GROUNDING THE ADVOCACY ARGUMENT:
MEASURING THE NATIONAL BUDGET EFFECTS OF IMPROVING ROMA EDUCATION LEVELS IN ALBANIA, SERBIA AND REPUBLIC OF MACEDONIA

May 2014

This report was prepared by Gabor Kezdi of Central European University and Gabor Kertesi of the Institute for Advanced Study at Central European University & CERS, Hungarian Academy of Sciences for JBS International under Contract No. AID-OAA-BC-10-00005.

GROUNDING THE ADVOCACY ARGUMENT:
MEASURING THE NATIONAL BUDGET EFFECTS OF IMPROVING ROMA EDUCATION LEVELS IN ALBANIA, SERBIA AND REPUBLIC OF MACEDONIA

Report on the exploratory analysis

May 2014

The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

Table of Contents

Acronyms iv
Introduction 1
1. The Hungarian study 1
2. Feasibility of similar studies in Albania, Republic of Macedonia and Serbia using existing data 3
3. Recommendations for data collection in Albania, Republic of Macedonia and Serbia 4
Appendix: Detailed results of the country studies 6
Albania 6
Republic of Macedonia 8
Serbia 10

Acronyms

CEU	Central European University
EC	European Community
EUR	Euros
UNDP	United Nations Development Programme
USAID	U.S. Agency for International Development
WB	World Bank

Introduction

A lack of data on Roma in most European countries hinders private and public efforts to improve the situation of the poorest and most marginalized minority on the continent. Hungary is a special case where detailed data are available on employment, earnings and consumption, there is a representative survey that helps assessing ethnic differences in these, and there is additional information on education outcomes, welfare receipt and other important factors. With these data, we developed a methodology to measure governmental cost savings and tax revenue generated through improving the education of the country's Roma citizens.[1] The methodology and framework, described below, can be replicated in other countries, including for other ethnic minorities in other geographic contexts, but only when data actually exist.

We carried out preliminary analysis exploring the feasibility of similar studies in Albania, Serbia and Macedonia.[2] The analyses comprise of a template for the replication of the analysis and a listing of the data sources. With the help of Central European University (CEU) students from the three countries we identified the source of all available data and carried out preliminary estimations where appropriate data were available. On May 9, 2014, we then held a workshop for representatives from government and statistical offices from each of Albania, Serbia and Republic of Macedonia to discuss the preliminary analyses and explore potential sources of data, as well as how these sources might be utilized or modified.

This report summarizes the lessons from the preliminary analysis and the workshop. The first section summarizes the methodology and data requirements of the benchmark Hungarian study. The second section summarizes the availability of data and the results of the preliminary analyses. The third section concludes and contains our recommendations. The report is kept concise; the country-specific technical reports are in the appendix.

1. The Hungarian study

In our Hungarian study (http://www.econ.core.hu/doc/bwp/bwp/bwp0605.pdf) we estimated the net benefit of a hypothetical investment, taking place at age four. More specifically, we researched the potential benefits of enabling a young Roma to successfully complete secondary school (with a closing exam called maturity examination or "eretttsegi") instead of stopping at grade eight (or dropping out of secondary school). The benefits were interpreted in a narrow budgetary sense (national budget, social security, public health care) and did not consider wider social benefits (such as positive human capital externalities or better citizenship). The analysis considered seven channels of possible individual contribution: personal income tax on income earned from registered full-time employment, social security (including public health care) contributions paid by employers and employees on earned income, unemployment benefits, means-tested welfare benefits, costs of public employment projects, value added and excise tax on consumption, and incarceration costs. The estimates were adjusted by the extra costs of secondary and college education.

The analysis used large sample surveys, aggregate administrative data, and tax and contribution rules to estimate the necessary parameters. The analysis required estimated levels of employment, earnings, unemployment, consumption and welfare receipts of Roma people of different ages and levels of education. These allowed for the estimation of net contributions to the national budget and social security by age and levels of education. The

1. That paper was commissioned by the Roma Education Fund, http://www.romaeducationfund.hu/sites/default/files/publications/kertesi-kezdi-budgetarybenefits.pdf

2. This workshop and the report were funded by USAID's Europe and Eurasia Bureau. The countries selected for the analysis have sizable Roma populations and are the countries where USAID currently works.

estimated age and education level – specific contributions to the national budget were then discounted those contributions to age four. To investigate the robustness of the results several discount rates were applied.

Ideally the analysis would have used large-sample data on Roma individuals to estimate required statistics by age and education level. While we had access to high quality and representative data on the Roma of Hungary, those data covered samples that were not large enough. The most important constraint of those data was the small number of Roma with secondary and higher education, due to the obvious fact that they are a small minority among the Roma and only very large datasets could provide adequate coverage of them.

Alternatively, in the Hungarian study we estimated employment, earnings, unemployment, consumption and welfare receipt by age and level of education for the entire population of Hungary using large nationally representative data. We then adjusted these national estimates by the average difference between Roma and non-Roma to arrive at the Roma estimates by age and level of education. The main sources of data for the national profiles were the Hungarian Labor Force Survey, the Hungarian Wage Survey and the Hungarian Household Budget Survey. Of these three, the first and the last are standard surveys that many countries administer (including Albania, Republic of Macedonia and Serbia). The Hungarian Wage Survey is unique in its coverage and information content (including the age and the level of education of the individual wage earners). The adjustments to the Roma minority were based on the Hungarian Roma Survey of 2003. Importantly, this survey was representative of all Roma in Hungary not only those that lived in segregated Roma communities.

Our results indicated that an investment that makes one young Roma successfully complete secondary school would yield significant direct long-term benefits to the national budget. According to our benchmark estimate, discounted to age four (a possible starting age for such an investment), the present value of the future benefits was about EUR 70,000 relative to the value the government would collect on the representative person in case if s/he had not continued studies after the primary school. The benefits were somewhat smaller if (without the suggested early childhood educational investment), a young Roma person finished vocational training school (EUR 55,000). The estimated returns were sensitive to the discount rate, the assumed wage growth, the college completion rate after secondary school, and the employment and wage differentials of Roma and the national averages. But even our most conservative estimates suggested that benefits were least EUR 21,000 to 27,000.

We formulated all results in terms of the benefits of an investment that makes one child successfully complete secondary school, for methodological convenience. Naturally, no investment is certain to bring such a result. When comparing benefits to costs, one has to factor in the success probabilities. For example, if an investment increases the chance of secondary school completion by 20 percentage points, i.e., one child out of five gets there as a result of the investment, benchmark benefits relative to 8 grades are EUR 14,000 (70,000/5). In other words, an investment of EUR 14,000 per child would break even with a 20 percent success rate. Even by looking at our most conservative estimates, any investment with such a success rate is almost sure to yield a positive return if costs are EUR 6,000 per child. An important caveat of the results is that those returns would occur in adulthood, which is typically many years after the investment takes place. It therefore takes a government that plans for the long run to apply the cost-benefit calculations implied by our estimates.

Our results showed that, overwhelmingly, the benefits would come from increased government revenues, from personal income tax and employer/employee contributions after earned income. Savings on unemployment insurance, welfare benefits and public employment projects would be negligible, and savings on incarceration costs would also be small. Larger contributions through value added tax on consumption would also be sizable.

2. Feasibility of similar studies in Albania, Republic of Macedonia and Serbia using existing data

Comparable data on the Roma population is available in all countries from the UNDP/WB/EC survey that was fielded in 2011. In each country the sample was a random selection of Roma households in neighborhoods populated exclusively or mostly by Roma households. Size of the sample in each country was approximately 750 Roma households.

The major advantage of the UNDP survey is its comparability across countries. Both the survey instrument and the sampling procedure were harmonized, and the field work took place in the same three months (May through July). The survey contains measures of employment, unemployment, earnings as well as consumption, making it a good candidate for our analysis.

The size of the sample does not allow for estimating age and education-specific statistics. On the other hand, it allows for either of two approaches: (1) estimating correction factors to national age-specific estimates (similarly to the Hungarian study), or (2) estimating differences by education only, ignoring the age factor. As will be explained later, age and education-specific estimates were impossible to produce from published data even at the national level in the three countries. As a result, we produced estimates using the second approach.

Our estimates are very preliminary. The UNDP data are rich but also complex and it requires in-depth analysis to estimate the statistics we need in robust ways. All surveys are subject to missing data and measurement errors, and the UNDP data set is no exception. Earnings data are notorious for having many missing answers in general, and in this survey as well, and dealing with those missing answers was outside the scope of our project.

Perhaps more importantly, estimates of the potential returns to education based on the UNDP data are biased downward, and the bias is potentially large. Young Roma who obtain higher levels of education tend to move out of segregated Roma neighborhoods to more mainstream neighborhoods. At the same time, the UNDP sampling frame contains neighborhoods only; it does not have any Roma that moved out of those segregated neighborhoods. As a result, our data have substantially fewer Roma with higher levels of education than in reality. Moreover, the Roma with higher levels of education in the UNDP data are those that remained in segregated neighborhood, and their employment and earnings prospects are likely substantially worse than the employment and earnings prospects of their peers who live in other neighborhoods. This is very unfortunate for our analysis: the differences between less educated Roma and more educated Roma in our sample are likely substantially smaller than the true differences in the entire society. For this reason estimates of the returns to education from the UNDP data are almost surely substantially smaller than the true returns.

The appendix shows our estimates for the three countries using the UNDP data. The estimates cover all accounts that proved to be important in the Hungarian study (revenues from personal income taxes, social security contributions and consumption taxes, as well as expenditures on unemployment benefits and welfare benefits). In each of the three countries the estimated budget effects of improving Roma education are an order of magnitude smaller than they were in Hungary. Although these results are very preliminary, and a more in-depth analysis of the data may reveal larger effects, they are in line with our concerns that the UNDP data are likely not suited for this analysis.

Unfortunately, we could not identify other information sources that could provide data on employment, earnings and consumption of the Roma in the three countries.[3] The national census data in Serbia may provide estimates for employment but it has no information about wages or welfare receipt.[4]

National estimates of employment, unemployment, earnings and consumption seem possible to obtain in Serbia, but the possibilities are more limited in Albania and Republic of Macedonia.[5] Unfortunately, micro-level data from censuses, labor force surveys, wage surveys and household budget surveys are not available for research in any of the countries. That is in contrast with Hungary where such data have been available to researchers for a long time. At the same time, the Statistical Office of the Republic of Serbia has developed a service that provides statistics from these surveys as specified by users. As a result, a limited version of the Hungarian study approach may be adequate in Serbia. But many uncertainties need yet to be resolve for that. First, wage data may not be available to the necessary details. Second, this approach would require adjustment factors to arrive at the Roma estimates, and apart from employment information from the census it is not clear how that could be achieved.

As participants of our workshop suggested, the national budget effects of improving Roma education levels in Albania, Serbia and Republic of Macedonia may be indeed substantially smaller than those effects in Hungary, for multiple reasons. First, the larger size of the informal economy means that a larger part of earnings and consumption may bring no revenues to the national budget. As a result, a smaller part of the benefits of increasing the levels of education would go to the national budget. Second, the tax burden is lower in these countries, resulting again in a smaller part of the benefits to go to the national budget. Third, a substantial fraction of the working population migrates to other countries. The incomes earned from such employment do not bring tax revenues to the sending countries. If more educated workers earn more abroad, they may send larger remittances that leads to higher levels of consumption, which may result in higher contributions to the national budget from consumption taxes. Unfortunately, these effects are very difficult to trace out.

The conclusion from our research is that the lack of data hinders from carrying out a credible estimation of the national budget effects of improving Roma education levels in Albania, Serbia or Republic of Macedonia. In the last part of this report we give some recommendations for data collection that can allow for estimating the object of our study and many more analyses on the situation of the Roma.

3. Recommendations for data collection in Albania, Republic of Macedonia and Serbia

The principles behind our recommendations are the following.

(1) Data on the Roma minority should be collected regularly on the most important information such as demographics, education, employment, earnings, consumption and housing.
(2) Data on the Roma minority should be directly comparable to data on the majority or other minorities.
(3) Ethnic affiliation should be measured by respondent identifying themselves.
(4) The data should yield indicators that are comparable across countries and with standard measures.

3. An exception is Albania where we identified another Roma survey (commissioned by the Open Society Foundations in Albania; http://www.soros.al/2010/foto/uploads/File/Dritan/ROMA%20situation%20in%20Albania/roma%20decade%20eng%20final.pdf). Unfortunately, the sampling frame and the size of this survey are very similar to the UNDP survey: it is relatively small and it covers segregated Roma communities.
4. There was no census in 2011 in Republic of Macedonia. The Albanian census has a common ethnic marker for Roma and Egyptian minorities and thus cannot provide estimates for the Roma separately.
5. Details of publicly available data that could be used for the analysis is provided in the country case studies in the Appendix.

The most straightforward and least expensive way to meet these principles is adding questions on ethnic identity to standard surveys. The most important of these are the Labor Force Surveys and the Household Budget Surveys, but surveys on incomes (such as the Statistics on Incomes and Living Conditions) or health (such as the European Health Interview Survey) would also be extremely informative.

In order to have information on earnings, the Labor Force Survey should ask about wage on top of employment. Labor Force Surveys interview their respondents multiple times (six times typically). Questions about earnings are usually thought as sensitive ones that would induce non-response to subsequent interviews. To mitigate such potential negative effects, the questions on earnings could be administered to respondents who participate in the survey for the last time. If questions about ethnicity are considered sensitive, those, too, could be asked on the last interview only.

Research is facilitated best if all such data are made available to researchers as micro-level datasets. If confidentiality issues arise, use of the data can be restricted and monitored.

It is important to note that while these recommendations may sound radical in the context of the three countries they are considered seriously, and most are already implemented, in other countries, including Hungary. Such data would allow for analyzing and monitoring the social situation of the Roma minority in the three countries in ways that are comparable to each other and other countries with a sizeable Roma minority in Central and Eastern Europe.

Appendix: Detailed results of the country studies

Albania

Challenges with publicly available data in Albania

The table below summarizes the institutions and the source documents that have publicly available data needed for the project. As apparent from the table, data by levels of education is not always available publicly in Albania.

Table 1: Publicly available data in Albania

Data	Source Documents	Institutions	By Educational Level	By Ethnicity
Employment	Labor Force Survey, 2012	INSTAT	Yes	No
Average wage	Only incomes available, 2012	ISNTAT	Yes (by education of family head)	No
Pensions	Labor Force Survey	INSTAT	No	No
Welfare assistance	Not available	Not available	No	No
Consumption	Household Budget Survey	INSTAT	No	No
Unemployment benefit receipt	Labor Force Survey, 2012	INSTAT	Yes	No
Welfare receipt	Not available	Not available	No	No

Main results of the analysis in Albania based on the UNDP Data

Table 2: Degree Type and Labor Market Entry/Exit in Albania

Degree Type	Description	Age of Entry to Labor Force	Age of Starting Pension
type 0	incomplete primary education	16	62.5
type 1	completed 9 grades of education	16	62.5
type 2	secondary and vocational education	18	62.5

Table 3: Discounted total revenues and expenditures (local currency units) in Albania

Account	type 0	type 1	type 2
Social Security & PIT revenues	579,684	757,962	750,508
Old-age pension expenditures (negative)	-53,523	-52,033	-19,753
Consumption tax revenues	67,731	78,845	64,813
Welfare expenditures (negative)	-28,063	-44,405	-4,024
SUM	565,829	740,369	791,544
ROUNDED SUM	**566,000**	**740,000**	**792,000**

Note: Estimates used a discount rate of 5 percent and discounted benefits up to age 4.

Table 4: Budget effects in Albania

Difference in educational attainment	Type Comparison	Local Currency	EUR
Primary education vs less than primary	type 1 vs 0	174,000	5,800
Secondary education vs less than primary	type 2 vs 0	226,000	7,533
Secondary education vs primary	type 2 vs 1	52,000	1,733

Note: Exchange rate calculated at 30 (local currency/EUR)

These calculations show our estimates of revenues from personal income taxes, social security contributions and consumption taxes, as well as expenditures on unemployment benefits and welfare benefits using the UNDP data. The estimates are carried out for a typical individual with three different levels of education: incomplete primary education, completed nine grades of education, and secondary education (including vocational degree). All money amounts refer to estimates over an individual's lifetime and are discounted to age four. The main results of the analysis are the estimated budget effects: these are calculated as the difference in the total contribution (net total expenditures) by a typical individual with higher level of education and a typical individual with lower level of education. These estimated effects are meant to approximate the gain of the national budget by increasing the level of education from the reference level (e.g., "less than primary" to the target level (e.g., "primary education").

The estimated budget effects of improving Roma education in Albania are an order of magnitude smaller than they were in Hungary. These results are very preliminary and may be biased downward as the UNDP data under-represents successful young Roma with higher levels. As a result, the estimated effects of higher levels of education in this sample are likely substantially smaller than the true differences in the entire society. Unfortunately, lack of adequate data prevents us from estimating more reliable effects in Albania.

Republic of Macedonia

Challenges with publicly available data in Republic of Macedonia

The table below summarizes the institutions and the source documents that have available data needed for the project. Even though these institutions have the main responsibility for data collection and make it available for the general public, frequently the data are not differentiated by education (and never by ethnicity).

Table 5: Publicly available data in the Republic of Macedonia

Data	Source Documents	Institutions	By Educational Level	By Ethnicity
Employment data	Labor Force Survey, 2012	Statistical Office, Republic of Macedonia	Yes	No
Earnings	Employees and net wages, 2012	Statistical Office, Republic of Macedonia	Yes	No
Pensions	Yearly report	Pension and Disability Insurance Fund of Republic of Macedonia	No	No
Consumption	Household consumption in RM, 2012	Statistical Office, Republic of Macedonia	No	No
Unemployment recipient & benefits	Statistical overview of the unemployed persons who are receiving benefits	Agency for Employment, Republic of Macedonia	Yes	No
Welfare recipients & benefits		Ministry of Labor and Social Affairs	Yes	No
Welfare receipt	Not available	Not available	No	No

Main results of the analysis in Republic of Macedonia based on the UNDP Data

Table 6: Degree Type and Labor Market Entry/Exit in Republic of Macedonia

Degree Type	Description	Age of Entry to Labor Force	Age of Starting Pension
type 0	less than primary	16	62
type 1	completed 8 grades of elementary	18	62
type 2	Secondary education (vocational + high school)	22	62

Table 7: Discounted total revenues and expenditures (local currency units) in Republic of Macedonia

Account	type 0	type 1	type 2
Social Security & PIT revenues	47,461	104,604	138,585
Old-age pension expenditures (negative)	-17,540	-33,494	-30,779
Consumption tax revenues	276,085	290,680	434,495
Welfare expenditures (negative)	-95,337	-96,962	0
SUM	210,669	264,829	542,302
ROUNDED SUM	**211,000**	**265,000**	**542,000**

Note: Estimates used a discount rate of 5 percent and discounted benefits up to age 4.

Table 8: Budget effects in Republic of Macedonia

Difference in educational attainment	Type Comparison	Local Currency	EUR
Primary education vs less than primary	type 1 vs 0	54,000	878
Secondary education vs less than primary	type 2 vs 0	331,000	5,382
Secondary education vs primary	type 2 vs 1	277,000	4,504

Note: Exchange rate calculated at 61.5 (local currency/EUR)

Similarly to the Albanian case study above, these calculations show our estimates of revenues from personal income taxes, social security contributions and consumption taxes, as well as expenditures on unemployment benefits and welfare benefits using the UNDP data. The estimates are carried out for a typical individual with three different levels of education: incomplete primary education, completed primary education (eight grades in Republic of Macedonia) and secondary education (including vocational degree). All money amounts refer to estimates over an individual's lifetime and are discounted to age four. The main results of the analysis are the estimated budget effects: these are calculated as the difference in the total contribution (net total expenditures) by a typical individual with higher level of education and a typical individual with lower level of education. These estimated effects are meant to approximate the gain of the national budget by increasing the level of education from the reference level (e.g., "less than primary" to the target level (e.g., "primary education").

As they are based on data from the same UNDP survey as in Albania, the estimated budget effects in Republic of Macedonia are also likely to be biased downward (the UNDP data under-represents successful young Roma with higher levels). As a result, the estimated effects of higher levels of education in this sample are likely substantially smaller than the true differences in the entire society. Unfortunately, lack of adequate data prevents us from estimating more reliable effects in Republic of Macedonia, too.

Serbia

Challenges with publicly available data in Serbia

The table below summarizes the institutions and the source documents that have available data needed for the project. The situation in Serbia is the best in comparison with the other two countries: almost all data is available by education.

Table 9: Publicly available data in Serbia

Data	Source Documents	Institutions	By Educational Level	By Ethnicity
Employment rate	Labor Force Survey, 2012	Statistical Office, Serbia	Yes	No
Average wage	Employees and net wages, 2012	Statistical Office, Serbia	Yes	No
Pensions	Yearly report	Pension and Disability Insurance Fund of Serbia	Yes	No
Welfare assistance	Welfare system in Serbia 2011,	Statistical Office,	Yes	No
Consumption	Household consumption	Statistical Office, Serbia	No	No
Unemployment recipient & benefits	Labor Force Survey, 2012	Statistical Office, Serbia	Yes	No
Welfare recipients & benefits		Ministry of Labor and Social Affairs	Yes	No

Main results of the analysis in Serbia based on the UNDP Data

Table 10: Degree Type and Labor Market Entry/Exit in Serbia

Degree Type	Description	Age of Entry to Labor Force	Age of Starting Pension
type 0	incomplete primary	15	62
type 1	completed 8 grades primary	18	62
type 2	secondary	22	62

Table 11: Discounted total revenues and expenditures (local currency units) in Serbia

Account	type 0	type 1	type 2
Social Security & PIT revenues	21,407	100,342	318,555
Old-age pension expenditures (negative)	-135,123	-157,643	-202,684
Consumption tax revenues	988,112	1,129,271	1,305,719
Welfare expenditures (negative)	-109,822	-104,359	-67,333
SUM	764,575	967,610	1,354,257
ROUNDED SUM	**765,000**	**968,000**	**1,354,000**

Note: Estimates used a discount rate of 5 percent and discounted benefits up to age 4.

Table 12: Budget effects in Serbia

Difference in educational attainment	Type Comparison	Local Currency	EUR
Primary education vs less than primary	type 1 vs 0	203,000	1,750
Secondary education vs less than primary	type 2 vs 0	589,000	5,078
Secondary education vs primary	type 2 vs 1	386,000	3,328

Note: Exchange rate calculated at 116 (local currency/EUR)

Similarly to the Albanian and Macedonian case studies above, these calculations show our estimates of revenues from personal income taxes, social security contributions and consumption taxes, as well as expenditures on unemployment benefits and welfare benefits using the UNDP data. Similarly to the Albanian and Macedonian case studies, the estimates are carried out for a typical individual with three different levels of education: incomplete primary education, completed primary education (eight grades in Serbia) and secondary education (including vocational degree). All money amounts refer to estimates over an individual's lifetime and are discounted to age four. The main results of the analysis are the estimated budget effects: these are calculated as the difference in the total contribution (net total expenditures) by a typical individual with higher level of education and a typical individual with lower level of education. These estimated effects are meant to approximate the gain of the national budget by increasing the level of education from the reference level (e.g., "less than primary" to the target level (e.g., "primary education").

As they are based on data from the same UNDP survey as in Albania and Republic of Macedonia, the estimated budget effects in Serbia are also likely to be biased downward (the UNDP data under-represents successful young Roma with higher levels). As a result, the estimated effects of higher levels of education in this sample are likely substantially smaller than the true differences in the entire society. Unfortunately, lack of adequate data prevents us from estimating more reliable effects in Serbia, too.