0.25	9,102	0.03	0.18

Appendix Table A1. (continued) Sample statistics: Panel sample

Variables	2005			2010		
	Obs.	Mean	Std. dev.	Obs.	Mean	Std. dev.
8. Proportion of hh. members openly unemployed (last 6 months)	8,897	0.02	0.09	8,869	0.01	0.07
9. Proportion of hh. members by sector of work (last 6 months)	8,897	0.98	0.09	8,869	0.99	0.07
Agriculture, forestry, fishing and mining	8,783	0.56	0.45	8,743	0.52	0.45
Manufacturing and construction	8,783	0.09	0.24	8,743	0.11	0.25
Services	8,783	0.33	0.41	8,743	0.36	0.42
Undefined private sector	8,783	0.02	0.13	8,743	0.02	0.10
10. Location and regional effects						
Dry Zone region (Dummy)	9,102	0.45	0.50	9,102	0.44	0.50
Coastal region (Dummy)	9,102	0.14	0.35	9,102	0.14	0.35
Delta region (Dummy)	9,102	0.26	0.44	9,102	0.27	0.44
Hills region (Dummy)	9,102	0.14	0.35	9,102	0.15	0.36
Village Tract/Wards: Inland plains (Dummy)	9,102	0.63	0.48	9,102	0.57	0.49
Village Tract/Wards: Hills (Dummy)	9,102	0.05	0.23	9,102	0.04	0.19
Village Tract/Wards: Mountains (Dummy)	9,102	0.02	0.15	9,102	0.05	0.22
Village Tract/Wards: Delta (Dummy)	9,102	0.13	0.34	9,102	0.15	0.35
Village Tract/Wards: Valley (Dummy)	9,102	0.03	0.17	9,102	0.02	0.13
Distance to nearest market (Miles)	9,102	4.63	10.37	9,102	4.09	4.54
Distance to nearest financial services (Miles)	9,102	8.31	10.80	9,102	9.96	11.05
Distance to nearest health services (Miles)	9,102	1.79	3.83	9,102	1.24	2.97
Distance to primary and monastic school (Miles)	9,102	0.19	1.74	9,102	0.37	2.08
Distance to lower secondary school (Miles)	9,102	1.97	3.79	9,102	2.28	4.08
Distance to upper secondary school (Miles)	9,102	3.82	5.55	9,102	4.05	6.36
11. Infrastructure and transport						
Road density by state and region	9,102	8.61	2.57	9,102	9.96	2.84
Bituminous road in the village/ward	9,102	0.39	0.49	9,102	0.38	0.49
Gravel roads in the village/ward	9,102	0.31	0.46	9,102	0.44	0.50
Laterite roads in the village/ward	9,102	0.17	0.37	9,102	0.28	0.45
Dirt roads in the village/ward	9,102	0.82	0.38	9,102	0.92	0.28
Months on road by vehicle and on water by boat	9,102	9.63	3.90	9,102	10.15	3.61
Water supply (Dummy)	9,102	0.21	0.40	9,102	0.29	0.45
Electricity supply (Dummy)	9,102	0.41	0.49	9,102	0.56	0.50
Normal transport taxi/bus (Dummy)	9,102	0.40	0.49	9,102	0.36	0.48
Normal transport taxi/bus ship/boat (Dummy)	9,102	0.18	0.39	9,102	0.15	0.36
Normal transport bullock cart (Dummy)	9,102	0.54	0.50	9,102	0.33	0.47
Normal transport horse (Dummy)	9,102	0.15	0.36	9,102	0.10	0.30

Note: [a] Open unemployment is defined in the poverty profile IHLCA (2011a) as the proportion of household members aged 15 and above who did not work in the 6 months prior to the survey period.

Source: Authors' estimations.

Appendix Table A2. Measures of real consumption expenditure per year by decile group: panel sample (Kyat, December 2009 prices, CPI deflator)

Consumption Deciles	Present study (including health and user-costs of durables)				IHLCA study (June, 2011) (excluding health and user-costs of durables)			
	2005	**2010**	**Difference**	**% change**	**2005**	**2010**	**Difference**	**% change**
1st decile (Lowest 10%)	249,966	277,517	27,551	11.0	239,790	267,673	27,883	11.6
2nd decile	319,980	344,352	24,372	7.6	305,365	332,519	27,154	8.9
3rd decile	368,799	389,739	20,940	5.7	349,991	372,096	22,105	6.3
4th decile	411,813	430,526	18,713	4.5	388,871	410,360	21,489	5.5
5th decile	453,603	467,974	14,371	3.2	427,577	444,633	17,056	4.0
6th decile	499,368	509,070	9,702	1.9	466,204	479,843	13,639	2.9
7th decile	553,135	556,491	3,357	0.6	511,031	521,750	10,719	2.1
8th decile	624,937	617,161	-7776	-1.2	568,884	573,356	4,472	0.8
9th decile	734,737	717,406	-17331	-2.4	659,807	654,608	-5200	-0.8
10th decile (Highest 10%)	1,191,310	1,111,034	-80276	-6.7	917,430	901,310	-16120	-1.8
National	**530,058 (13,081)**	**542,505 (8,987)**	**12,447**	**2.3**	**483,406 (10,505)**	**495,775 (6,980)**	**12,369**	**2.6**

Notes:
1. See notes to Table 1. Kyat is the currency unit of Myanmar.
2. The calculations for consumption deciles of IHLCA data (without health expenditures and user costs of durables) are also weighted by (survey weights X household size). Accordingly, the results are slightly different from IHLCA (2011a).
3. Standard errors (round brackets) for poverty incidence and Gini coefficients are based on the STATA code of Jenkins (2008), based on Kovacevic and Binder (1997).

Source: Authors' calculations, using data from IDEA and IHLCA (2007a, b and c) and IHLCA (2011a, b and c).

Appendix Table A3. Regression Results: Panel sample

Dependent Variable = Ln household expenditure per adult equivalent	2005		2010	
	Coefficient	Std. Errors.	Coefficient	Std. Errors.
1. Characteristics of household head				
Age (years)	-0.0027	0.0025	-0.0013	0.0030
Age squared (years)	0.0000	0.0000	0.0000	0.0000
Gender (Dummy)	0.0325*	0.0183	0.0181	0.0112
Ethnicity (Myanmar =1, other = 0)	0.0332	0.0240	0.0243	0.0233
Years of non-agricultural business operation	0.0059***	0.0009	0.0047***	0.0009
2. Household size and its squared				
Number of persons	-0.0902***	0.0082	-0.0998***	0.0099
Number of persons squared	0.0035***	0.0005	0.0041***	0.0008
3. Proportion of hh. members by age groups				
Proportion of members under 6	0.0749	0.0710	0.1351***	0.0489
Proportion of members 6-10	0.3301***	0.0716	0.0814	0.0598
Proportion of members 11-15	0.1775***	0.0661	0.2059***	0.0588
Proportion of members 16- 65	0.1033*	0.0572	-0.0374	0.0454
4. Proportion of hh. members by level of education (last 6 months)				
Proportion of members with tertiary education	0.3505***	0.0549	0.4136***	0.0464
Proportion of members with upper secondary	0.1463***	0.0518	0.2427***	0.0362
Proportion of members with lower secondary	0.0536	0.0500	0.1544***	0.0320
Proportion of members with primary education	-0.0295	0.0471	0.1013***	0.0370
Proportion of members illiterate	-0.1272**	0.0600	-0.0127	0.0442
5. Health condition of hh. members (last 30 days)				
Proportion of members sick/ ill/ injured	0.2053***	0.0411	0.2337***	0.0378
6. Proportion of hh. members by occupation (Last 7 days)				
Legislators, senior officials and managers	0.4951***	0.0649	0.3482***	0.0690
Professionals	0.2282***	0.0689	0.0957**	0.0418
Technicians and associate professionals	0.2515***	0.0868	0.2419***	0.0832
Service workers and market sales workers	0.3792***	0.0670	0.2059***	0.0582
Skilled agricultural and fishery workers	0.2565***	0.0496	0.1348**	0.0602
Craft and related trades workers	0.1608***	0.0462	0.1045*	0.0593
Plant and machine operators and assemblers	0.2913***	0.0582	0.1581**	0.0597
Elementary occupations	0.1366***	0.0498	0.0466	0.0539
7. Land ownership and access / crop type				
Owned and accessed irrig. land area per capita	0.0677***	0.0063	0.0361***	0.0068
Owned and accessed unirrig. land per capita	0.0467***	0.0072	0.0279***	0.0074
Landless (Dummy)	-0.0295	0.0196	-0.0499***	0.0173
Cultivation of cereal crops (Dummy)	0.0094	0.0152	0.0332**	0.0130
Cultivation of pulses (Dummy)	0.0236	0.0222	0.0121	0.0155
Cultivation of oilseed crops (Dummy)	-0.0192	0.0180	0.0164	0.0159
Cultivation of tuber/root crops, spices /medicinal plants and vegetables (Dummy)	0.0358	0.0248	0.0025	0.0265
Cultivation of fruit crops (Dummy)	0.0543	0.0380	0.0624	0.0530
Cultivation of industrial crops (Dummy)	0.0079	0.0272	0.0494*	0.0257

Appendix Table A3. (continued) Regression Results: Panel sample

Dependent Variable = Ln per adult equivalent household expenditure	2005		2010	
	Coefficient	Std. Errors.	Coefficient	Std. Errors.
8. Proportion of hh. members openly unemployed (last 6 months)	-0.1773*	0.0975	-0.2986***	0.0941
9. Proportion of hh. members by sector of work (last 6 months)				
Agriculture, forestry, fishing and mining	0.0886**	0.0438	0.0383	0.0401
Manufacturing and construction	0.1105**	0.0430	-0.0416	0.0402
Services	0.1317***	0.0483	0.0343	0.0402
10. Location and regional effects				
Dry Zone region (Dummy)	-0.0005	0.0398	0.0589	0.0397
Coastal region (Dummy)	0.0185	0.0461	-0.0301	0.0371
Delta region (Dummy)	0.2448***	0.0788	0.0466	0.0574
Village Tract/Wards: Inland plains (Dummy)	0.0295	0.0275	0.0278	0.0259
Village Tract/Wards: Hills (Dummy)	0.0446	0.0475	-0.0091	0.0275
Village Tract/Wards: Mountains (Dummy)	-0.2224***	0.0390	-0.0129	0.0517
Village Tract/Wards: Delta (Dummy)	-0.1842***	0.0650	-0.0622	0.0456
Village Tract/Wards: Valley (Dummy)	0.0578	0.0409	-0.0115	0.0363
Distance to nearest market (Miles)	0.0004	0.0004	-0.0034	0.0022
Distance to nearest financial services (Miles)	-0.0016**	0.0008	-0.0004	0.0008
Distance to nearest health services (Miles)	-0.0013	0.0019	-0.0047	0.0028
Distance to primary and monastic school (Miles)	0.0012	0.0023	0.0014	0.0024
Distance to lower secondary school (Miles)	-0.0024	0.0020	-0.0035**	0.0017
Distance to upper secondary school (Miles)	-0.0006	0.0014	0.0005	0.0012
11. Infrastructure and transport				
Road Density by state and region	-0.0210***	0.0073	-0.0140***	0.0049
Bituminous (Dummy)	0.0034	0.0170	0.0775***	0.0175
Gravel roads (Dummy)	0.0002	0.0205	-0.0116	0.0161
Laterite roads (Dummy)	-0.0420	0.0267	-0.0127	0.0173
Dirt roads (Dummy)	-0.0413	0.0310	-0.1135**	0.0451
Months on road (vehicle) and on water (boat)	0.0045*	0.0025	-0.0025	0.0019
Water supply (Dummy)	0.0549*	0.0287	0.0312*	0.0177
Electricity supply (Dummy)	0.0149	0.0138	0.0336*	0.0170
Normal transport taxi/bus (Dummy)	0.0473***	0.0176	0.0448***	0.0167
Normal transport taxi/bus ship/boat (Dummy)	0.0235	0.0221	0.0257	0.0201
Normal transport taxi/bus bullock cart (Dummy)	-0.0480***	0.0165	-0.0217	0.0153
Normal transport horse (Dummy)	0.0203	0.0173	-0.0261	0.0229
Constant	**13.0503*** **	**0.1548**	**13.3754*** **	**0.1483**
Number of Observations	**8,337**		**8,670**	
***F* -statistic**	**0.0001**		**0.0002**	
R^2	**0.3648**		**0.3442**	

Notes:

1) All estimates are computed using probability weights calculated as the inverse of the sampling fraction. Calculations are weighted by (survey weights X household size).
2) Linearized standard errors are reported and ***, ** and * indicate significance at the 99%, 95% and 90% confidence levels, respectively.
3) The education variables denote the proportion of household members of working age (15 to 64) who have completed the five levels of education. Primary education denotes 5 years of schooling or less, lower secondary education denotes between 6 and 8 years of schooling, upper secondary education denotes 9 and 10 years of education. Tertiary education denotes undergraduate diploma, bachelor degree,

and post graduate diploma. The reference category of education is the share of household members of working age with unclassified/unknown education.

4) The health variables denote the proportion of household members sick/ ill/ injured in the last 30 days. The reference category of health condition is the share of household members who were not sick/ ill/ injured in the last 30 days.
5) Variables for occupation status denote the proportion of household members of working age (15 to 64) in 9 categories: legislators, senior officials and managers, professionals, technicians and associate professionals, service workers and shop and market sales workers, skilled agricultural and fishery workers, craft and related trades workers and plant and machine operators and assemblers, and elementary occupations. The reference category in occupation variables is the share of household members who were clerks.
6) The variable for open employment denotes the proportion of household members aged 15 to 64, who looked for, but could not find one in the 6 months prior to the administrations of IHLCA surveys in 2005 and 2010.
7) Variables for industry denote the proportion of household members of working age (15 to 64) in 4 sectors combining of 11 categories. The reference category is: the share of household working members engaging in the activities of private work as employers and undifferentiated production activities.
8) Of the region dummies, the reference category is residents in the Hill region. At community level, dummies for different types of topography are included.
9) Of the infrastructure dummies, different types of roads and types of the most common mode of transportation in the community level are included.

Source: Authors' estimations.