

Beautiful Serbia*

Holger Bonin[†]
ZEW Mannheim, IZA

Ulf Rinne[‡]
IZA

January 26, 2009

Abstract

This paper studies the causal impact of participation in an active labor market program—the Beautiful Serbia program providing training and temporary work in the construction sector in Serbia—on labor market outcomes as well as on measures of subjective well-being approximating individual welfare. According to our estimates, the positive impact of this particular program appears much stronger judged by subjective well-being than judged by the immediate labor market effect.

Keywords: Program Evaluation, Active Labor Market Policy, Subjective Well-Being, Propensity Score Matching

JEL Classification: J68, H43, P27

*The data used in this paper originate from the evaluation of the Beautiful Serbia program on behalf of UNDP Serbia and Montenegro. We would like to thank Mihael Arandarenko, Mariana Kotzeva, Pavle Golcin, Nenad Moslavac, and participants at the IZA/World Bank Conference on Employment and Development in Berlin for helpful discussions and comments. All remaining errors are our own.

[†]Holger Bonin, ZEW, P.O. Box 103443, D-68034 Mannheim. Email: bonin@zew.de

[‡]Ulf Rinne, IZA, P.O. Box 7240, D-53072 Bonn. Email: rinne@iza.org

1 Introduction

This paper studies the causal impact of participation in an active labor market policy (ALMP) measure—the Beautiful Serbia program providing training and temporary work in the construction sector in Serbia—on labor market outcomes as well as on subjective well-being approximating individual welfare. According to our estimates, the positive impact of this particular program appears much stronger judged by subjective well-being than judged by the immediate labor market effect.

Our study goes beyond the scope of traditional evaluation analysis which focuses on economic outcomes, i.e., judges the success of a labor market program by comparing the employment rates, unemployment rates, or wages of individuals who participate to the outcomes of comparable individuals who do not.¹ But a program may be beneficial for participants even if it does not immediately improve their labor market situation. It may reduce the psychic costs of being unemployed by strengthening participants' self-confidence or social contacts, and thus improve the subjective level of well-being. Lechner and Wunsch (2006) give specific examples for different spheres—other than earnings and employment in the primary labor market—treatment effects can materialize in. They mention the facts of receiving earnings from work instead of benefits and having a daily routine in this context. More generally, in the economic literature on happiness a variety of measures of subjective well-being frequently serve as proxies for individual welfare.²

Nonetheless, so far the literature evaluating specific policies with respect to their impact on individual well-being is rather scarce. For instance, Gruber and Mullainathan (2002) assess the impact of a higher tax on cigarettes on the happiness of smokers, Di Tella et al. (2003) look at the impact of changes in unemployment benefits, and Frey and Stutzer (2000) analyze the role of direct democracy for subjective well-being. In the context of ALMP, Korpi (1997) shows that program participants indicate a higher level of subjective well-being than the openly unemployed who do not attend such a program. However, our study can still be considered as one of the

¹See, e.g., Heckman et al. (1999), Martin and Grubb (2001) and Kluve (2006) for surveys of the international literature on the evaluation of ALMP measures.

²Frey and Stutzer (2002), Clark et al. (2006, 2008), and Di Tella and MacCulloch (2006) review the literature on happiness. Hayo and Seifert (2003) and Hayo (2004) show that most of the findings known from studies on the U.S. or Western Europe carry over to transition economies.

first that explicitly incorporates subjective well-being into the evaluation of ALMP.

This paper also contributes to the still relatively small literature analyzing the effectiveness of ALMP in transition economies. Papers evaluating labor market programs in Eastern Europe—with rather mixed results—include studies focusing on Poland (Puhani and Steiner, 1997; O’Leary, 1998b; Kluge et al., 1999; Puhani, 2002), Romania (Rodriguez-Planas and Benus, 2006; Rodriguez-Planas, 2007), Slovakia (Lubyova and van Ours, 1999), Hungary (O’Leary, 1998a; Micklewright and Nagy, 2005), and Estonia (Leetmaa and Vörk, 2004). In this context, our paper provides the first evaluation of an ALMP measure implemented in Serbia.

The remainder of this paper is organized as follows: Section 2 describes the background and design of the Beautiful Serbia program. Section 3 discusses our data. After explaining the evaluation strategy in Section 4, program impacts on labor market outcomes and subjective well-being are quantified in Section 5. Finally, Section 6 concludes.

2 The Beautiful Serbia Program

The Serbian economy is still considered to pass through a transitional phase. Although the country has initiated a package of economic reforms aimed at restructuring and liberalizing the economy and some positive results already materialized, high unemployment is very persistent. This is supposed to be partly an inherited problem and partly due to the prolonged and until 2000 highly irregular transition process. But also after the democratic changes of October 2000—the fall of the Milošević regime—unemployment has further increased. Table 1 displays the economic development of Serbia between 2000 and 2006. Although the economy has been improving in terms of real GDP and GDP per capita during this period, this process has surprisingly not translated into greater employment. The share of unemployed individuals among the economically active population continuously increased from about 12 (27) percent in 2001 to roughly 22 (33) percent in 2006 according to the labor force survey (administrative data).³ This picture becomes even more

³The differences in unemployment rates between administrative data and data from the labor force survey are mainly due to different underlying definitions of employment, unemployment, and participation (Arandarenko and Jovicic, 2007).

Table 1: Economic indicators in Serbia (2001–2006)

	2001	2002	2003	2004	2005	2006
GDP per capita ^a	1757	2242	2408	2643	2833	3424
GDP real growth	5.1	4.5	2.4	9.3	6.3	5.7
Unemployment rate (LFS) ^b	12.2	13.3	14.6	18.5	21.8	21.6
Unemployment rate (admin.) ^c	26.8	29.0	31.7	31.6	32.4	33.2

Source: Arandarenko and Jovicic (2007), Table 1.

Notes: ^a in Euros at exchange rate; ^b LFS: Labor Force Survey, in percent, October; ^c in percent, end of period, excl. agricultural self employment.

dramatic for an earlier period: Arandarenko (2004) reports that unemployment in Serbia was 73 percent higher in 2000 than in 1993. Therefore, the issue of active labor market programs as temporary measures to alleviate the unemployment impact of the economic transition process is ranked high on the political agenda in Serbia, at least until the conditions of a rapid and sustained economic growth—accompanied by increasing employment—are established. The program under study represents one of the first policies implemented in the country for this purpose.

The Beautiful Serbia program operated in 2004 and 2005. It was administered by the United Nations Development Program, UNDP. The program was implemented with the support of the Ministry of Labor, Employment and Social Policy, MLESP, and fully incorporated into the National Employment Service of Serbia.⁴ Due to limited financial means, the program was run on a small scale. It first started operating only in the capital city of Belgrade. In a second stage, which took place mostly in 2005, the geographic focus of the program shifted to the major cities of Niš and Zrenjanin. These three cities are economically quite heterogeneous (see Table 2). Overall, Belgrade appears to be the economically more advantaged region both in terms of the unemployment rate and the GDP level. However, in all three cities the unemployment rate reached a peak a few years ago (2002 in Belgrade, 2003 in Niš and Zrenjanin) and decreased afterwards. Additionally, the GDP levels somehow converged in recent years and are fairly equal in 2006.

The basic design of the Beautiful Serbia program was intended to replicate the Beautiful Bulgaria program, an active labor market and refurbishing program

⁴Besides UNDP and MLESP, also the governments of Canada, the Netherlands, Austria and Greece, as well as city beneficiaries contributed to financing the program.

Table 2: Regional economic indicators (2001–2005)

	2001	2002	2003	2004	2005
<i>Unemployment rate by district</i>					
City of Belgrade	20.3	24.7	22.9	21.1	18.9
Nišavski (incl. Niš)	36.5	38.9	41.9	38.2	32.4
Srednje-banatski (incl. Zrenjanin)	33.6	38.2	41.6	40.2	38.5
<i>GDP level by city (Republic of Serbia = 100)</i>					
Belgrade	157.23	164.26	152.00	142.50	119.40
Niš	121.46	119.02	105.89	114.80	109.50
Zrenjanin	111.50	116.11	91.98	110.40	126.90

Source: Arandarenko and Jovicic (2007), Table 2, and Statistical Office of the Republic of Serbia.

which had run on a nation-wide scale in Bulgaria. The apparent success of this earlier program led officials to the assumption that it could be adapted to successfully work also in Serbia. The Beautiful Serbia program consists of two different components: (a) provision of vocational training for disadvantaged unemployed persons, and (b) subsequent provision of temporary jobs restricted to the (any) disadvantaged unemployed. The two components of the program are basically independent. Participation in the vocational training stage is neither a necessary nor a sufficient condition for obtaining a job offer in the temporary employment stage.

To be specific, the vocational training measure of the Beautiful Serbia program lasts for three months and is full-time. It provides certified vocational training for the constructional sector as mason, carpenter or painter. Its intended target group consists of long-term and otherwise disadvantaged unemployed persons, identified as such by the National Employment Service. No sanctions are applied if a person refuses to participate, and participation in the training measure can be considered as voluntary. Participants in vocational training receive a compensation amounting to about 30 percent of the average national wage. As only a very small fraction of the job seekers in Serbia are entitled to income support, this appears like a substantial incentive to take up the program.⁵ Nonetheless it turned out difficult to attract individuals to the vocational training stage. One possible explanation is that participants supposedly face substantial opportunity costs in terms of forgone wage

⁵The participants who were entitled to any kind of income support before the training received 110 percent of this amount during the period of training.

earnings. A large share of the unemployed in Serbia actually make their living from informal activities.⁶ As the vocational training in the Beautiful Serbia program is full-time, participation is difficult to reconcile with these activities. We would thus expect that only those individuals expecting to recover the opportunity costs of their investment into human capital (self-)select into the training measure.

The second component of the Beautiful Serbia program is the provision of temporary jobs in the construction sector. Typical for a transition economy at its first growth stage, construction still plays an important role in the Serbian economy. Naturally temporary project-based jobs show a high incidence in this sector. The Beautiful Serbia program creates additional demand for these jobs by financing the refurbishment of selected public buildings and spaces. In the refurbishment projects, private firms are contracted under the condition that they employ a specified share (40–60 percent) of workers who are identified by the National Employment Service as previously unemployed and otherwise disadvantaged. Firms receive a fixed payment for conducting the refurbishment project. Projects are assigned to firms on a competitive basis, i.e., the firm offering the best quality-price ratio wins the tender. This procedure should guarantee that wages paid on the jobs in the temporary employment stage of the Beautiful Serbia program are competitive. In particular, firms do not receive a special wage subsidy for hiring the mandatory number of previously unemployed workers. The contracted firms can select among the pool of people who meet the criteria of the National Employment Service. Individuals run through an ordinary application procedure. Hence one would expect that successful candidates are hired in accordance with the needs of the company, and represent the most competent and capable among the unemployed individuals firms can choose among. This means that it is neither necessarily the case that participants in the vocational training part of the Beautiful Serbia program later on work in the sponsored refurbishment projects, nor that the previously unemployed workers hired for these projects did participate in the training measure before.

In total the Beautiful Serbia program provided vocational training to 252 unemployed individuals. The drop-out rate at this stage was very low. Almost 95 percent of the enrolled completed the training. In the 35 refurbishment projects fi-

⁶Schneider (2007) presents estimates for the size of the Serbian shadow economy. It amounts to 37.3 percent of the official GDP in 2004/05.

nanced by the program—managed by 16 contracted private companies—321 men found temporary employment.⁷ About half of them had participated in the vocational training stage before.

3 Data

Our data come from a special survey of 363 individuals who were registered as unemployed at the National Employment Service when the program started (January 2004) and who either participated in at least one stage of the Beautiful Serbia program or did not participate at all. The interviews were conducted face-to-face by a professional survey agency, GfK Belgrade, shortly after the final refurbishment project of the program had been completed, during October and November 2005.

In principle, the survey was constructed such as to mimic an experimental design *ex post*. For each individual who participated in the Beautiful Serbia program, a matched partner with the same observable characteristics was drawn from the unemployment registry and scheduled for interview. The intention was to create a control group, which would resemble the treatment group as much as possible, with a limited number of interviews. Unfortunately, due to deficiencies of unemployment registries at the National Employment Service, only few individual characteristics were available to implement this strategy. In effect, the one-to-one pre-matching routine to create a control group was only based on the following individual characteristics: age, education, and place of residence (Belgrade, Niš, or Zrenjanin). In particular, the (un-)employment history which appears extremely relevant for the success of active labor market policies could not be controlled for.⁸ In the accomplished survey, systematic drop-outs may further reduce the effective quality of the matches between program participants and non-participants. A sizeable number of persons scheduled for interview—around 40 percent—could either not be found or refused to participate in the interviews.

⁷In principle, the program was available for women, too, but actually none participated—neither in the vocational training nor in the temporary employment stage.

⁸Kluve et al. (1999) demonstrate that pre-unemployment labor market careers are also extremely important when assessing active labor market policies in transition economies, where variations in these histories tend to be smaller than in Western economies because they start from a situation of no formal unemployment.

Table 3: Distribution of observations across participation status

	Participation in temp. employment?		
Participation in training?	No	Yes	Total
No	146	28	174
Yes	48	66	114
Total	194	94	288

We observe data on 168 participants, while the potential control group of non-participants consists of 195 individuals. After dropping records with missing values on key characteristics (e.g., employment history, unemployment duration) we are left with a sample of 288 individuals.⁹ Table 3 illustrates the distribution of the retained sample regarding participation in either of the two program stages. Among the 142 participants, about one in three was exposed to the Beautiful Serbia program only through the vocational training stage, whereas about one in five was exposed to the program only through the temporary employment stage. The ratio of participants to non-participants in our working sample is close to one.

In Table 4 we present some descriptive statistics of the individuals subject to one of the three possible treatments—participation in the complete program, participation in the vocational training stage only, and participation in the temporary employment stage only—and of the individuals not participating at all who are potential controls. Substantial differences between participants and non-participants arise in our sample. In particular, across all treatments participants appear to be significantly younger than non-participants, better educated, more likely to belong to the ethnic group of Roma, and more likely to live in Belgrade. Furthermore, in January 2004 when the program started, the participants had experienced shorter spells of unemployment, had more frequently been employed in the past 36 months, and more often actively searching for a job than non-participants.

The substantial differences in observed characteristics indicate that the pre-matching routine to create a control group has not worked satisfactorily. One potential explanation would be that the selection of the control group was based on planned rather than on accomplished interviews. An alternative—and probably

⁹We also drop 7 individuals in the potential control group who exit the labor market by turning into pensioners or students.

Table 4: Descriptive statistics (selected variables)

	CP	VT	TE	NP
Age	31.09 (9.84)	31.85 (10.20)	33.36 (10.60)	34.23 (11.79)
Married	0.3182 (0.4693)	0.5000 (0.5053)	0.6786 (0.4756)	0.5822 (0.4949)
Roma	0.1061 (0.3103)	0.2083 (0.4104)	0.2143 (0.4179)	0.0822 (0.2756)
Belgrade	0.4848 (0.5036)	0.5000 (0.5053)	0.4285 (0.5040)	0.3151 (0.4661)
Education: primary school or less	0.3182 (0.4693)	0.4167 (0.4982)	0.3571 (0.4880)	0.2877 (0.4542)
Education: vocational school (3 years)	0.3333 (0.4750)	0.3333 (0.4764)	0.3571 (0.4880)	0.4110 (0.4937)
Previous unemployment duration (in months)	31.33 (37.67)	36.83 (41.78)	42.68 (50.07)	60.05 (54.69)
Employed at all in last 3 years	0.7424 (0.4407)	0.7292 (0.4491)	0.8214 (0.3900)	0.5685 (0.4970)
Actively searching for a job	0.8485 (0.3613)	0.8125 (0.3944)	0.8571 (0.3563)	0.6370 (0.4825)
# observations	66	48	28	146

Note: Mean values of selected variables (standard deviation in brackets). CP indicates participation in both the vocational training and the temporary employment stage of the program, VT (TE) indicates participation in the vocational training (temporary employment) stage of the program only, NP indicates non-participation.

more relevant—explanation would be that the probabilities to participate in the program were actually affected by individual characteristics beyond those few used by the pre-matching routine, see above. In any case, the observed characteristics of the program participants altogether appear to give them a comparative advantage over the non-participants—in particular concerning potential labor market success. Therefore, one would expect that a comparison of mean outcomes between the two groups overestimates the positive program effects. In order to avoid this bias, we need to rely on econometric techniques for constructing a control group that is effectively comparable to the treatment group.

4 Evaluation Approach

Ideally, we would like to compare the outcomes for the individuals participating in the Beautiful Serbia program (Y^1) with the outcomes for the same individuals if they had not participated (Y^0). If D denotes participation, where $D = 1$ if a person participates in the program and $D = 0$ otherwise, the actual outcome for individual i can be written as:

$$Y_i = Y_i^1 \cdot D_i + Y_i^0 \cdot (1 - D_i) . \quad (1)$$

The individual treatment effect would be given by the difference $\Delta_i = Y_i^1 - Y_i^0$. However, it is impossible to calculate this difference because one of the outcomes is unobservable. Instead, the evaluation literature concentrates on population average gains from treatment—usually on the average treatment effect on the treated (ATT or Δ_{ATT}) which is formally given by:

$$\Delta_{ATT} = E(\Delta|D = 1) = E(Y^1|D = 1) - E(Y^0|D = 1) . \quad (2)$$

It is the principle task of any evaluation study to find a credible estimate for the second term on the right hand side of equation (2), which is unobservable.

As mentioned above, one possible solution could be to simply compare the mean outcomes of participants and non-participants. But if $E(Y^0|D = 1) \neq E(Y^0|D = 0)$, estimating the ATT by the difference between the sub-population means of these two

groups will yield a selection bias—which is likely the case in our data. On the other hand, if treatment assignment is *strongly ignorable*, i.e., if selection is on observable characteristics X (unconfoundedness or conditional independence assumption), and if observable characteristics of participants and non-participants overlap (common support), the matching estimator is an appealing choice to estimate the desired counterfactual (Rosenbaum and Rubin, 1983). Under these conditions, the distribution of the counterfactual outcome Y^0 for the participants is the same as the observed distribution of Y^0 for the comparison group *conditional on the vector of covariates* X . Formally,

$$E(Y^0|X, D = 1) = E(Y^0|X, D = 0). \quad (3)$$

Entering this relation into (2) allows estimating the ATT. Rosenbaum and Rubin (1983) show that if treatment assignment is strongly ignorable *given* X , it is also strongly ignorable *given any balancing score that is a function of* X .¹⁰ One possible balancing score is the propensity score $P(X)$, i.e., the probability of participating in a given program.

There are several propensity score matching methods suggested in the literature.¹¹ Based on the characteristics of our data, we opt to apply nearest-neighbor matching with replacement. This matching method has the advantage of being the most straightforward matching estimator: a given participant is matched with a non-participant who is closest in terms of the estimated propensity score. As the participants and non-participants in our sample appear quite different, we allow matching with replacement to avoid bad matches between high-score participants and low-score non-participants. The disadvantage of this approach is that the variance of the estimator increases as the constructed counterfactual outcome is based on less distinct non-participants (Smith and Todd, 2005).

For the variance of the estimated ATT, we apply the approximation suggested by Lechner (2001, 2002). The following formula applies for nearest neighbor match-

¹⁰When there are many covariates, it is impractical to match directly on covariates because of the curse of dimensionality. See, e.g., Zhao (2008) for some comments on this problem.

¹¹See, e.g., Caliendo and Kopeinig (2008) for an overview.

ing with replacement:

$$Var(\hat{\Delta}_{ATT}) = \frac{1}{N_1} \cdot Var(Y^1|D = 1) + \frac{(\sum_{j \in \{D=0\}} (w_j)^2)}{(N_1)^2} \cdot Var(Y^0|D = 0), \quad (4)$$

where N_1 is the number of matched treated individuals and w_j is the number of times individual j from the control group is used. We checked the accuracy of this approximation by also calculating the variance of the estimated treatment effects based on bootstrapping procedures. Although nearest neighbor matching does not satisfy the basic conditions for the bootstrap and the bootstrap variance diverges from the actual variance (Abadie and Imbens, 2006), this alternative method gives similar variances of the estimated treatment effects and does not change the implications presented below.

We estimate the probability of treatment in the Beautiful Serbia program conditional on observable characteristics—the propensity score—using binary probit models with participation as the dependent variable.¹² The potential control group always consists of the individuals who did not participate in the program at all. Our preferred specifications of the propensity score include a full range of personal characteristics. We measure regional variation in program participation rates by including an index variable taking the value of one if an individual lives in Belgrade and zero otherwise. However, as all participants in Belgrade entered the program in 2004, and almost all participants outside Belgrade entered in 2005, this variable also captures the variation related to the timing of program entry.¹³

Table 5 depicts the marginal effects of the probit estimates underlying the propensity scores for the three treatments. The results basically confirm the impression from the descriptive statistics. It appears that individuals relatively close to the labor market—i.e., individuals of younger age, relatively short-term unemployed, recently employed or actively searching for a job—had a higher chance to participate in the Beautiful Serbia program.

¹²Estimations are done using the PSMATCH2 Stata ado-package by Leuven and Sianesi (2003).

¹³We have tried several specifications of the probit model. The results did not change qualitatively. For instance, including the number of (small) children living in the household does not change the predictions since all individuals in our sample are men for whom age and marital status already capture most of the effect possibly associated with children. Our preferred specifications appear to deliver the best overall predictions of program participation rates.

Table 5: Probit estimates (marginal effects)

	CP vs. NP	VT vs. NP	TE vs. NP
ln(Age)	-98.8535*	-40.0550	7.4593
ln(Age) ²	29.3678*	12.2453	-1.7225
ln(Age) ³	-2.8845*	-1.2343	0.1222
Married	-0.2717***	-0.1577*	0.0331
Roma	0.1607	0.2182	0.1851*
Belgrade	0.0538	0.1339*	0.0658
Homeowner	0.1540	0.1931**	0.4381***
Education: primary school or less	0.1157	0.1577*	0.0442
Education: vocational school (3 years)	0.0127	0.1022	0.0534
Disabled		-0.0254	-0.0046
Mobile	-0.1169	-0.1675**	-0.0634
Unemployed \leq 12 months	0.4141***	0.2587**	0.2331***
Unemployed 13–24 months	0.2749***	0.2156***	0.0178
Unemployed 25–36 months	0.2165**	0.4246**	0.1222
Unemployed 37–48 months	0.2712**	0.1826	0.1062
Employed in last 3 years	0.0959	0.0975	0.2091***
Share of employment in last 3 years	-0.1704	-0.2355*	-0.2510**
Other income	-0.2669	-0.2005	-0.1196
Jobsearcher	0.1971**	0.1296*	0.1061**
ALMP	-0.2503**	-0.1240	
Jobdesire	0.1386	0.0597	0.0612
Jobchances	0.0594	0.2020**	-0.0479
Jobchances \times Jobdesire	0.0551		
Jobchances \times Employed in last 3 years	-0.0381		
Jobchances \times Roma	-0.0246		
Roma \times Belgrade	0.5449	0.0419	
Roma \times Homeowner	-0.2600	-0.1661**	
Roma \times Married		0.0991	
Mobile \times Education: primary school or less	-0.2494		
Jobsearcher \times Unemployed 25–36 months		-0.1029	
Employed in last 3 years \times Homeowner			-0.1156**

Note: *** significant at 1%; ** significant at 5%; * significant at 10%. CP indicates participation in both the vocational training and the temporary employment stage of the program, VT (TE) indicates participation in the vocational training (temporary employment) stage of the program only, NP indicates non-participation.

The distributions of the propensity scores obtained from the probit estimates are on display in Figure 1. A comparison of participants and non-participants reveals that the latter tend to be endowed with characteristics that make them systematically less likely to be selected for participation in the Beautiful Serbia program. Among the individuals participating in both stages of the program, 4 have a higher propensity score than the individual with the highest estimated propensity score among the non-participants. Hence these individuals are off support according to the usual ‘Minmax’ criterion and need to be excluded for the computation of the ATT. To achieve common support, we need to exclude 5 (3) observations when evaluating participation in the vocational training (temporary employment) stage only.

After forming the matched pairs, a suitable way to assess the matching quality is comparison of the standardized bias before matching, SB^b , to the standardized bias after matching, SB^a . The standardized biases are defined as

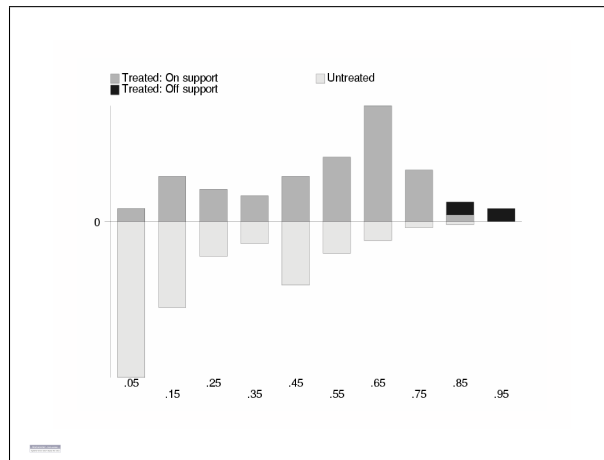
$$SB^b = \frac{(\bar{X}_1 - \bar{X}_0)}{\sqrt{0.5 \cdot (V_1(X) + V_0(X))}} ; SB^a = \frac{(\bar{X}_{1M} - \bar{X}_{0M})}{\sqrt{0.5 \cdot (V_{1M}(X) + V_{0M}(X))}} , \quad (5)$$

where X_1 (V_1) is the mean (variance) in the treated group before matching and X_0 (V_0) the analogue for the comparison group. X_{1M} (V_{1M}) and X_{0M} (V_{0M}) are the corresponding values after matching (Rosenbaum and Rubin, 1985). Following the example of Sianesi (2004) we also re-estimate the propensity score on the matched sample to compute the pseudo- R^2 before and after matching.

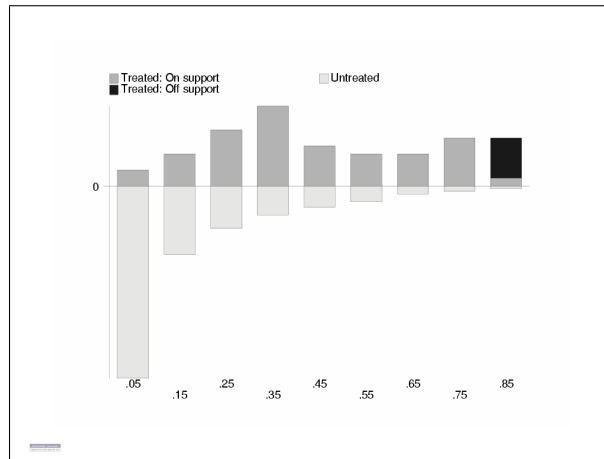
These measures (see Table 6) suggest that the quality of our matching procedures is quite satisfactory. The standardized bias of the matched sample is markedly smaller than that of the unmatched sample. Likewise, the pseudo- R^2 after matching are fairly low and decrease substantially compared to before matching. This is what we should expect considering that after matching, there should not be any systematic differences in the distribution of observable characteristics between participants and matched non-participants. Therefore, the test of the matching quality makes us confident to estimate meaningful treatment effects on the basis of nearest neighbor matching with replacement, despite of the rather small sample available for constructing the matched pairs.

Figure 1: Distribution of propensity scores, common support

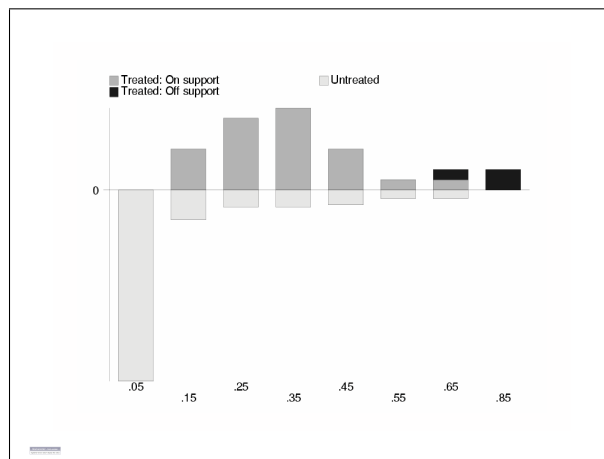
(a) CP vs. NP



(b) VT vs. NP



(c) TE vs. NP



Note: CP indicates participation in both the vocational training and the temporary employment stage of the program, VT (TE) indicates participation in the vocational training (temporary employment) stage of the program only, NP indicates non-participation.

Table 6: Matching quality

	CP vs. NP	VT vs. NP	TE vs. NP
# treated individuals	66	48	28
# treated individuals off support	4	5	3
# matched pairs	62	43	25
Mean SB before matching	0.1962	0.2467	0.1965
Mean SB after matching	0.0862	0.0882	0.1001
Pseudo- R^2 before matching	0.2573	0.2872	0.2890
Pseudo- R^2 after matching	0.1363	0.1688	0.1139

Note: The mean SB is calculated as the mean of the single characteristics' SB (in percent). CP indicates participation in both the vocational training and the temporary employment stage of the program, VT (TE) indicates participation in the vocational training (temporary employment) stage of the program only, NP indicates non-participation.

5 Treatment Effects

In the following we first adopt the conventional perspective of ALMP evaluation studies and assess the causal impact of the Beautiful Serbia program on labor market outcomes, i.e., on unemployment and employment probabilities. In a second step, we will look at subjective well-being variables at the core of our interest.

5.1 Labor Market Outcomes

The survey data do not trace individuals' employment histories. Hence our outcome variables are based on the labor market status at the time of the interview. More precisely, we consider four different labor market states: (a) unemployment, (b) employment in a regular job including self-employment, (c) employment in a seasonal job, and (d) employment in another active labor market program implemented by the National Employment Service ('ALMP job'). Table 7 summarizes the estimated ATT for these four different labor market outcomes and the three possible treatments.

Our point estimates suggest that participation in both stages of the program reduces the probability of being unemployed at the survey date by about 13 percentage points, compared to not participating in the program. Participation in the training stage only reduces the unemployment rate by 7 percentage points, whereas participation in the temporary employment stage apparently has no effect on the

Table 7: ATT labor market outcomes

	CP vs. NP	VT vs. NP	TE vs. NP
Unemployment	-0.1290	-0.0698	0.0000
Regular job	0.1290	0.0465	0.1200
Seasonal job	-0.0161	0.0930	-0.1600
ALMP job	0.0323	-0.0698	0.0400
# matched pairs	62	43	25

Note: *** significant at 1%; ** significant at 5%; * significant at 10%. CP indicates participation in both the vocational training and the temporary employment stage of the program, VT (TE) indicates participation in the vocational training (temporary employment) stage of the program only, NP indicates non-participation.

propensity of being unemployed. The latter result is estimated on very few observations, however. In fact, none of the estimated ATT is statistically significant at conventional levels. In general, the small scale of the program and therefore small sample sizes will only yield significant ATT if participants and matched non-participants exhibit very distinct outcomes.

Considering overall employment, the ATT basically mirror those concerning unemployment. Participation in the program is generally associated with a higher employment rate. However, some differences appear between the different treatments concerning the type of employment.¹⁴ Participation in the complete program positively affects the chances of working in a regular job while the chances of seasonal employment or of an ALMP job basically remain unaffected. In contrast, for participation in the training stage of the program only, the strongest program impact is on employment in a seasonal job. The effect on seasonal employment is even stronger than the overall employment effect: Participation in the training measure reduces the chance to become employed in another active labor market program. Finally, while participation in the temporary employment stage of the program only has basically no effect on the overall employment rate, it seems to affect the type of employment: Treated individuals appear to work more frequently in regular jobs and less frequently in seasonal jobs.

In sum, our findings concerning the impacts of the Beautiful Serbia program on labor market outcomes suggest that both the vocational training and the temporary

¹⁴The overall ATT concerning employment is the sum of the ATT regarding the three different types of employment.

employment part (and therefore the program taken as a whole) exert a positive influence on the employment prospects of the participants. However, the positive effects are not sufficiently strong or clear-cut to be considered statistically significant.

5.2 Subjective Well-Being

Even if an active labor market program does not immediately raise the employment probabilities of participants, a social planner may find it beneficial if it manages to improve the individual welfare of the target group. The survey data collected in connection with the Beautiful Serbia program provide us with the unique opportunity to study program impacts also on various dimensions of life that may serve to approximate individual well-being or ‘happiness’.

In the literature, happiness is usually measured by the answer to a very broad question. For instance, the U.S. General Social Survey asks: *Taking things all together, how would you say you are these days—would you say you are very happy, pretty happy, or not too happy?* The individuals in our data were not asked for such a global assessment of their whole sphere of life. Instead, we observe answers to a set of questions relating to items that give a reasonable picture of how their personal situation concerning various aspects of life has changed over time. Individuals were requested to compare their situation at the time of the interview with that before the Beautiful Serbia program came into effect, and had to judge whether their situation has strongly or somewhat improved, has stayed more or less the same, or has strongly or somewhat deteriorated.¹⁵ In detail, the survey requested a self-assessment of changes concerning self-confidence, the desire to find a job, social contacts, health status, the family income situation, personal qualification and skills, and the chances to find a regular job. These items have been identified as determinants of personal happiness (Frey and Stutzer, 2002). However, the extent to which the different items are related to subjective well-being varies. For example, personal health ratings and happiness appear to be highly correlated, whereas changes in income are considered to have only temporary impacts on sub-

¹⁵Individuals from Belgrade where the program was introduced earlier were asked to compare their situation to that in the beginning of 2004, and individuals from in Niš and Zrenjanin where the program started later were asked to compare their situation to that in the beginning of 2005.

Table 8: ATT indicators of subjective well-being.

	CP vs. NP	VT vs. NP	TE vs. NP
Self-confidence	0.1129	0.2093 *	0.2800 **
Job desire	0.2419 **	0.2558 **	0.1200
Social contacts	0.1451	0.1860	0.2800 **
Qualification and skills	0.3387 ***	0.5116 ***	0.2400 **
Health	0.1774 **	0.0233	0.0000
Job chances	0.1129	0.0698	0.2400 ***
# matched pairs	62	43	25

Note: *** significant at 1%; ** significant at 5%; * significant at 10%. CP indicates participation in both the vocational training and the temporary employment stage of the program, VT (TE) indicates participation in the vocational training (temporary employment) stage of the program only, NP indicates non-participation.

jective well-being, probably due to the phenomenon of adaptation (Layard, 2006). The dimensions of ‘qualification and skills’ and ‘job chances’ contain information on *employability*, which is a more general concept than actual employment.¹⁶ An improvement of subjective employability probably reduces the psychic cost of being unemployed, and thus may put individuals higher on the happiness scale.

In our subsequent analysis, we apply our matching approach to the subjective data. As outcome variables, we define dummy variables that take the value of one if individuals report that their personal situation has strongly or somewhat improved, and take a value of zero otherwise. In this way, the ATT measure the change in the percentage share of individuals judging their personal as improved because of program participation. Table 8 summarizes our findings.

The general impression based on the point estimates is that program participation has improved the personal situation with regard to all aspects of life considered. In contrast to the impact on labor market outcomes, the program effects often appear so substantial that the estimated ATT are statistically significant despite the small sample sizes on which they are estimated. For all treatments, the strongest program impact is on the subjective rating of qualification and skills. The impact is particularly strong for participants in the vocational training stage only, followed by participants in both stages of the program. This means that the vocational train-

¹⁶Following McQuaid and Lindsay (2005), the concept of employability we have in mind is a narrow one—thereby focusing on individual factors, i.e., essential attributes and personal competencies—and close to the operational versions of the ‘socio-medical employability’ and the ‘manpower policy employability’.

ing content of the program is viewed positively from the participants' perspective even when it does not immediately raise the employment rate. Among those individuals who participated in both stages of the Beautiful Serbia program, the share with improved job desire and improved self-assessed health is significantly higher than among comparable individuals who were not affected by the program. Similar positive effects arise considering those participating only in the training stage, which furthermore appears to significantly improve self-confidence. A strong self-confidence effect also occurs for those participating in the temporary employment stage only. Personal relations established at the work or training place are probably responsible for the clear growth of social contacts (15–28 percentage points) achieved through the Beautiful Serbia program.

It is interesting to note that the program—though focused on the construction sector offering probably relatively poor working conditions—if anything positively impacts on subjective health status. Among the individuals participating in the whole program, the rate of those reporting an improvement in health compared to the pre-program situation is 18 percentage points larger than among non-participants, and the effect is statistically significant. The ATT concerning health status are much smaller for the other two treatments, but, judged by the point estimates, they are at least not negative.

Taken together, the positive program effects considering individuals' subjective assessment of conditions of life appear to be larger than the program impacts when considering their objective labor market status. This suggests that the program improves subjective well-being through other channels than the labor market. The impacts we find are strong for all treatments considered. Even the subjective health rating—a key determinant of happiness—significantly increases for those going through the complete program.

6 Conclusion

This paper evaluates the Beautiful Serbia program providing vocational training and temporary employment to disadvantaged unemployed. While using standard matching techniques to bring out causal average treatment effects on the treated, the

analysis deviates from routine program evaluation by considering subjective measures of individual well-being as possible outcomes. Hence this paper is linked to the rising economic literature focused on the concept of happiness as an approximation for the individual welfare scale.

Given that the ultimate goal of social policies is the improvement of individual welfare, subjective well-being clearly is a relevant dimension for a full impact assessment of an active labor market program. The evaluation results obtained from the Beautiful Serbia program indeed provide an example that the positive effects of a policy can appear stronger if it is judged by subjective well-being rather than by labor market effects. The program probably impacted on individual welfare through other channels than the immediate economic status, notably by strengthening self-confidence, job desire and social inclusion of the participants.

Unfortunately, due to the small scale of the program and certain deficiencies in the accomplished survey, the treatment effects estimated for the Beautiful Serbia program overall allow only tentative conclusions. The systematic inclusion of subjective measures of well-being into the evaluation of a larger-scale program as well as the inclusion of more direct measures of the individual happiness scale that are also tested for behavioral relevance thus remain a challenge for future research.

References

- Abadie, A. and G. W. Imbens (2006). On the Failure of the Bootstrap for Matching Estimators. NBER Technical Working Paper 325, National Bureau of Economic Research, Inc.
- Arandarenko, M. (2004). International Advice and Labor Market Institutions in South-East Europe. *Global Social Policy* 4(1), 27–53.
- Arandarenko, M. and M. Jovicic (2007). Regional Labour Market Differences in Serbia: Assessment and Policy Recommendations. *The European Journal of Comparative Economics* 4(2), 299–317.
- Caliendo, M. and S. Kopeinig (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys* 22(1), 31–72.
- Clark, A. E., P. Frijters, and M. A. Shields (2006). Income and Happiness: Evidence, Explanations and Economic Implications. PSE Discussion Paper 2006-24, PSE (Ecole normale supérieure).
- Clark, A. E., P. Frijters, and M. A. Shields (2008). Relative Income, Happiness and Utility: An Explanation for the Easterlin Paradox and Other Puzzles. *Journal of Economic Literature* 46(1), 95–144.
- Di Tella, R. and R. J. MacCulloch (2006). Some Uses of Happiness Data in Economics. *The Journal of Economic Perspectives* 20(1), 25–46.
- Di Tella, R., R. J. MacCulloch, and A. J. Oswald (2003). The Macroeconomics of Happiness. *The Review of Economics and Statistics* 85(4), 809–827.
- Frey, B. and A. Stutzer (2000). Happiness, Economy, and Institutions. *The Economic Journal* 110(466), 918–938.
- Frey, B. and A. Stutzer (2002). *Happiness and Economics: How the Economy and Institutions Affect Human Well-Being*. Princeton and Oxford: Princeton University Press.
- Gruber, J. and S. Mullainathan (2002). Do Cigarette Taxes Make Smokers Happier? NBER Working Paper 8872, National Bureau of Economic Research, Inc.
- Hayo, B. (2004). Happiness in Eastern Europe. Marburg Working Papers on Economics 200412, Philipps-Universität Marburg, Faculty of Business Administration and Economics, Department of Economics.
- Hayo, B. and W. Seifert (2003). Subjective Economic Well-being in Eastern Europe. *Journal of Economic Psychology* 24(3), 329–348.
- Heckman, J. J., R. J. LaLonde, and J. A. Smith (1999). The Economics and Econometrics of Active Labor Market Programs. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3. Amsterdam: Elsevier.
- Kluge, J. (2006). The Effectiveness of European Active Labor Market Policy. IZA Discussion Paper 2018, Institute for the Study of Labor (IZA), Bonn.
- Kluge, J., H. Lehmann, and C. M. Schmidt (1999). Active Labor Market Policies in Poland: Human Capital Enhancement, Stigmatisation, or Benefit Churning? *Journal of Comparative Economics* 27, 61–89.

- Korpi, T. (1997). Is Utility Related to Employment Status? Employment, Unemployment, Labor Market Policies and Subjective Well-Being Among Swedish Youth. *Labour Economics* 4(2), 125–147.
- Layard, R. (2006). Happiness and Public Policy: A Challenge to the Profession. *The Economic Journal* 116(510), C24–33.
- Lechner, M. (2001). Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption. In M. Lechner and F. Pfeiffer (Eds.), *Econometric Evaluation of Labour Market Policies*, pp. 1–18. Heidelberg: Physica-Verlag.
- Lechner, M. (2002). Some Practical Issues in the Evaluation of Heterogeneous Labour Market Programmes by Matching Methods. *Journal of the Royal Statistical Society, Series A, Statistics in Society* 165(1), 59–82.
- Lechner, M. and C. Wunsch (2006). Active Labour Market Policy in East Germany: Waiting for the Economy to take Off. IZA Discussion Paper 2363, Institute for the Study of Labor (IZA), Bonn.
- Leetmaa, R. and A. Võrk (2004). Evaluation of Active Labour Market Programmes in Estonia. *mimeo*.
- Leuven, E. and B. Sianesi (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Statistical Software Components, Boston College Department of Economics. Available at <http://ideas.repec.org/c/boc/bocode/s432001.html>.
- Lubyova, M. and J. C. van Ours (1999). Effects of Active Labor Market Programs on the Transition Rate from Unemployment into Regular Jobs in the Slovak Republic. *Journal of Comparative Economics* 27(1), 90–112.
- Martin, J. P. and D. Grubb (2001). What Works and for Whom: A Review of OECD Countries' Experiences with Active Labor Market Policies. *Swedish Economic Policy Review* 8(2), 9–56.
- McQuaid, R. W. and C. Lindsay (2005). The Concept of Employability. *Urban Studies* 42(2), 197–219.
- Micklewright, J. and G. Nagy (2005). Job Search Monitoring and Unemployment Duration in Hungary: Evidence from a Randomised Control Trial. IZA Discussion Paper 1839, Institute for the Study of Labor (IZA), Bonn.
- O'Leary, C. J. (1998a). Evaluating the Effectiveness of Active Labor Market Programs in Hungary. Upjohn Institute Technical Report 98-013, W.E. Upjohn Institute for Employment Research, Kalamazoo, Michigan.
- O'Leary, C. J. (1998b). Evaluating the Effectiveness of Active Labor Market Programs in Poland. Upjohn Institute Technical Report 98-012, W.E. Upjohn Institute for Employment Research, Kalamazoo, Michigan.
- Puhani, P. (2002). Advantage through Training in Poland? A Microeconomic Evaluation of the Employment Effects of Training and Job Subsidy Programmes. *Labour* 16(3), 569–608.

- Puhani, P. and V. Steiner (1997). The Effectiveness and Efficiency of Active Labour Market Policies in Poland. *Empirica* 24(3), 209–231.
- Rodriguez-Planas, N. (2007). What Works Best for Getting the Unemployed Back to Work: Employment Services or Small-Business Assistance Programmes? Evidence from Romania. IZA Discussion Paper 3051, Institute for the Study of Labor (IZA), Bonn.
- Rodriguez-Planas, N. and J. Benus (2006). Evaluationg Active Labor Market Programs in Romania. IZA Discussion Paper 2464, Institute for the Study of Labor (IZA), Bonn.
- Rosenbaum, P. R. and D. B. Rubin (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70(1), 41–55.
- Rosenbaum, P. R. and D. B. Rubin (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score. *The American Statistician* 39(1), 33–38.
- Schneider, F. (2007). Shadow Economies and Corruption All Over the World: New Estimates for 145 Countries. *Economics – The Open-Access, Open-Assessment E-Journal* 1(9).
- Sianesi, B. (2004). An Evaluation of the Active Labor Market Programmes in Sweden. *The Review of Economics and Statistics* 86(1), 133–155.
- Smith, J. A. and P. E. Todd (2005). Does Matching Overcome LaLonde’s Critique of Non-Experimental Estimators? *Journal of Econometrics* 125(1-2), 305–353.
- Zhao, Z. (2008). Sensitivity of Propensity Score Methods to the Specifications. *Economics Letters* 98(3), 309–319.