DISCUSSION PAPER SERIES

IZA DP No. 2262

A Note on Decomposing Differences in Poverty Incidence Using Regression Estimates: Algorithm and Example

Sumon Kumar Bhaumik
Ira N. Gang
Myeong-Su Yun

August 2006

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

A Note on Decomposing Differences in Poverty Incidence Using Regression Estimates: Algorithm and Example

Sumon Kumar Bhaumik
Brunel University
and IZA Bonn

Ira N. Gang
Rutgers University
and IZA Bonn

Myeong-Su Yun
Tulane University
and IZA Bonn

Discussion Paper No. 2262
August 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

IZA Discussion Paper No. 2262
August 2006

ABSTRACT

A Note on Decomposing Differences in Poverty Incidence Using Regression Estimates: Algorithm and Example*

This paper decomposes differences in poverty incidence (head count ratio) using estimates from a regression equation, synthesizing the approaches proposed in World Bank (2003) and Yun (2004). A significance test is developed for characteristics and coefficients effects when decomposing differences in poverty incidence. The proposed method is implemented for studying differences in poverty incidence between Serbians and Albanians in Kosovo using Living Standard Measurement Survey.

JEL Classification: C20, I30

Keywords: poverty incidence, head count ratio, OLS, probit, decomposition

Corresponding author:

Myeong-Su Yun
Department of Economics
Tulane University
206 Tilton Hall
New Orleans, LA 70118
USA
E-mail: msyun@tulane.edu

* The authors wish to thank John Landon-Lane, Kunal Sen and Hiroki Tsurumi for helpful discussions.

1. Introduction

Poverty incidence can be computed utilizing estimated coefficients from the regression of the ratio of per capita income or expenditure to the poverty line on its correlates. The computation relies on calculating the probability of being in poverty, i.e., computing the cumulative distribution function using transformed regression coefficients, the (negative of) estimated coefficients divided by the standard deviation of the error term (β/σ). Once poverty incidence is measured using the transformed regression coefficients, one can answer "what if" questions by simulating the impact of various policies on poverty incidence. This approach has been heavily promoted by World Bank (2003) when the underlying variable of poverty incidence, i.e., income, is observed.[1]

We may want to compare poverty incidence across groups or periods. In doing so we may be interested in the sources of differences in poverty incidence among groups or over time. In principle, the simulation framework in World Bank (2003) can provide a partial answer by using the counterfactual value of independent variables (characteristics) that are provided from the comparison group. This simple simulation gauges the impact of changing the values of a few policy variables while keeping others constant. By tediously re-doing the simulation one can see the full impact of differences in mean characteristics between the two groups. This approach focuses on the impact of changes in characteristics given regression estimates. An earlier literature on comparing two group's outcomes applied simulation to coefficients with given values of characteristics (e.g., Abowd and Killingsworth, 1984, and Fairlie, 2005).

While these simulation methods are useful, they have an important conceptual limitation. Successive substitution simulation cannot avoid a well-known path dependency problem, that is, sequential substitution provides different pictures depending on the order of substitution, as

[1] Below we will compare this approach to the binary choice model (e.g., probit or logit).

discussed in Ham, Svejnar and Terrell (1998, pp. 1137). This shortcoming can be easily overcome by adopting a more systematic Oaxaca-type decomposition algorithm which also allows us to obtain a fuller explanation of the sources of differences in poverty incidence.[2]

We show how seamlessly the World Bank's (2003) approach of computing poverty incidence using regression estimates can be synthesized with the Oaxaca-type decomposition method proposed by Yun (2004) when studying sources of differences in poverty incidence across socio-demographic groups. The synthesized decomposition equation is easy to calculate, overcomes the path-dependency problem of the simulation method, and provides the effect of not only changes in socio-economic and demographic characteristics but also changes in their coefficients. Furthermore, we develop a significance test on whether the changes in characteristics or coefficients are significantly contributing to the differences in poverty incidence between the two groups/periods. The significance test for decomposition equation is developed by deriving the asymptotic distribution of β/σ. Using the 2001 Living Standard Measurement Survey (LSMS) for Kosovo, we demonstrate how the World Bank (2003) and Yun (2004) methodologies can be used in tandem to examine the difference in poverty incidence among Serbian and Albanian households.

2. Decomposing Differences in Poverty Incidence using Regression Coefficients: Algorithm

According to World Bank (2003), poverty incidence can be computed by first constructing R, the ratio of per capita income or expenditure (Y) to the poverty line (Z), i.e., $R = Y/Z$. Second, the regression equation, $\log R = X\beta + e$, is estimated, where R, X, and β are, respectively, an $N{\times}1$

---

[2] In the regression based models, especially when studying wage differentials, these questions are systematically handled by Oaxaca-type decomposition analyses. The Oaxaca decomposition methodology is used extensively in the labor literature to explain the wage differentials across gender and race, arising from differences in characteristics and differences in coefficients (Oaxaca, 1973).

vector, an $N \times K$ matrix of independent variables, and a $K \times 1$ vector of estimated coefficients, and e is the error term. The probability of being poor is obtained by computing $Pr(\log R < 0)$. In practice, the probability is calculated using the standard normal distribution function, $\Phi(\cdot)$, that is, $Pr(e < -X\beta) = \Phi(X\tilde{\beta})$, where $\tilde{\beta} = -\beta/\sigma$ and σ is the standard deviation of error term (e). $Pr(\log R < 0)$ is calculated separately for each group of interest.

This computation of the probability of poverty incidence provides a basis for examining differences in poverty incidence across groups or time periods applying the decomposition method developed by Yun (2004). Yun (2004) extends the Oaxaca decomposition algorithm to nonlinear models, including probit, and both the probit model and the World Bank's computation of the probability of poverty incidence use the standard normal distribution function.

Using the transformed regression coefficients, $\tilde{\beta}$, a Oaxaca-type decomposition can be done by first noting the head count ratio (H) is asymptotically equivalent to the sample average of poverty incidence (P), i.e., $H_j = \overline{P}_j = \overline{\Phi(X_j \tilde{\beta}_j)}$, where $j = A$ and B, and the "over bar" represents the value of the sample's average. Therefore, the difference in head count ratios between groups A and B can be decomposed as

$$H_A - H_B = \overline{P}_A - \overline{P}_B = \left[\overline{\Phi(X_A \tilde{\beta}_A)} - \overline{\Phi(X_B \tilde{\beta}_A)}\right] + \left[\overline{\Phi(X_B \tilde{\beta}_A)} - \overline{\Phi(X_B \tilde{\beta}_B)}\right]. \quad (1)$$

The above decomposition is done at the aggregate or overall level, which is widely accepted as a way to decompose the differences in the first moment into a linear combination of differences in characteristics $\left[\overline{\Phi(X_A \tilde{\beta}_A)} - \overline{\Phi(X_B \tilde{\beta}_A)}\right]$ and differences in coefficients $\left[\overline{\Phi(X_B \tilde{\beta}_A)} - \overline{\Phi(X_B \tilde{\beta}_B)}\right]$.

The next step is determining the contribution of each variable to the overall difference in the poverty incidence between the two groups (detailed decomposition). The key issue is the identification of the proper weight associated with the contribution of each variable to the characteristics and coefficients effects.[3] Yun (2004) proposes the detailed decomposition equation,

$$\overline{P}_A - \overline{P}_B = \sum_{k=1}^{K} W_{\Delta X}^{k} \left[\overline{(\Phi(X_A\tilde{\beta}_A)} - \overline{\Phi(X_B\tilde{\beta}_A)} \right] + \sum_{k=1}^{K} W_{\Delta\tilde{\beta}}^{k} \left[\overline{\Phi(X_B\tilde{\beta}_A)} - \overline{\Phi(X_B\tilde{\beta}_B)} \right], \quad (2)$$

where

$$W_{\Delta X}^{k} = \frac{(\overline{X}_A^k - \overline{X}_B^k)\ \tilde{\beta}_A^k}{(\overline{X}_A - \overline{X}_B)\ \tilde{\beta}_A},\ W_{\Delta\tilde{\beta}}^{k} = \frac{\overline{X}_B^k\ (\tilde{\beta}_A^k - \tilde{\beta}_B^k)}{\overline{X}_B\ (\tilde{\beta}_A - \tilde{\beta}_B)}, \text{ and } \sum_{k=1}^{K} W_{\Delta X}^{k} = \sum_{k=1}^{K} W_{\Delta\tilde{\beta}}^{k} = 1.$$

Equation (2) is a generalization of Even and Macpherson's (1990, 1993) methodology for generating only the characteristics effect in contexts where the underlying regression models use probit instead of OLS. One major merit of equation (2) is that it is free from the path dependency that arises when sequentially replacing the value associated with one of the groups with the corresponding values of the other (or comparison) group in order to compute the contribution of an individual variable or its coefficient towards the overall difference in the poverty incidence.

In addition, we can test for the statistical significance of the characteristics and coefficients effects in equation (2) using the so-called *delta* method (Yun, 2005). To compute the asymptotic variances of the characteristics and coefficients effects we need to compute the covariance matrix of $\tilde{\beta}$. World Bank (2003) uses OLS to obtain estimates of (β, σ). However, a problem associated with the use of OLS in this context is that OLS produces only a covariance matrix for β, while one

---

[3] In order to obtain a proper weight, the following are used; first, an approximation of the value of the average of the function, $\overline{\Phi(X\tilde{\beta})}$, with that of the function evaluated at the average value of exogenous variables $\Phi(\overline{X}\tilde{\beta})$; second, a first order Taylor expansion to linearize the characteristics and coefficients effects around $\overline{X}_A\tilde{\beta}_A$ and $\overline{X}_B\tilde{\beta}_B$, respectively. See Yun (2004) for details.

requires a covariance matrix for (β, σ) to calculate the covariance matrix of $\tilde{\beta} = -\beta/\sigma$. Several methods including bootstrapping can be used for this task. Yet, obtaining a covariance matrix for (β, σ) of the underlying regression model via maximum likelihood (ML) estimation seems to be simplest way to address this issue. The covariance matrix for (β, σ) obtained via ML can be used to calculate the covariance matrix of $\tilde{\beta}$.

Suppose that the covariance matrix for (β, σ) is generated by the ML estimation of the regression model is given by

$$\Sigma(\beta, \sigma) = \begin{bmatrix} \Sigma(\beta) & cov(\beta, \sigma) \\ cov(\sigma, \beta') & V(\sigma) \end{bmatrix},$$

where $\Sigma(\beta)$, $V(\sigma)$, and $cov(\beta,\sigma)$ are $K \times K$ covariance matrix of β, variance of σ, and $K \times 1$ covariance vector of β and σ. The asymptotic covariance matrix of $\tilde{\beta}$ that is generated from the covariance matrix for (β, σ) using the *delta* method is given by

$$\Sigma(\tilde{\beta}) = \left(\frac{1}{\sigma^2}\Sigma(\beta)\right) + \left(\frac{\beta}{\sigma^2}V(\sigma)\frac{\beta'}{\sigma^2}\right) - \left(\frac{1}{\sigma}cov(\beta,\sigma)\frac{\beta'}{\sigma^2}\right) - \left(\frac{\beta}{\sigma^2}cov(\sigma,\beta')\frac{1}{\sigma}\right).$$

Once the asymptotic covariance matrix of $\tilde{\beta}$ is obtained, the significance test can be easily implemented as explained in Yun (2005). Let $C = \overline{\Phi(X_A\tilde{\beta}_A)} - \overline{\Phi(X_B\tilde{\beta}_A)}$ and $D = \overline{\Phi(X_B\tilde{\beta}_A)} - \overline{\Phi(X_B\tilde{\beta}_B)}$ be the characteristics and coefficients effects, respectively. In order to construct a test for the aggregate effects, the asymptotic variances of C and D are computed using $\Sigma(\tilde{\beta}_j)$, where $j = A$ and B. The asymptotic variances of the C and D are calculated as

$$\sigma_C^2 = G_C\,\Sigma(\tilde{\beta}_A)\,G_C' \text{ and}$$

$$\sigma_D^2 = G_D \begin{bmatrix} \Sigma(\tilde{\beta}_A) & 0 \\ 0 & \Sigma(\tilde{\beta}_B) \end{bmatrix} G_D' = \frac{\partial D}{\partial \tilde{\beta}_A'}\,\Sigma(\tilde{\beta}_A)\,\frac{\partial D'}{\partial \tilde{\beta}_A} + \frac{\partial D}{\partial \tilde{\beta}_B'}\,\Sigma(\tilde{\beta}_B)\,\frac{\partial D'}{\partial \tilde{\beta}_B},$$

where $G_C = \frac{\partial C}{\partial \tilde{\beta}_A'}$ and $G_D = \left[\frac{\partial D}{\partial \tilde{\beta}_A'} \vdots \frac{\partial D}{\partial \tilde{\beta}_B'}\right]$, $1 \times K$ and $1 \times 2K$ vectors of gradients, respectively.

The exercise then involves testing whether components C and D are statistically meaningful in explaining the differences in the average probability of households in groups A and B being in poverty. The test statistics under the null hypotheses ($C = 0$ and $D = 0$) are $t_C = \frac{C}{\sigma_C}$ and $t_D = \frac{D}{\sigma_D}$, and these test statistics are asymptotically normally distributed.

The significance tests for the characteristics and coefficients effects at the individual variable level can be implemented in a similar fashion. Let $C_k = W_{\Delta X}^k \left[(\overline{\Phi(X_A\tilde{\beta}_A)} - \overline{\Phi(X_B\tilde{\beta}_A)}\right]$ and $D_k = W_{\Delta\tilde{\beta}}^k \left[\overline{\Phi(X_B\tilde{\beta}_A)} - \overline{\Phi(X_B\tilde{\beta}_B)}\right]$ be contributions of variable k to the differences in the likelihood of poverty incidence as part of C and D, respectively. The asymptotic variances of C_k and D_k are defined as

$$\sigma^2_{C_k} = \frac{\partial C_k}{\partial \tilde{\beta}'_A} \Sigma(\tilde{\beta}_A) \frac{\partial C'_k}{\partial \tilde{\beta}_A}$$

and $$\sigma^2_{D_k} = \frac{\partial D_k}{\partial \tilde{\beta}'_A} \Sigma(\tilde{\beta}_A) \frac{\partial D'_k}{\partial \tilde{\beta}_A} + \frac{\partial D_k}{\partial \tilde{\beta}'_B} \Sigma(\tilde{\beta}_B) \frac{\partial D'_k}{\partial \tilde{\beta}_B},$$

where $\frac{\partial C_k}{\partial \tilde{\beta}'_j}$ and $\frac{\partial D_k}{\partial \tilde{\beta}'_j}$ are $1 \times K$ vectors of gradients. The null hypotheses concerning variable k, i.e.,

$C_k = 0$ and $D_k = 0$, can be tested using t-test as, $t_{C_k} = \frac{C_k}{\sigma_{C_k}}$ and $t_{D_k} = \frac{D_k}{\sigma_{D_k}}$, where the test statistics

are, as before, asymptotically normally distributed.

3. An Example: Poverty Incidence in Kosovo

We use the 2001 Living Standards Measurement Survey (LSMS) data for Kosovo to demonstrate the synthesized method by sequentially employing the World Bank (2003) and Yun (2004) methods to generate the likelihood that each household *i* is in poverty and then decompose the difference in the average likelihood of poverty incidence between Serb and Albanian households.[4] The survey, which was carried out between September and December of 2000, collected data from 2,880 households. After accounting for missing values, the survey provides information on 2101 Kosovo

[4] Bhaumik, Gang and Yun (forthcoming) use this data to study poverty another measure of the living standard, per capita expenditure.

Albanian households and 416 Kosovo Serbian households.[5]

As discussed, we first compute R, the ratio of household consumption expenditure per adult equivalent and the World Bank determined poverty line of 3.499DM per adult per day. The log-transformed R is the dependent variable in our maximum likelihood estimation. For our purposes here we choose a parsimonious vector of household characteristics as the explanatory variables, as laid out in Table 1. The maximum likelihood estimates are reported in Table 2. Log-likelihood ratio tests of the joint hypothesis $\beta = 0$ (not including the constant term) rejects the null hypotheses in both Serb and Albanian regressions.

Using the estimates of regression equation of log R, we can study the impacts of changes in coefficients and characteristics on poverty incidence. Here, we compare the gap in poverty incidence between Serbs (group A) and Albanians (group B). The differences in the average probability of being poor between groups A and B, ($\overline{P}_A - \overline{P}_B$), can be algebraically decomposed into two components which represent the characteristics and coefficients effects. The 10.56% difference in predicted poverty incidence between Serbs (55.98%) and Albanians (45.41%) are decomposed in Table 3.[6] The overall characteristics effect is –0.035. Of the 10.56 percentage point gap, -3.54 percentage points is the characteristics effect, or -3.54/10.56 = -33.55% of the gap in poverty incidence is due to characteristics differences. The overall coefficients effect is 0.141. Of the 10.56 percentage point gap, 14.11 percentage points or 14.11/10.56 = 133.55% of the gap in poverty

[5] This is an over-sampling of Serbian households. In a sample of only Serbs and Albanians, Serbs should account for 7.36% and Albanians 92.6% of the observations. In our data, 83% of the households are Albanians, and the rest are Serbs. We use weights to account for this difference between the population and the sample. See http://www.worldbank.org/LSMS/guide/select.html for details of the survey (last accessed July 31, 2006).

[6] The observed difference in poverty rates are 11.87% between Serbs (57.38%) and Albanians (45.52%).

incidence. In other words, Serbs would be *worse off* if the differences between their characteristics and those of the Albanian households disappear, and Serbs would be *better off* if there is no difference in the poverty mitigating effectiveness of those characteristics between the Serbian and Albanian households. When we look at detailed decomposition, it becomes clear that the main reason why Serbs have higher poverty incidence is due to coefficients effect of constant term. Even though Serbs have better characteristics which can lower poverty incidence, and enjoy stronger poverty mitigating effect of these characteristics relative to Albanians, there is huge baseline gap in poverty incidence between the two ethnic groups, captured by the coefficients effect of the constant term.

4. World Bank (2003) vs. the Binary Choice Model, and Decomposition Analysis

Another approach to predicting poverty incidence is to probit estimation to examine the determinants of being in poverty. World Bank (2003) notes that binary choice models (e.g., probit) typically have better predictive power in classifying households as poor or non-poor than models that generate the likelihood of a household being in poverty using regression estimates of the continuous dependent variable (log R). World Bank (2003) views predictive power as less of a problem than the sensitivity of coefficient estimates in bivariate models to specification. However, even though theoretically the continuous variable contains more information than using just binary information, and regressions that use continuous dependent variables are not as sensitive to regression specifications, it is not clear whether the gains associated with not using a probit approach outweighs the inferior fit of their underlying regression models.

We estimated the probit model whose dependent variable takes the value 1 when a household's per adult equivalent consumption is lower than the World Bank's poverty line (Table 4), and we generate the characteristics and coefficients effects (Table 5) using the approach we described earlier (Yun, 2004, 2005). The probit decomposition shows qualitatively similar results. They are also qualitatively similar except for variables wealth/assets and geographic characteristics. Our results also show that probit produces a better predictor of poverty incidence than the continuous variable regression approach: the actual gap is 11.87%, the predicted gap using probit is 11.79%, while that using the World Bank (2003) approach is10.56%.

5. Conclusion and Discussion

We bring together the methodologies of the World Bank (2003) and Yun (2004). Using the former, we are able to generate the likelihood of each household in the sample being in poverty. We then use Yun's (2004) methodology to decompose the average (estimated) likelihood of poverty incidence of two population groups into characteristics and coefficients effects. In addition, we are able to compute the statistical significance of these characteristics and coefficients effects (Yun, 2005). We use these methodologies to generate the likelihood of poverty incidence among Serb and Albanian households sampled in the 2001 LSMS for Kosovo, and then decompose the difference in the average likelihood of being in poverty for Serbs and Albanians. The proposed decomposition equation is easy to calculate and overcomes the path-dependency problem.

We also show that the decomposition results obtained using probit based decomposition of differences in headcount ratios of two population groups are similar to those obtained by combining the World Bank and Yun's approaches.

REFERENCES

Abowd, J.M. and M.R. Killingsworth, 1984, Do minority/white unemployment differences really exist?, Journal of Business and Economic Statistics, 2, 64-72.

Bhaumik, S.K., I.N. Gang and M.-S. Yun, forthcoming, Ethnic conflict and economic disparity: Serbians and Albanians in Kosovo, Journal of Comparative Economics.

Blinder, A.S., 1973, Wage discrimination: reduced form and structural estimates, Journal of Human Resources, 8, 436-455.

Doiron, D.J. and W.C. Riddell, 1994, The impact of unionization on male-female earnings differences in Canada, Journal of Human Resources, 29, 504-534.

Even, W.E. and D.A. Macpherson, 1990, Plant size and the decline of unionism, Economics Letters, 32, 393-398.

Even, W. E. and D. A. Macpherson, 1993, The decline of private-sector unionism and the gender wage gap, Journal of Human Resources, 28, 279-296.

Fairlie, R. W., 2005, An extension of the Blinder-Oaxaca decomposition technique to logit and probit models, Journal of Economic and Social Measurement, 30, 305-316.

Ham, J.C., J. Svejnar and K. Terrell, 1998, Unemployment and the social safety net during transitions to a market economy: evidence from the Czech and Slovak Republics, American Economic Review, 88, 1117-1142.

Oaxaca, R., 1973, Male-female wage differentials in urban labor markets, International Economic Review, 14, 693-709.

Ravallion, M., 1994, Poverty comparisons, Philadelphia: Harwood Academic Publishers.

World Bank, 2003, Poverty Reduction Strategy Sourcebook, http://www.worldbank.org/poverty/strategies/sourcons.htm. Last accessed July 21, 2006.

Yun, M.-S., 2004, Decomposing differences in the first moment, Economics Letters, 82, 275-280.

Yun, M.-S., 2005, Hypothesis tests when decomposing differences in the first moment, Journal of Economic and Social Measurement, 30, 295-304.

Table 1
Characteristics of the Households

	Albanians			Serbs		
	All	**Non-poor**	**Poor**	**All**	**Non-poor**	**Poor**
Expenditure and poverty						
Per adult equivalent expenditure (DM)	128.29	173.71	73.92	111.23	163.55	72.37
	(73.35)	(70.82)	(19.78)	(66.99)	(72.04)	(20.55)
Poverty Rate		0.46 (0.50)			0.57 (0.49)	
Demographic characteristics of households						
Proportion aged 15 or below	0.32	0.29	0.36	0.18	0.18	0.18
	(0.21)	(0.21)	(0.21)	(0.20)	(0.19)	(0.21)
Proportion aged above 65	0.04	0.04	0.05	0.10	0.06	0.14
	(0.12)	(0.10)	(0.13)	(0.23)	(0.16)	(0.27)
Proportion of adults who are male	0.48	0.49	0.46	0.48	0.51	0.46
	(0.15)	(0.15)	(0.15)	(0.20)	(0.20)	(0.21)
Households with male head	0.93	0.94	0.93	0.87	0.93	0.83
	(0.25)	(0.24)	(0.26)	(0.34)	(0.25)	(0.38)
Education of adults						
Proportion with primary education	0.45	0.39	0.52	0.31	0.19	0.40
	(0.30)	(0.29)	(0.30)	(0.34)	(0.27)	(0.36)
Proportion with secondary education	0.29	0.33	0.25	0.51	0.58	0.46
	(0.26)	(0.26)	(0.25)	(0.35)	(0.35)	(0.35)
Proportion with vocational training	0.08	0.09	0.07	0.07	0.09	0.06
	(0.17)	(0.18)	(0.16)	(0.18)	(0.20)	(0.16)
Proportion with tertiary education	0.09	0.12	0.05	0.08	0.13	0.04
	(0.19)	(0.23)	(0.14)	(0.20)	(0.26)	(0.11)
Labor market characteristics						
Proportion of working adults	0.41	0.47	0.34	0.47	0.51	0.43
	(0.29)	(0.28)	(0.28)	(0.38)	(0.37)	(0.39)
Proportion of households with members working in family farms & businesses	0.27	0.29	0.24	0.34	0.33	0.34
	(0.28)	(0.29)	(0.28)	(0.40)	(0.39)	(0.40)
Wealth/Assets						
Acreage of land household owns (000)	0.07	0.07	0.07	0.12	0.16	0.08
	(0.10)	(0.10)	(0.09)	(0.62)	(0.93)	(0.13)
Value of animals household owns (000 DM)	0.56	0.57	0.55	0.46	0.39	0.51
	(0.78)	(0.81)	(0.73)	(0.75)	(0.73)	(0.76)
Transfers						
Households at least one of whose members has a disability card	0.10	0.09	0.12	0.10	0.08	0.11
	(0.30)	(0.28)	(0.32)	(0.30)	(0.28)	(0.32)
Household at least one of whose members receive private transfers	0.44	0.44	0.43	0.05	0.06	0.04
	(0.50)	(0.50)	(0.50)	(0.21)	(0.23)	(0.19)
Geographic characteristics						
Urban households	0.28	0.31	0.24	0.42	0.44	0.40
	(0.40)	(0.41)	(0.38)	(0.49)	(0.49)	(0.49)
Number of households	2101	1136	965	416	180	236

Notes: The figures within the parentheses are standard deviations.

Table 2
Determinants of Ratio of Per Capita Expenditure to Poverty Line (ML estimation)

	Albanians		**Serbs**	
	Estimate	S.E.	Estimate	S.E.
Constant	- 0.33***	(0.09)	- 1.10***	(0.21)
Demographic characteristics of households				
Proportion aged 15 or below	- 0.58***	(0.06)	- 0.17	(0.12)
Proportion aged above 65	- 0.10	(0.11)	- 0.06	(0.13)
Proportion of adults who are male	0.04	(0.09)	0.23	(0.16)
Households with male head	- 0.06	(0.05)	0.06	(0.09)
Education				
Proportion of adults with primary education	0.18**	(0.08)	0.31	(0.19)
Proportion of adults with secondary education	0.58***	(0.08)	0.92***	(0.20)
Proportion of adults with vocational training	0.52***	(0.10)	0.91***	(0.23)
Proportion of adults with tertiary education	0.75***	(0.10)	1.46***	(0.21)
Labor market characteristics				
Proportion of working adults	0.45***	(0.06)	0.22**	(0.11)
Proportion of households with members working in family farms & businesses	- 0.00	(0.07)	- 0.04	(0.11)
Wealth/Assets				
Acreage of land household owns (000)	0.17	(0.15)	0.01	(0.01)
Value of animals household owns (000 DM)	0.03	(0.02)	0.04	(0.03)
Transfers				
Households at least one of whose members has a disability card	0.02	(0.04)	- 0.10	(0.07)
Household at least one of whose members receive private transfers	0.09***	(0.02)	0.33***	(0.11)
Geographic Characteristics				
Urban households	0.05	(0.03)	0.06	(0.06)
Standard deviation of error term (σ)	0.46***	(0.01)	0.46***	(0.03)
Log-likelihood (*L*)	-150785.98		-19300.24	
Constrained Log-likelihood (*L0*)	-180607.61		-25300.89	
Number of households	2101		416	

Note: *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. Weights are used in estimation. Standard errors which are robust to mis-specification are reported. Constrained log-likelihood is calculated only when constant and standard deviation of error term are estimated. Likelihood ratio test, 2*(*L* - *L0*), rejects the null hypothesis that coefficients except for the constant are zero for both Serbs and Albanians.

Table 3
Decomposing Difference in Poverty Rates of 10.56% between Serbs and Albanians using Estimates of Per Capita Expenditure Regression Equations

	Characteristics Effect		Coefficients Effect	
	Estimate	Share	Estimate	Share
Aggregate Effect	- 0.035	- 33.55	0.141***	133.55
Aggregate Effect Without Constants	- 0.035	- 33.55	- 0.429**	- 405.66
Constant			0.570***	539.21
Demographic characteristics of households	- 0.016	- 15.04	- 0.244**	- 231.24
Proportion aged 15 or below	- 0.021*	- 19.52	- 0.095***	- 90.19
Proportion aged above 65	0.003	2.90	- 0.001	- 0.97
Proportion of adults who are male	- 0.001	- 1.33	- 0.068	- 64.11
Proportion with male head	0.003	2.92	- 0.080	- 75.96
Education	- 0.113***	- 106.66	- 0.191	- 180.66
Proportion of adults with primary education	0.034*	32.64	- 0.044	- 41.81
Proportion of adults with secondary education	- 0.165***	- 155.94	- 0.076*	- 72.17
Proportion of adults with vocational training	0.006***	5.33	- 0.024	- 22.42
Proportion of adults with tertiary education	0.012***	11.32	- 0.047***	- 44.25
Labor market characteristics	- 0.008*	- 7.64	0.074**	70.41
Proportion of working adults	- 0.010*	- 9.67	0.067*	63.30
Proportion of households with members working in family farms & businesses	0.002	2.02	0.008	7.11
Wealth/Assets	0.003	2.46	0.003	2.88
Acreage of land household owns (000)	- 0.000	- 0.36	0.008	7.39
Value of animals household owns (000 DM)	0.003	2.82	- 0.005	- 4.51
Transfers	0.106**	100.00	- 0.068*	- 64.28
Proportion of households at least one of whose members has a disability card	- 0.000	- 0.03	0.009	8.65
Proportion of household at least one of whose members receive private transfers	0.106**	100.03	- 0.077**	- 72.92
Geographic Characteristics				
Urban households	- 0.007	- 6.67	- 0.003	- 2.77

Note: Share is the ratio of the contribution of each factor to the "predicted" overall difference in poverty rate (10.56%) between Serbs (55.98%) and Albanians (45.41%), in percentage terms. The observed overall difference in poverty rate are 11.87% between Serbs (57.38%) and Albanians (45.52%). The predicted poverty rate is computed using estimates from the per capita expenditure regression. The details of the computation using the per capita expenditure regression is explained in the main text. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Table 4
Determinants of Poverty Incidence (Probit)

	Albanians		**Serbs**	
	Estimate	S.E.	Estimate	S.E.
Constant	0.71***	(0.26)	2.98***	(1.07)
Demographic characteristics of households				
Proportion aged 15 or below	1.30***	(0.16)	0.60*	(0.35)
Proportion aged above 65	0.42	(0.29)	0.17	(0.38)
Proportion of adults who are male	- 0.09	(0.23)	- 0.09	(0.42)
Households with male head	- 0.02	(0.13)	- 0.35	(0.26)
Education				
Proportion of adults with primary education	- 0.27	(0.23)	- 1.43	(1.05)
Proportion of adults with secondary education	- 1.13***	(0.24)	- 2.73***	(1.06)
Proportion of adults with vocational training	- 0.98***	(0.30)	- 2.69**	(1.09)
Proportion of adults with tertiary education	- 1.56***	(0.30)	- 3.91***	(1.09)
Labor market characteristics				
Proportion of working adults	- 1.12***	(0.19)	- 0.59*	(0.32)
Proportion of households with members working in family farms & businesses	0.15	(0.19)	0.20	(0.34)
Wealth/Assets				
Acreage of land household owns (000)	- 0.24	(0.40)	- 1.04*	(0.61)
Value of animals household owns (000 DM)	- 0.05	(0.05)	0.15	(0.13)
Transfers				
Households at least one of whose members has a disability card	0.05	(0.11)	0.18	(0.23)
Household at least one of whose members receive private transfers	- 0.17***	(0.07)	- 0.69*	(0.36)
Geographic Characteristics				
Urban households	- 0.07	(0.09)	0.15	(0.19)
Pearson's Chi-Square test	231878.76***		30563.67***	
Number of households	2101		416	

Note: *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. Weights are used in estimation. Standard errors which are robust to mis-specification are reported.

Table 5
Decomposing Difference in Poverty Rates of 11.79% between Serbs and Albanians using Probit Estimates

	Characteristics Effect		Coefficients Effect	
	Estimate	Share	Estimate	Share
Aggregate Effect	- 0.051	- 43.58	0.169***	143.58
Aggregate Effect Without Constants	- 0.051	- 43.58	- 0.562*	- 476.41
Constant			0.731**	619.99
Demographic characteristics of households	- 0.015	- 12.79	- 0.178*	- 151.07
Proportion aged 15 or below	- 0.024*	- 20.21	- 0.074*	- 62.60
Proportion aged above 65	0.003	2.36	- 0.004	- 3.07
Proportion of adults who are male	- 0.000	- 0.16	- 0.000	- 0.39
Proportion with male head	0.006	5.23	- 0.100	- 85.01
Education	- 0.094***	- 79.66	- 0.428	- 362.80
Proportion of adults with primary education	0.054	45.59	- 0.167	- 141.41
Proportion of adults with secondary education	- 0.164**	- 139.20	- 0.151	- 127.83
Proportion of adults with vocational training	0.006**	4.76	- 0.044	- 36.93
Proportion of adults with tertiary education	0.011***	9.19	- 0.067**	- 56.63
Labor market characteristics	- 0.005*	- 4.66	0.074**	62.37
Proportion of working adults	- 0.009*	- 7.82	0.069	58.78
Proportion of households with members working in family farms & businesses	0.004	3.16	0.004	3.59
Wealth/Assets	- 0.017**	- 14.83	0.019	16.07
Acreage of land household owns (000)	- 0.013*	- 11.41	- 0.018	- 14.99
Value of animals household owns (000 DM)	- 0.004	- 3.42	0.037	31.06
Transfers	0.074**	63.13	- 0.068	- 57.51
Proportion of households at least one of whose members has a disability card	- 0.000	- 0.02	0.004	3.69
Proportion of household at least one of whose members receive private transfers	0.074**	63.15	- 0.072**	- 61.21
Geographic Characteristics				
Urban households	0.006	5.22	0.020	16.53

Note: Share is the ratio of the contribution of each factor to the "predicted" overall difference in poverty rate (11.79%) between Serbs (57.30%) and Albanians (45.51%), in percentage terms. The observed overall difference in poverty rate are 11.87% between Serbs (57.38%) and Albanians (45.52%). The predicted poverty rate is computed using probit estimates. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.