Human Capital Externalities:
Evidence from the Transition Economy of Russia

Alexander Muravyev[1]

Abstract

The paper tests for the existence of human capital externalities, more precisely those stemming from higher education, using a micro-level approach: the Mincerian wage regression augmented with the average level of education in a local geographical area (city). To solve identification problems arising due to endogeneity of average education the study exploits a natural experiment provided by the process of economic transition in the former communist economies. We argue that the educational structure of cities under the central planning was determined by the government rather than the market; thus the average educational attainment in cities at the end of communism can be regarded as exogenous with respect to wages prevailing after the start of transition. The identification strategy based on the use of the pre-transition average education is applied to data from the Russia Longitudinal Monitoring Survey, RLMS. Empirical results are consistent with the presence of significant human capital (educational) externalities in the Russian economy. According to the estimates, one percent increase in the college share in a city results in the increase of city residents' wages by 1-2 percent. The result proves to be robust to several changes in the empirical specification.

JEL classification: I2, J31
Keywords: Human Capital Externalities, Cities, Russia

This version: April 2006

---

[1] Department of Economics, European University Institute, Florence.
E-mail: alexander.muravyev@iue.it

I would like to thank Andrea Ichino, Francis Vella, Mark Schaffer, Randall Filer, John Earle, Tom Cupé, and Stepan Jurajda for discussions and comments at various stages of work on this paper. All remaining mistakes are my own. I acknowledge financial support from the Economics Education and Research Consortium (EERC), grant No. R04-0241.

1. Introduction

Human Capital Theory states that individuals invest in their human capital – by spending money and time on education and training – in order to enhance own productivity which is in turn rewarded by higher wages. Consistent with this supposition is a vast amount of empirical studies demonstrating that the private return to education – an increase of individual's earnings resulting from one additional year of schooling – falls in the range of 6 to 10% (Card, 1999). The theory asserts that investments in human capital are undertaken by individuals until the point where the marginal productivity gained equals the marginal opportunity cost (from the individual's viewpoint).

Benefits of human capital accumulation[2] by a person need not pertain to that person solely. An individual's investment in her own human capital may also increase productivity of the other factors of production, i.e. physical capital or human capital of others. Importantly, the channels of such influence – the most prominent of which is sharing of knowledge and skills trough formal and informal interaction between people in the same industry, city, region or economy – may not be internalised within firms or families. This gives rise to human capital externalities. Since Lucas (1988) contribution, the hypothesis on human capital externalities has become a standard modelling tool in the New Growth Theory where they are regarded as a major factor of sustainable growth (Spagat, 2002).[3]

Much of the interest in this area is explained by important policy implications of human capital externalities. This primarily concerns education which is often regarded as a primary means of human capital accumulation. If education has the characteristics of a public good, the private returns to education may underestimate the full returns to the society – an increase in the total earnings resulting from a one-year increase in the average schooling. In such case, education is not provided at the efficient scale and public investment in education is desirable.

The existence of educational and, more generally, human capital externalities remains a questionable issue from both theoretical and empirical perspectives. In theory it has been suggested, for example, that education may play a pure signalling role. If

[2] In addition to education and training, human capital accumulation may occur through the learning that experience yields.

[3] Besides static productivity externalities such as those suggested by Lucas, higher stock of human capital may facilitate creation and adoption of new technologies or make learning-by-doing more effective thus leading to dynamic human capital externalities (Venniker, 2000). In addition, there may be other external

education is nothing else but a signal of individual's innate ability, the social returns to education are zero: the aggregate income stays unchanged when all workers increase their schooling by one year.

Empirical evidence on human capital spillovers remains scarce and inconclusive, as emphasized in several recent surveys (e.g., Moretti, 2004b; Psacharopoulos and Patrinos, 2002, Sianesi and Van Reenen 2002; Venniker, 2000). Two types of studies may be distinguished depending on whether they are based on micro- or macro-data. In the empirical macro-economic literature, the stock of human capital is typically used to explain either the long-run level or the long-run growth rate of the economy. The micro approach explores whether, given a worker's educational level (and possibly other characteristics), her wage rises with the average educational level attained in a relevant geographical area, usually a city.

Krueger and Lindahl (2001) suggest that the micro approach is less suitable for uncovering the social returns to education since it defines educational externalities in a limited way. Indeed, education may affect national income in ways that are not fully reflected in wages – through lower crime, improved political participation, etc. Moreover, the focus on a local geographical area prevents identification of externalities that arise if more skilled workers generate ideas used in other regions of the country.[4] Also, spill-over effects may (partly) accrue to employers instead of workers. In this light, macro-level analysis is a better tool to reveal these wider effects of such investments on economic growth (Sianesi and Van Reenen, 2002). However, compared with the micro analysis the macro approach faces many more methodological problems in interpreting the coefficient on education. These include measurement of human capital over time and across countries, causality issues, mostly *ad hoc* nature of model specifications and high sensitivity of estimates to the choice of additional regressors (Sianesi and Van Reenen, 2002). Krueger and Lindahl (2001) when pointing out the fragility of the macroeconomic evidence that is based on cross-country studies suggest that a focus on growth across regions of countries with reliable data is more promising.

social impacts, which can in turn have indirect economic effects, for example, lower crime, reduced welfare dependence and enhanced political behaviour.

[4] It may be argued however, that the level of cities and not regions or countries is particularly relevant for identification of human capital spillovers. First and foremost, the externalities may be one of the reasons underlying the formation of cities; second, urban areas represent a natural economic unit of analysis in contrast to administrative regions or states which are often arbitrary defined (see, e.g., Duranton, 2004).

This paper tests for the existence of human capital externalities – more precisely those stemming from higher education – using the abovementioned micro approach and data from a transition country (Russia). From a research perspective, the focus on a transition country may help circumvent a number of methodological problems that complicate identification of human capital externalities in the empirical micro-studies. In particular, we exploit the idea that the transition economies offer a unique natural experiment where market forces are imposed on the environment shaped by the central planning mechanism. The basic assumption underlying this paper is that the average educational attainment in the Russian cities at the end of the Soviet time was exogenous with respect to wages prevailing after the start of transition.[5] From a policy viewpoint, knowing the size of the externalities may be useful in determining the optimal level of public support of education.

This paper is organised as follows. Section 2 focuses on the theoretical foundations and design of the micro-studies and outlines several identification problems typical of such studies. Section 3 summarises the existing empirical evidence. Section 4 explains the plausibility and potential gains from implementing a micro-level study using data from a transition country. Estimation framework and data are described in section 5 followed by empirical analysis in section 6. Section 7 concludes.

2. Identification and measurement of human capital externalities in micro-level studies

2.1. Basic framework

The simplest framework that establishes the relationship between individual's earnings and the average stock of human capital in a relevant geographical area comes from a model developed by Lucas (1988). In this model, externalities are built into the aggregate production function in the form of technological increasing returns while the exact mechanism that generates externalities remains uncovered. This model is sketched below to provide a baseline for subsequent discussion. It is built on the following key assumptions: there exists a competitive economy, production takes place in several regions (cities), output is produced by identical agents (which differ only with respect to their human capital) and is traded on the national market rather than locally.

[5] A similar empirical strategy to study human capital externalities is employed by Jurajda (2004) who relies on exogeneity of the pre-transition distribution of the average education or exogeneity of the his-

Let c index cities, i – workers and L_c be the number of workers living in a city c. In the simplest model without capital, let y_{ic} be the output (and also the marginal product) of a worker i living in city c. Assume that it depends on the stock of human capital the worker possesses h_{ic} and on the city productivity shifter A_c so that

$$y_{ic} = w_{ic} = A_c h_{ic}, \tag{1}$$

where w_{ic} denotes earnings of worker i. The aggregate production function for a city can be written the following way:

$$Y_c = A_c \sum_{i=1}^{L_c} h_{ic} . \tag{2}$$

To formalise the idea that interactions among workers raise their productivity, A_c is assumed to depend on the aggregate human capital in a relevant city. As in the Lucas (1988) model, the aggregate human capital can be measured as the average human capital in the city, H_c=E_c(h_{ic}), so that

$$A_c = B_c H_c^{\alpha}, \tag{3}$$

where B_c measures a city-specific effect and human capital externalities are captured in elasticity parameter α.[6] Worker wage can therefore be written as

$$w_{ic} = A_c h_{ic} = B_c H_c^{\alpha} h_{ic} . \tag{4}$$

By taking logarithms, one transforms the above expression into

$$\ln w_{ic} = \ln B_c + \alpha \ln H_c + \ln h_{ic} . \tag{5}$$

The last equation provides rationale for using a standard Mincerian regression augmented with the average stock of human capital in a relevant city in order to identify human capital externalities.

More recent models draw the existence of externalities from the process of market interactions (e.g., Acemoglu, 1996). In the Acemoglu model an increase in the average education of the workforce raises equilibrium wages due to complementarity between human and physical capital even in the absence of technological or learning externalities. Importantly, the earnings equation resulting from this model is very similar to one obtained from the Lucas model.

---

torical location of colleges across the 74 districts of the Czech Republic.

[6] Another possible assumption is that the skills of the most talented individuals create externalities (Murphy, Shleifer and Vishny, 1991).

2.2. Extensions

An ideal framework for identifying human capital externalities would be a random assignment of different overall levels of human capital across cities, finding identical individuals in the cities and measuring difference in their wages before any sorting occurs. Yet, such experimental framework is not available. The literature on human capital externalities offers several extensions of the basic model that highlight identification problems accompanying empirical analysis. They primarily focus on unobserved heterogeneity of individuals (e.g., innate ability), selective migration and imperfect substitution among workers with different educational attainments (see, e.g., Acemoglu and Angrist, 2000; Moretti, 2004a; Ciccone and Peri, 2002). These issues are briefly reviewed below.

Heterogeneous individuals. Let workers be heterogeneous in terms of their unobserved ability with higher ability causing higher earnings. In such case, as argued in the empirical literature on human capital externalities (e.g., Acemoglu and Angrist, 2000; Silva, 2002), estimation may be based on the following model:

$$\ln w_{ic} = \alpha + \ln B_c + \beta_A A_c + \beta_S S_c + \gamma_a a_{ic} + \gamma_h s_{ic} + u_{ic}\,, \qquad (6)$$

where B_c is a city fixed effect, A_c is the average ability in city c, S_c is the average schooling of people living in city c, a_{ic} is ability and s_{ic} – schooling of individual i living in city c.

Ability at the individual and city levels is not observed. This poses no problem in empirical analysis as long as ability does not affect wages or is uncorrelated with schooling and other explanatory variables. If this is not the case, OLS estimate of the parameter of interest, β_S , is generally inconsistent.

Selective migration. Identification problems arising from selective migration are discussed in Acemoglu and Angrist (2000), Moretti (2004a) and others. For example, Acemoglu and Angrist (2000) use a simple model to show that positive city-specific shocks to wages attract more educated workers to the respective cities and increase average human capital through migration. This generates positive correlation between average education and wages across cities and may bias OLS estimates.

Similarly, Moretti (2004a) argues that unobserved characteristics of cities may be correlated with the share of people with higher educational attainment. Cities with particularly high productivity of skilled workers pay higher wages and therefore attract more skilled workforce. In this case the causal relationship runs from high wages to the average level of education of the labour force, rather than the other way around.

Imperfect substitutability of workers with different level of education. If workers with different educational attainment are imperfect substitutes in production (for which there is ample evidence, e.g., Katz and Murphy, 1992) wage changes may capture the complementarity between skilled and unskilled workers. In particular, under imperfect substitutability an increase in the share of educated workers may raise wages of unskilled workers due to the supply effect even in the absence of any externality. In contrast, wage of skilled workers will tend to go down. Hence, the existence of externalities is firmly established only if an increase in average education is related to an increase of wages of *more educated* workers. Based on this, Moretti (2004a) estimates external returns to education separately for each educational group and compares results for high and low education individuals. This approach, while providing evidence on the existence of externalities, hardly says anything about their magnitude, as emphasised in Ciccone and Peri (2002).

3. Empirical evidence from micro-level studies

The first study that attempted testing for and measuring of human capital externalities using the Mincerian approach is a paper by Rauch (1993). Based on a cross-sectional analysis of 1980 data from the US, this seminal paper finds that one year increase in average schooling leads to 3-5% increase in wages. The average level of education in this study is treated as historically predetermined, which evokes criticism in subsequent analyses.

Using panel data from 1960-1980 Censuses in the US, Acemoglu and Angrist (2000) reported that each additional year of average schooling in a state raised individual wages by 7% (OLS); however, IV estimation – which was intended to circumvent a bunch of the identification problems outlined above, in particular endogeneity of the individual and average educational attainments – resulted in coefficients that were small and insignificantly different from zero.

Rudd (2000) tests for the existence of human capital spillovers in the US on the state level with a panel dataset. Using OLS estimation and controlling for state fixed effects he finds that the observed correlation between the state educational attainment and individual earnings stems from the fact that the average level of education proxies for other, truly productive factors. In other words, the study finds no support for the hypothesis that human capital spillovers affect individual earnings.

Moretti (2004a) notes that OLS estimates show a large positive relationship between the share of college graduates in US cities and individual wages of the respective cities' residents. He then attempts to control for unobservable individual characteristics and unobservable city-specific shocks that may raise wages and attract people with higher educational attainment to different cities. He finds that a one percentage point increase in the labour force share of college graduates increases the wages of high-school dropouts and college graduates by 1.9% and 1.6% respectively while wages of college graduates raise by 0.4%. The result that an increase in the supply of college graduates raises their wages is consistent with the presence of human capital externalities.

To account for imperfect substitutability of workers with different human capital Ciccone and Peri (2002) propose a constant-composition approach to estimating human capital spillovers which, in contrast to the standard approach based on the Mincerian model, does not require estimation of the return to schooling at the individual level. Based on city level data from the US, the abovementioned study finds no evidence of positive human capital spillovers.

In this strand of literature Jurajda (2004) is the only (to the best knowledge of the author) study of human capital spillovers that exploits the "natural experiment" feature of the transition process in the former communist economies. The paper is built on the assumption of exogeneity of the historical location of colleges (which is typical in the literature) or, given the nature of the central planning system, exogeneity of the pre-transition distribution of human capital across the districts of the Czech Republic. The study finds no evidence of increasing returns from local concentration of human capital. This result, however, may be driven by the fact that the economic units of analysis in the paper are districts which do not necessarily coincide with local labour markets. Given a relatively small size of these administrative units in the Czech Republic, commuting may represent a particular problem for identification of human capital spillovers if the district of work and district of residence differ for a substantial number of workers (this problem is acknowledged by the author of the paper). In this light, focusing on cities in a bigger transition country may be a better approach.

Overall, the available evidence is contradictory and inconclusive. OLS estimates typically show a positive and significant impact of average education on individual wages. However, when IV estimation is used to circumvent various identification problems, the coefficient on average education remains positive but statistically insignificant. This raises the question about the quality of instruments: weak instruments result in the inflation of standard errors and insignificant coefficients. Indeed, many of the instruments used in the previous analysis – variations in compulsory schooling laws across the states, the presence of land-grant colleges and the demographic structure of cities are likely to be weak. To summarise, endogeneity issues and the quality of the instruments remain the major issues in the empirical analysis of human capital externalities on the micro-level.

4. Identification of human capital spillovers using Russian data

In order to identify human capital externalities, this paper exploits a natural experiment provided by the process of economic transition in Russia. We argue that the educational structure of cities under the central planning was determined by the government rather than the market; in particular, skill-biased migration was virtually non-existent as wages were equalized across regions through the so-called wage grid: an engineer in a highly skilled city received virtually the same wage as his colleague in a low-skilled city. As a result, the average educational attainment in cities at the end of communism can be regarded as exogenous with respect to market wages that prevailed after the start of transition.

Indeed, the specifics of the wage setting mechanism in the former USSR play a crucial role in justifying our approach. The Soviet economy was characterized by huge job vacancies with no open unemployment. Earnings of workers and salaried employees were determined according to the wage grid which primarily took into account position's respective skill level and the responsibilities it required (Geisheckerb and Haisken-DeNew, 2002). For ideological reasons mainly, the grid implied extremely small wage differentials and low returns to education; in fact, it resulted in the most egalitarian distribution of income in the world (Munich et al., 2002). The ultimate effect of the grid was that wages had little to do with actual productivity of workers, including potential productivity gains from local concentration of human capital. This effectively precluded any sorting based on productivity and wage differences in the Soviet time.

The differences in the price levels across regions were not conducive to selective migration either. In the USSR, full employment and wage compression were coupled with only slightly varying prices, rents and infrastructure costs across the regions and implied a rather uniform standard of living in different parts of the country. When wages and costs of living are centralized and do not vary much, individual preferences with respect to the place of residence depend on location-specific amenities which include climate, urban conditions and environmental quality. In general, the differences in these pure consumption amenities could induce migration, even a skill-biased one, if for some reason more able workers value specific consumption amenities more than their less able counterparts. We maintain, however, that such population movement was substantially reduced, if not eliminated, by the Soviet government through the system of migration controls that included different mechanisms at different periods of the USSR's existence. In particular, until the late 1950s the government denied internal passports for citizens from rural areas thus effectively preventing them from moving anywhere, most importantly to the cities. Later and until the very collapse of the USSR, seventy seven cities, mostly large ones, were subject to immigration restrictions (hence, a notion of "restricted cities" in the literature – see, e.g., Gang and Stuart, 1999). To redistribute labour from surplus to deficit regions, the government widely used a system of organized recruitment for labour ("orgnabor"). Non-market mechanisms such as those mentioned above were in the later years supplemented with attempts to emulate market forces (regional wage differentials, housing subsidies, paid moving expenses, etc.) in order to reallocate labour to the areas which were considered by the communist government as particularly important (in various senses, not necessarily economically).

The argument for treating the average educational attainment in cities as exogenous can be made more explicit by referring to the Soviet education system and administrative allocation of college graduates in particular. In the USSR, the government regulated the number of specialists produced by the educational institutions in accordance with the accounted needs of the planned economy and determined not only the overall composition of the workforce with respect to education, but also the distribution across the regions. For example, administrative allocation of labour was applied to graduates of higher education and secondary special educational institutions: after completion of their studies, the graduates received their first allocation, often in another region, where they had to work for three years. Apparently, the idea behind was that the graduates would settle in the destination regions for good (Clarke, 1999). This allowed the state to easily maintain a "target" skill level in the cities.[7]

The validity of this description of the centrally planned economy can be illustrated using regional level data from Russia.[8] In what follows we examine changes in the regional educational composition as well as their relationship with population change/migration before and after the start of transition. The analysis is based on data from the USSR censuses held in January 1979 and January 1989; the Russia's micro-census of February 1994 that embraced 5% of the population and the last census held in October 2002.[9] Table 1 provides descriptive statistics for the regional college share between 1979 and 2002. The mean college share increased steadily over the period, from 6.7% in 1979 to 14.2% in 2002.[10] Though the range has been increasing over time, the coefficient of variation has been falling which can be interpreted as a sign of regional convergence. Another important result is that the regional educational attainment seems to have depended on the presence of regional colleges only in the post-Soviet period. Before that, the presence of colleges appears to have played little role in determining the regional college share.

Table 1. College share in Russian regions, census data

Statistics	1979	1989	1994	2002
mean*	0.067 (88)	0.100 (88)	0.119 (88)	0.142 (89)
*- regions with own college(s)**	0.067 (76)	0.101 (76)	0.120 (76)	0.145 (81)
*- regions without own college(s)**	0.067 (12)	0.099 (12)	0.112 (12)	0.111 (8)
median	0.062	0.093	0.114	0.140
sd	0.024	0.029	0.033	0.033
cv	0.362	0.287	0.277	0.235
min	0.026	0.043	0.054	0.060
max	0.208	0.264	0.299	0.312
range	0.182	0.221	0.245	0.252

* Figures in brackets show the corresponding number of regions.

[7] Some authors argue that this system was not very stringent since it did not prevent people from finding another job or migrating into another area (see, e.g., Clarke, 1999). Nevertheless, even if people with certain qualification were leaving a particular area, the "central planner" could easily substitute similar individuals for them.
[8] Data limitations prevent such analysis on the city level.
[9] Hereafter all statistics on education refer to adult population, i.e. those of 15 years and older.
[10] Note that the means are not weighted by size (population) of the regions. This explains the discrepancies between the reported figures and the average college shares in the whole country according to the censuses (7.6% in 1979, 11.3% in 1989, 13.3% in 1994 and 16% in 2002).

Finally, the correlation coefficients between the regional college shares in adjacent periods are very high – 0.99 (1979 and 1989), 0.94 (1989 and 1994) and 0.86 (1994 and 2002). This indicates a high (though declining since after the start of transition) persistency of the regional educational attainment.

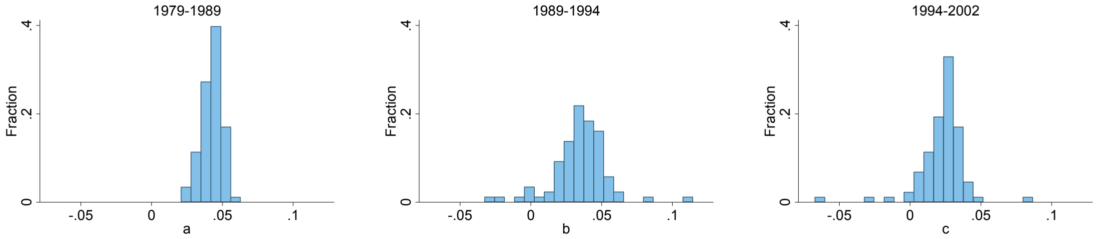
Table 2 provides summary statistics for the annualized rate of growth of the college share across regions in the three periods spanned by the censuses. As follows from the table, the last decade of the USSR's existence was characterized by a fairly homogenous growth of the college share across regions: similar growth rates are observed both in the regions that did not have own higher education establishments (12 such regions out of 88 in 1979) and in those that had at least one college/university; the figures are 4.1% and 4.3% respectively. After the start of transition there is a remarkable divergence among the college share growth rates with some regions experiencing a negative change in the college share. A substantially higher variance of the growth rates shows up in the coefficient of variation which jumps from 0.17 in 1979-1989 to 0.56 in 1989-1994 and 0.72 in 1994-2002. This widening of the distribution is illustrated in Figure 1. Also, according to Table 2, the presence of colleges/universities in regions has become an important determinant of the college share growth rate after the start of transition: between 1989 and 1994, the growth rate was 3.6% in the regions with at least one college (at the beginning of the period) and 2.4% in the regions without a college; in 1994-2002 the figures were 2.4% and 1.5% respectively. In contrast, the presence of colleges in the regions had little effect on the college share growth rate in the Soviet period.

Table 2. Annualized rate of growth of the regional college share

Statistics	1979-1989	1989-1994	1994-2002
mean	0.042 (88)	0.035 (87)	0.023 (88)
*- regions with own college(s)**	0.043 (76)	0.036 (75)	0.024 (76)
*- regions without own college(s)**	0.041 (12)	0.024 (12)	0.015 (12)
median	0.043	0.036	0.025
sd	0.007	0.019	0.016
cv	0.167	0.557	0.717
min	0.021	-0.032	-0.067
max	0.062	0.113	0.082
range	0.041	0.145	0.149

* Figures in brackets show the corresponding number of regions.

Figure 1. Distribution of the college share growth rates across regions in different periods.



The above analysis has established two important facts: 1) the variation in the college share growth rates has increased substantially in the post-Soviet period; 2) the presence of colleges in the regions, which was irrelevant for the regional college share as well as its growth in the Soviet time, has become an important determinant of both variables since after the start of transition. It appears that the Soviet government was able to maintain a target composition of skills in the regions regardless of the local production of education. Overall, these data may be interpreted as evidence of government controls over allocation of resources, including human capital, in the pre-transition period and absence of such controls thereafter.

Another piece of relevant evidence comes from the analysis of migration in the pre- and post-transition periods. The existing studies of migration in Russia document lower inter-regional migration rates since after the start of transition compared with other countries and with the Soviet period (see e.g., Heleniak, 1997; Andrienko and Guriev, 2002; Hill, 2004). One could probably argue that these low rates imply that the regional skill (educational) composition underwent little change since 1989 when it was determined by the central planner and therefore the contemporaneous average education in cities can be treated as an exogenous variable. However, the argument is probably flawed as what really matters is not the overall, but selective, or skill-biased, migration. Therefore, it is important to see how the regional college share growth rate was related to the regional migration rate. A simple regression analysis below helps shed some light on this issue. Table 3 shows the relationship between the college share growth rate and population growth (migration) rate.[11] To avoid potential criticism that the results are

[11] Migration rate over the period is defined as the average of the annual net migration rates (the ratio of migration growth to the mid-year resident population by current estimate) reported by the Federal State Statistic Service. The population growth rate is calculated using census population numbers and represents the annualized rate over the period between two adjacent censuses.

driven by outliers (clearly visible in Figure 1 for the college share growth rate), the relationship is derived from the quantile (median) regression.[12] It appears that since after the start of transition the college share growth rate was positively related to the population growth and migration rates while between 1979 and 1989 the relationship was the opposite (and weaker).[13]

Table 3. College share growth rate and population growth: quantile regressions

	College share growth rate				
	1979-89	1989-94	1994-02	1989-94	1994-02
Intercept	0.043**	0.035**	0.028**	0.037**	0.024**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Population growth rate	-0.121**	0.338**	0.461**		
	(0.034)	(0.084)	(0.066)		
Migration rate				0.499**	0.715**
				(0.090)	(0.089)
Pseudo R2	0.030	0.039	0.185	0.073	0.244
Number of obs.	88	87	88	87	88

Standard errors are reported in parenthesis. Asterisks indicate significance levels: 1% (**) and 5% (*).

Table 4 contains results of similar analysis but with additional controls for the initial college share (it is arguably easier to increase college share in a relatively low-skilled region and more difficult in a high-skilled one) as well as for the presence of at least one college in the region. These regressions show that between the last censuses held in the USSR, changes in the regional share of college graduates had little to do with population growth (migration). In contrast, the collapse of the centralized planning has lead to a positive association between the college share growth and immigration in the Russian regions. This is evidence of skilled-biased migration in the post-Soviet period and absence of it under the central planning. The analysis also provides evidence (though rather weak) that the presence of regional colleges started to affect regional college share growth only in the 1999s: the coefficient on the college dummy is negligible

[12] The results change little with OLS estimation.

[13] The author does not have data on migration in the pre-transition period; however, this seems to be a minor issue as population growth rates provide a very good proxy. In fact, the correlation coefficient between the migration rate and population growth rate is 0.90 in 1989-1994 and 0.91 in 1994-2002 (the figures are based on 87 and 88 observations respectively, excluding the conflict-ridden Checheno-Ingush/Chechen Republic characterized by large refugee flows). Importantly, the data show that such a high correlation is not driven by outliers.

in 1979-89, positive in the other two periods but statistically significant at 10% level in 1989-94 only.

Table 4. College share growth rate and population growth: quantile regressions

	College share growth rate				
	1979-89	1989-94	1994-02	1989-94	1994-02
Intercept	0.057**	0.033**	0.047**	0.037**	0.041**
	(0.004)	(0.006	(0.003)	(0.005)	(0.004)
College share, beg. of period	-0.215**	-0.057	-0.180**	-0.072	-0.124**
	(0.041)	(0.047)	(0.025)	(0.037)	(0.025)
Population growth rate	0.010	0.261*	0.464**		
	(0.066)	(0.122)	(0.043)		
Migration rate				0.448**	0.624**
				(0.110)	(0.059)
College in the region	0.000	0.009	0.002	0.007	-0.002
	(0.003)	(0.005)	(0.003)	(0.004)	(0.003)
Pseudo R2	0.286	0.080	0.297	0.114	0.332
Number of obs:	88	87	88	87	88

Standard errors are reported in parenthesis. Asterisks indicate significance levels: 1% (**) and 5% (*).

The above stylized facts on wage determination, administrative allocation of labour and migration controls in the Soviet period coupled with the analysis of the regional level data in Russia can be summarized in the following way. The pre-transition educational structure of cities is a valid and strong instrument for the (supposedly endogenous) average education in cities after the collapse of communism. Moreover, it can be considered not only as a valid and strong instrument, but also as a predetermined measure of average human capital that can be used in place of the respective contemporaneous measure (given that the regional college share is highly persistent). Clearly, such identification strategies are not available in the established market economies. Finally, we note that the identification strategy based on the historical locations of higher education establishments (the respective variable has been extensively used in previous studies as instrument for the average level of education) remains highly valid in the transition context as these establishments were set up by the Soviet government and had little to with the demands for skilled labour in the emerging market economy of the 1990s. However, the quality of this instrument is inferior to that of the historical average education, as follows from the above analysis.

5. Data description

Empirical analysis is based on data from RLMS (Russia Longitudinal Monitoring Survey), one of the few representative surveys of the Russian population which embraces about 8000 adults living in 32 out of 89 administrative regions of the country[14]. The RLMS sampling sites include 39 cities of which 15 are administrative centres with population (as of 1994) varying between 230 and 8630 thousand inhabitants and other, generally smaller, cities with population in the range of 11 to 275 thousand people. We use data from the fifth round of the survey implemented in 1994, which is a compromise solution in view of the following trade-off. With data from earlier waves that were conducted at the very beginning of the transition process, the argument for treating the contemporaneous level of education in cities as exogenous becomes stronger. Similarly, if the contemporaneous average education is instrumented with the pre-transition average education, the instrument becomes stronger when earlier data are chosen. However, the assumption that wages reflect marginal productivity of workers is quite problematic in the early transition period: following liberalization of prices and wages, adjustment of wages was hardly an instant process. Finally, there are data constraints related to the design of RLMS. In particular, the first phase of the survey (rounds 1 to 4 conducted in 1992-1994) embraced 12 cities only (compared with 39 in the second phase), which are too few for identification of the coefficient on the average level of education (especially when additional city-level covariates are added). Note that the effective sample size for identification of the coefficient on average education with RLMS is 39 at best.

There are several problems with specific variables that measure wages and education. As regards wages, the earlier rounds of RLMS, including round 5, contain information on the amount of money actually received from the employer(s) during the calendar month preceding the interview rather than the monthly contractual wage. The latter is available starting with round 8 only (i.e., from 1998 on). While this is hardly a drawback in the studies that use data from established market economies (the two estimates of wages differ little), unavailability of data on contractual wages entails problems in the Russian case because of the wage arrears that were widespread in the economy in the 1990s. Therefore, we redefine monthly wages using a simple procedure that takes account of wage arrears as in Earle and Sabirianova (1998) – see definitions of

[14] This is true of the second phase of the survey, which has been implemented since 1994. In rounds 1-4 that were conducted in 1992-1994 the sampling sites (cities) are different from those in the later rounds.

variables in Appendix 1. Up to round 6, RLMS provides education data in categorical form only with no information on actual years of schooling. The same applies to actual experience which is reported starting from round 8 only. Therefore, we impute years of schooling based on the highest degree obtained (as in Konstantinova-Vernon, 2002) and use potential experience which is defined as age minus years of schooling minus seven.

The RLMS data were supplemented with 1989 census data on average education across cities that were obtained from the respective regional branches of the Federal State Statistic Service (city-level data on education are not publicly available). These are the most precise data since they reflect educational attainment (though self-reported) of all the inhabitants in each city.

Table 5 provides information on city size, share of adults with higher education in 1989 (census data) and 1994 (estimates from RLMS), hourly wages and the number of adult respondents in each city.[15] The table reveals substantial differences in the share of adults with higher education across the cities in 1989 with variation between 6 and 27%. The size of the city and the college share are highly correlated; yet in 1989 Moscow was not ranked first with respect to the latter characteristic (it was surpassed by a smaller city in Moscow region). Next, the table shows that the RLMS estimate of the college share in 1994 and the exact college share in 1989 are highly correlated; though a high discrepancy between the two is also apparent (Figure 2 illustrates the relationship). This, of course, may be related to the skill-biased migration between 1989 and 1994, of which we found some evidence in the previous section. A complementary (and quite plausible) explanation is low precision of the estimates based on the RLMS data (the survey is not representative on the regional and city levels; the number of surveyed adults in each city is several hundred at best and in some cases falls to dozens only). For example, the true (population) values for the 1994 college share is 0.299 for Moscow and 0.247 for St. Petersburg (the figures are taken from the 5% micro-census held in 1994) while the RLMS estimates are 0.312 and 0.344 respectively. Thus, instrumenting contemporaneous average education in 1994 (measured with error on the basis of the RLMS data) with 1989 education can also help to reduce the measurement error bias.[16]

[15] City names are not shown due to the RLMS confidentiality policy.

[16] Note that the measurement error attenuates the coefficient on the average education while selective migration results in an upward bias in the coefficient. The two biases work in different directions; it is therefore possible that they (partially) compensate each other.

Figure 2. City college shares in 1989 (census data) and 1994 (RLMS estimates)

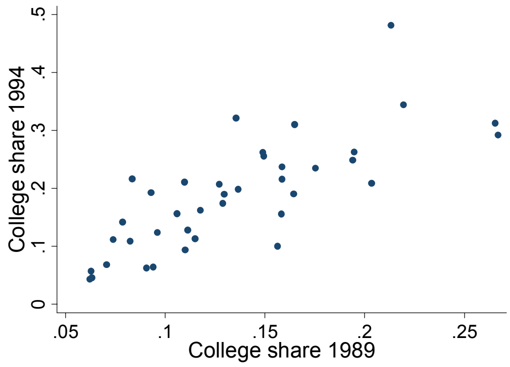
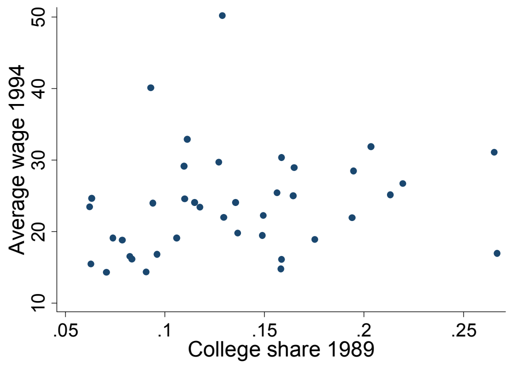


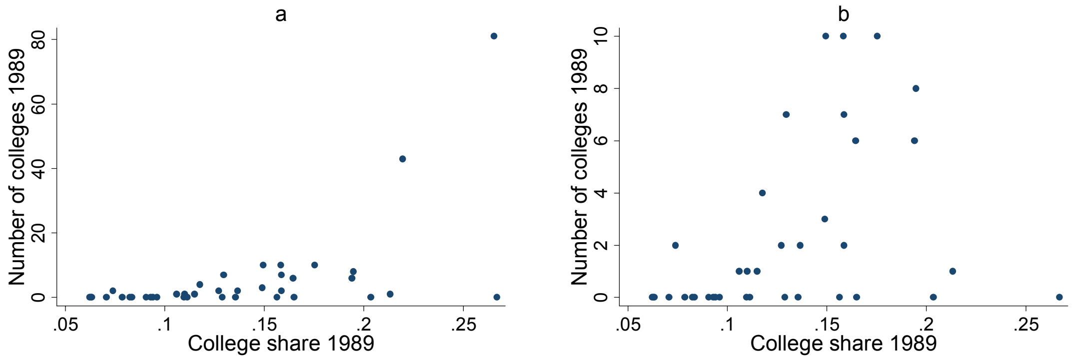
Table 5 also shows that the average hourly earnings differ a lot across the RLMS cities: the ratio of the maximum to minimum average wages in the sampled cities reaches 3.5. Not surprisingly, the highest earnings are found in the cities whose economies are centred upon the oil extraction industry (three top cities in the RLMS data). Figure 3 shows a positive correlation between the average hourly wage and college share in the RLMS cities. The correlation coefficient is 0.20 and not statistically significant; however, removing the three "oil cities" raises it to 0.41 (significant at 5% level).

Figure 3. City college share in 1989 (census data) and the average hourly wage in 1994.



The relationship between the college share and the number of higher education establishments in a city in 1989 is somewhat peculiar. Figure 4 show a positive correlation, which is particularly strong due to outliers – Moscow and St. Petersburg.

Figure 4. City college share and the number of institutes/universities in 1989: a) all RLMS cities b) excluding Moscow and St. Petersburg

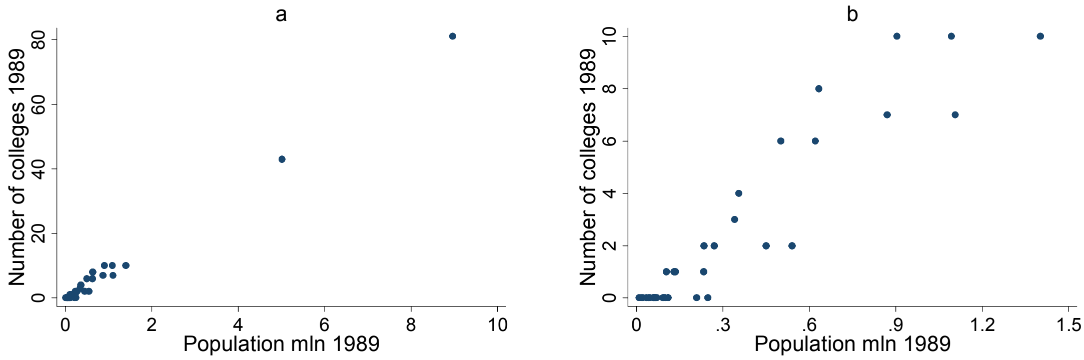


Once the metropolitan areas are excluded, the correlation becomes weaker (it falls from 0.55 to 0.4, but remains statistically significant at 1% level). For example, the city with the highest share of college graduates in 1989 with population slightly exceeding 0.1 million inhabitants did not host a higher education establishment in the Soviet time.[17] Again, this is a good example of how the centrally planned economy functioned: the government could achieve a target level of skill concentration in a city by using administrative allocation of resources (e.g., allocation of graduates) and not necessarily by establishing an institute or university in the city. Thus, the number of universities/colleges in a city in 1989 is a valid but not very strong instrument for the city college share, at least at the early years of transition, when the educational structure of cities was still pretty close to one created by the "central planner".

Also interesting is the relationship between city size and the number of higher education establishments in the cities in 1989. The correlation between the two is quite strong (0.99 for all cities and 0.94 if Moscow and St. Petersburg are excluded); moreover, the relationship between these variables virtually implies one university/institute per 110 thousand inhabitants which is likely to be a result of the specific policies of the central planner (see Figure 5).

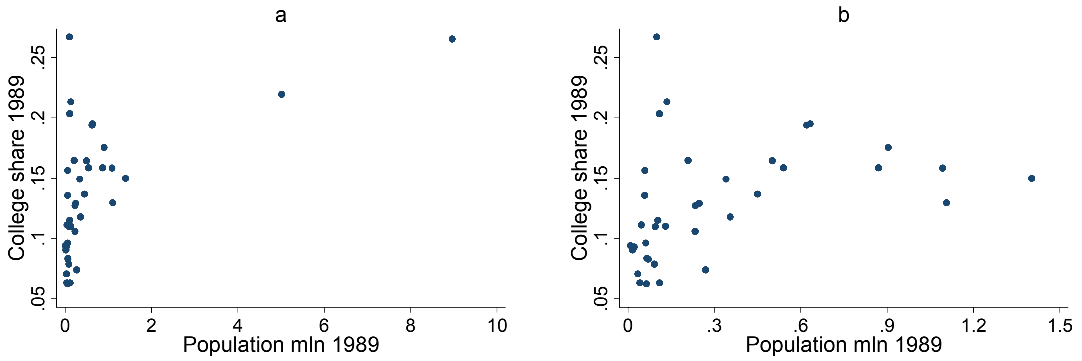
[17] High fraction of people with college degree in this city is due to the fact that it hosts several research institutes related to the defence industry.

Figure 5. City population and the number of institutes/universities in 1989: a) all RLMS cities b) excluding Moscow and St. Petersburg



Our final remark is on the relationship between city size and city college share – presumably highly correlated variables. City size may be though of as a natural control in the regressions as it is a proxy for agglomeration effects that may confound identification of the effect of average education on individual earnings. Figure 6 illustrates the relationship for the RLMS cities.

Figure 6. City population and college share in 1989 a) all RLMS cities b) excluding Moscow and St. Petersburg



As Figure 6 shows, there is a positive correlation between city population and college share which weakens if both Moscow and St. Petersburg are excluded (0.54 and 0.38; for log population the respective correlations are 0.62 and 0.48). This indicates the importance of controlling for city size; however, given the small number of cities (39), these high correlations may prevent disentangling the effects of college share on the one hand and city size on the other hand on individual wages.

6. Empirical analysis

We start with a simple OLS regression of log wages on individual characteristics (years of schooling, experience, experience squared and gender) and the RLMS-based estimate of the city college share in 1994. Regression results are reported in Table 6 (model 1 reported in the first column). The coefficient on the college share is 0.86 and statistically significant with the standard Huber/White/sandwich estimator of variance. The latter may not be appropriate due to correlation of the error terms corresponding to individuals living in same city. Therefore model 2 estimates the variance-covariance matrix under the assumption that observations are independent across cities, but not necessarily within cities (e.g., a standard cluster robust variance estimator is used).[18] The coefficient on the college share becomes statistically insignificant. These first results should be considered with caution due to potential endogeneity of the collsh variable (which would imply an upward bias in the coefficient) and the imprecise estimate of the college share based on the RLMS data, as discussed in the previous section (which would imply a downward bias).[19] Model 3 is an IV regression in which the 1994 share of people with higher education is instrumented with the respective share in 1989. The estimate of the coefficient of interests is 1.35 and significant at 5% level with cluster robust estimator of variance. Model 4 is identical to model 3 except for the instrument which is now the number of higher education establishments in the cities. The results are similar to those from model 3. Finally, in model 5 we use the share of adults with institute diploma in 1989 instead of the contemporaneous college share. The results are similar to those obtained by instrumenting for collsh: an increase in the share of people with higher education (in 1989!) by one percentage point raises average wages in the respective city (in 1994) by 1.67 percent. Note that interpretation of this coefficient should take into account the increase in the college share in the country from 11.3% in 1989 to 13.3% in 1994.

There are several weaknesses in this analysis. In general, there are two potentially endogenous regressors in the models: own education and average education and only the latter problem is addressed in models 3-5. Unfortunately, RLMS does not con-

[18] This is a somewhat problematic approach since estimation of the variance-covariance matrix relies on the asymptotic result with number of clusters going to infinity while we have only 39 cities (clusters).

tain any instrument for individual schooling; therefore the two admittedly endogenous variables have to be treated asymmetrically. Second, the above models ignore city heterogeneity (in particular, in terms of production amenities – natural resources, infrastructure, etc. that may have a direct impact on wages)[20]. This creates no problems as long as the average education in cities is uncorrelated with their production amenities, but this assumption may be questioned. Insrumenting for the contemporaneous average education with the pre-transition values or with the number of colleges/universities in the cities provides a solution to the problem; another way to check robustness of the results is by introducing various city-level control variables to the model. However, the best way to verify our results is perhaps by dropping observations from Moscow or Moscow and St. Petersburg taken together. The rationale is that these cities are obvious outliers in a number of important dimensions, e.g., size, educational attainment, number of higher education establishments, infrastructure quality, cultural amenities, etc.[21]

To check robustness of our results we use four important characteristics of cities: location (geographical region defined in the RLMS), status (whether it is an administrative centre of a region or not), the prevalence of the oil extraction industry in the city economy and city size measured by 1989 population.[22] Besides that, we re-estimate the equations using two sub-samples that exclude Moscow and Moscow/St. Petersburg respectively.

Estimation results are shown in Table 7. Model 1 is the same as model 5 in Table 6 and is replicated to serve as benchmark (thus, analysis that follows is based on the assumption of exogeneity of the city-level education in 1989). Model 2 introduces regional dummies. They are jointly statistically significant; however, they do not change the coefficient of interest. Model 3 introduces a dummy variable for the cities whose economies are centred upon the oil extraction industry; there are three such cities in RLMS: one in Komi Republic and the other two in Khanty-Mansiysk Autonomous Re-

---

[19] It is important to note that most of the literature is concerned about the upward bias due to endogeneity of the average education and less attention is paid to the potential downward bias which results when the average education is measured with error (especially when survey rather than census data are used).

[20] While some amenities can be measured, most are unobservable. City characteristics cannot be controlled for with dummy variables since the latter will eliminate the effect of average education and the coefficient of interest is unidentified.

[21] A similar approach is used by Jurajda (2004) who tests robustness of results by dropping data from Prague and Brno, the two largest cities in the Czech Republic.

[22] While many factors can be related to local wages, it is important to restrict the set of controls so that not to include possibly endogenous regressors in the model. For example, it is likely that such a natural control variable as cost of living in a city is an endogenous one (see, e.g., Gianetti, 2003).

gion. The dummy is highly significant; however, the coefficient on the college share in cities stays virtually unchanged. The same occurs when a dummy for city status is included – it turns out that there is no wage premium in regional capitals (see model 4). Model 5 contains a city size control (population in millions) – there is evidence that wages are higher in larger cities, but the effect of college share stays unchanged. Models 6 and 7 are estimated by dropping observations from Moscow and Moscow/St. Petersburg respectively. Again, the coefficient of interest stays nearly the same as before.

Next we consider imperfect substitutability across workers with different educational attainments by dividing the 1994 sample into two sub-samples: one containing individuals with college degree and the other one with people who obtained less schooling. Estimation results using these two sub-samples (Tables 8 and 9) show that the coefficient on average education is positive and significant in both sub-samples which is consistent with the existence of educational externalities. It is somewhat smaller in the more educated sub-sample; however, the differences between the estimates based on the two sub-samples are statistically insignificant.

As mentioned in section 5, one of the concerns related to the use of data from the early transition period is that wages may not properly reflect marginal productivity of workers. One possibility to check the validity of this concern is to replicate the above analysis using more recent data. Tables 10 and 11 show regression results based on the RLMS data from 2002 (the specifications are equivalent to those reported for 1994 in Tables 6 and 7). The relationship between individual wages and city share of people with higher education in 2002 is similar to that in 1994 and is consistent with the presence of human capital spillovers.

7. Conclusions

This paper centres upon the idea that the transition economies offer a unique natural experiment that makes it possible to shed some light on the controversial issue of human capital (educational) externalities. The basic assumption underlying this work is that the average educational attainment in Russian cities at the end of the Soviet time was exogenous with respect to the demand for skilled labour and city-specific productivity shocks in the emerging market economy of the 1990s. Such interpretation of the pre-transition average educational attainment solves the problem of endogeneity of average human capital which is encountered in the empirical micro-level studies based on data from the established market economies.

The estimation framework in this paper is the standard Mincerian wage regression augmented with the city college share variable. Our results are consistent with the presence of human capital externalities. In particular, we find that one percent increase in the share of people with higher education in cities results in the increase of earnings of the respective cities' residents by 1.5-2 percent. These results hold in both 1994 and 2002, i.e., early in the transition process and a decade after its start, are robust to the inclusion of several city-level controls and, more important, exclusion of observations from Moscow and St. Petersburg from the sample. Interestingly, the estimated magnitude of the externality is quite similar to one found by Moretti (2004) whose study is based on US data. It is also broadly consistent with macro-data based conjecture by Acemoglu and Angrist (2000) who suggest (by examining cross-country evidence) that the coefficient on average schooling measuring human capital externalities may be of the order of 25 to 30 percent on top of the 6-10 percent private returns.

Several qualifications should be made, however. The main caveat is that the results are based on a single cross-section and not a panel. This does not allow controlling for individual unobserved characteristics, such as ability, in the wage regression. Nor can we instrument for individual schooling due to the unavailability of relevant instruments in RLMS. Thus, the analysis proceeds under the assumption of exogeneity of individual schooling in the wage regression. Also, the cross-sectional structure of the data prevents introducing city fixed effects that would ideally capture city time-invariant characteristics that may affect both wages and average education. However, this problem has been addressed in the paper by applying a series of robustness checks.

Table 5. College share and wages in the RLMS cities

City No.	Population in 1989 (pop89)	College share in 1989 (collsh89)	College share in 1994 (collsh)	Wage in 1994 (Dec 2001 prices)	No. of obs. (adults)
1	5.020	0.219	0.344	26.7	369
2	8.967	0.265	0.312	31.1	590
3	0.235	0.127	0.207	29.7	203
4	0.047	0.111	0.128	32.9	164
5	0.341	0.149	0.262	19.4	229
6	0.070	0.083	0.108	16.5	157
7	0.540	0.159	0.237	16.1	225
8	1.403	0.150	0.256	22.2	227
9	0.042	0.063	0.057	15.4	141
10	0.450	0.137	0.198	19.8	232
11	0.034	0.071	0.068	14.3	133
12	1.094	0.158	0.156	14.8	244
13	0.905	0.175	0.234	18.9	209
14	0.066	0.083	0.216	16.2	134
15	0.092	0.079	0.142	18.8	240
16	0.621	0.194	0.249	21.9	189
17	0.063	0.096	0.124	16.8	98
18	1.107	0.130	0.189	22.0	227
19	0.356	0.118	0.162	23.4	204
20	0.104	0.115	0.113	24.0	169
21	0.271	0.074	0.111	19.1	217
22	0.110	0.063	0.046	24.6	175
23	0.502	0.164	0.190	25.0	180
24	0.248	0.129	0.174	50.1	186
25	0.233	0.106	0.156	19.1	205
26	0.870	0.159	0.216	30.4	213
27	0.634	0.195	0.262	28.5	227
28	0.064	0.062	0.043	23.5	140
29	0.022	0.093	0.192	40.1	26
30	0.009	0.094	0.064	24.0	47
31	0.136	0.213	0.481	25.1	27
32	0.101	0.267	0.292	17.0	24
33	0.059	0.156	0.100	25.4	20
34	0.058	0.135	0.321	24.1	28
35	0.130	0.110	0.094	24.5	36
36	0.109	0.203	0.208	31.9	24
37	0.209	0.165	0.310	28.9	29
38	0.095	0.110	0.211	29.2	19
39	0.016	0.091	0.063	14.3	16

Table 6. Estimation results: alternative specifications

Model :	1	2	3	4	5
Dep. var.: ln(wage)	OLS, rob. s.e.	OLS, clust. rob. s.e.	IV (collsh89)	IV (noinst)	OLS (college share 1989)
intcpt	1.968**	1.968**	1.914**	1.879**	1.906**
	(0.080)	(0.116)	(0.124)	(0.135)	(0.124)
collsh	0.859**	0.859	1.350*	1.672*	
	(0.164)	(0.498)	(0.558)	(0.685)	
schn	0.042**	0.042**	0.039**	0.036**	0.042**
	(0.006)	(0.006)	(0.006)	(0.008)	(0.006)
expp	0.022**	0.022**	0.022**	0.023**	0.022**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
expp2	-0.055**	-0.055**	-0.057**	-0.058**	-0.055**
	(0.008)	(0.009)	(0.009)	(0.008)	(0.009)
sex	0.299**	0.299**	0.299**	0.299**	0.30**
	(0.027)	(0.030)	(0.030)	(0.030)	(0.029)
collsh89					1.672**
					(0.551)
R-sq	0.104	0.104	0.101	0.096	0.111
# obs :	2801	2801	2801	2801	2801

Notes:

Standard errors are reported in parentheses with p<0.05 = *, p<0.01 = **; in all regressions except for (1) they are obtained using cluster robust estimator of variance. Standard errors in (1) are based on the Huber/White/sandwich heteroscedasticity-robust estimator of variance.

The instruments collsh89 and noinst are significant at 1% level in the first-stage regressions, with t-statistics 7.75 and 4.74 respectively.

With two instruments, noinst becomes insignificant in the first stage at 5% level and has wrong (negative) sign indicating that the variable just picks up noise. The Hansen J statistic for overidentification does not reject the null hypothesis that the instruments are independent of the second-stage disturbance term at usual significance levels (J= 0.448, p = 0.504). The estimated coefficient of interest in this case is 1.29 and significant at 5% level.

Table 7. Estimation results: robustness check for the basic specification

Model :	1	2	3	4	5	6	7
Dep. var.: ln(wage)	OLS	OLS	OLS	OLS	OLS	OLS	OLS
intcpt	1.906**	1.876**	1.839**	1.814**	1.687**	1.620**	1.645**
	(0.124)	(0.189)	(0.171)	(0.180)	(0.151)	(0.136)	(0.140)
collsh89	1.672**	1.780**	1.979**	2.163**	1.816*	1.782*	1.742*
	(0.551)	(0.621)	(0.512)	(0.724)	(0.823)	(0.826)	(0.833)
schn	0.042**	0.044**	0.045**	0.045**	0.045**	0.048**	0.046**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
expp	0.022**	0.020**	0.018**	0.018**	0.018**	0.020**	0.021**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
expp2	-0.055**	-0.051**	-0.047**	-0.047**	-0.046**	-0.050**	-0.053**
	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
sex	0.30**	0.308**	0.308**	0.308**	0.308**	0.30**	0.295**
	(0.029)	(0.026)	(0.026)	(0.026)	(0.026)	(0.028)	(0.029)
oil			0.626**	0.614**	0.618**	0.613**	0.60**
			(0.159)	(0.164)	(0.164)	(0.165)	(0.170)
captlr				-0.022	-0.017	-0.022	-0.057
				(0.067)	(0.068)	(0.069)	(0.093)
pop89					0.027*	0.034*	0.103
					(0.013)	(0.017)	(0.129)
Regg		Yes	Yes	Yes	Yes	Yes	Yes
		---	---	---	---	---	---
R-sq	0.111	0.19	0.218	0.218	0.219	0.221	0.224
# obs :	2801	2801	2801	2801	2801	2511	2348

Notes:
Standard errors are in parentheses with p<0.05 = *, p<0.01 = ** and obtained using cluster robust estimator of variance. Model 1 is the basic specification without additional controls.

Table 8. Estimation results: test for imperfect substitutability, sub-sample of workers without university degree

Model :	1	2	3	4	5	6	7
Dep. var.: ln(wage)	OLS	OLS	OLS	OLS	OLS	OLS	OLS
intcpt	2.056**	2.063**	2.067**	2.032**	1.839**	1.845**	1.839**
	(0.143)	(0.197	(0.172)	(0.196)	(0.156)	(0.160)	(0.167)
collsh89	1.658**	1.822**	1.983**	2.242*	1.814	1.800	1.741
	(0.612)	(0.684	(0.583)	(0.894)	(0.982)	(0.987)	(0.987)
schn	0.021	0.020	0.018*	0.018*	0.019*	0.020*	0.022*
	(0.012)	(0.010)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)
expp	0.025**	0.025**	0.022**	0.022**	0.022**	0.023**	0.023**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
expp2	-0.062**	-0.060**	-0.055**	-0.055**	-0.055**	-0.058**	-0.057**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)
sex	0.330**	0.332**	0.328**	0.328**	0.326**	0.312**	0.317**
	(0.038)	(0.034)	(0.034)	(0.034)	(0.034)	(0.035)	(0.036)
oil			0.622**	0.607**	0.612**	0.610**	0.597**
			(0.182)	(0.185)	(0.186)	(0.186)	(0.190)
captlr				-0.029	-0.024	-0.024	-0.061
				(0.080)	(0.079)	(0.081)	(0.105)
pop89					0.038**	0.037	0.109
					(0.014)	(0.020)	(0.126)
Regg		Yes	Yes	Yes	Yes	Yes	Yes
		---	---	---	---	---	---
R-sq	0.097	0.187	0.217	0.217	0.218	0.219	0.225
# obs :	2076	2076	2076	2076	2076	1892	1798

Notes:
Standard errors are in parentheses with p<0.05 = *, p<0.01 = ** and obtained using cluster robust estimator of variance. Model 1 is the basic specification without additional controls.

Table 9. Estimation results: test for imperfect substitutability, sub-sample of workers with university degree

Model :	1	2	3	4	5	6	7
Dep. var.: ln(wage)	OLS	OLS	OLS	OLS	OLS	OLS	OLS
intcpt	2.003*	2.192**	1.947*	1.925*	1.875*	1.193	1.945**
	(0.795)	(0.819)	(0.801)	(0.790)	(0.786)	(0.766)	(0.690)
collsh89	1.619**	1.177	1.489**	1.744**	1.533*	1.350	1.586*
	(0.543)	(0.665)	(0.568)	(0.570)	(0.730)	(0.741)	(0.737)
schn	0.045	0.041	0.051	0.051	0.052	0.09	0.036
	(0.053)	(0.053)	(0.052)	(0.052)	(0.053)	(0.055)	(0.047)
expp	0.021*	0.018	0.018	0.018	0.018	0.023*	0.031**
	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.008)
expp2	-0.062**	-0.054*	-0.054*	-0.054*	-0.054*	-0.061*	-0.078**
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.025)	(0.022)
sex	0.20**	0.218**	0.228**	0.227**	0.227**	0.224**	0.195**
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.049)	(0.048)
oil			0.629**	0.605**	0.607**	0.618**	0.579**
			(0.111)	(0.131)	(0.132)	(0.132)	(0.143)
captlr				-0.038	-0.035	-0.043	-0.088
				(0.061)	(0.064)	(0.064)	(0.090)
pop89					0.012	0.030*	0.102
					(0.010)	(0.013)	(0.161)
Regg		Yes	Yes	Yes	Yes	Yes	Yes
		---	---	---	---	---	---
R-sq	0.055	0.123	0.149	0.149	0.149	0.16	0.166
# obs :	725	725	725	725	725	619	550

Notes:
Standard errors are in parentheses with p<0.05 = *, p<0.01 = ** and obtained using cluster robust estimator of variance. Model 1 is the basic specification without additional controls.

Table 10. Estimation results: alternative specifications, 2002 data

Model : Dep. var.: ln(wage)	1 OLS, rob. s.e.	2 OLS, clust. rob. s.e.	3 IV (collsh89)	4 IV (noinst)	5 OLS (college share 1989)
intcpt	1.579**	1.579**	1.425**	1.280**	1.50**
	(0.075)	(0.114)	(0.133)	(0.169)	(0.113)
collsh	1.747**	1.747**	2.925**	4.034**	
	(0.146)	(0.622)	(0.715)	(0.687)	
schn	0.053**	0.053**	0.045**	0.037**	0.053**
	(0.005)	(0.007)	(0.007)	(0.007)	(0.006)
expp	0.017**	0.017**	0.019**	0.020**	0.018**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
expp2	-0.043**	-0.043**	-0.047**	-0.051**	-0.046**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)
sex	0.267**	0.267**	0.266**	0.265**	0.271**
	(0.023)	(0.032)	(0.031)	(0.029)	(0.031)
collsh89					2.980**
					(0.413)
R-sq	0.127	0.127	0.109	0.059	0.162
# obs :	3101	3101	3101	3101	3101

Notes:
Standard errors are reported in parentheses with $p<0.05$ = *, $p<0.01$ = **; in all regressions except for (1) they are obtained using cluster robust estimator of variance. Standard errors in (1) are based on the Huber/White/sandwich heteroscedasticity-robust estimator of variance.
The instruments collsh89 and noinst are significant at 1% level in the first-stage regressions, with t-statistics 6.12 and 4.89 respectively.
With two instruments, noinst has wrong (negative) sign in the first stage regression and is significant at 1% level. The Hansen J statistic for overidentification rejects the null hypothesis that the instruments are independent of the second-stage disturbance term at 10% significance level (J=2.765, p = 0.0963).

Table 11. Estimation results: robustness check for the basic specification, 2002 data

Model :	1	2	3	4	5	6	7
Dep. var.: ln(wage)	OLS	OLS	OLS	OLS	OLS	OLS	OLS
intcpt	1.500**	1.689**	1.643**	1.575**	1.586**	1.342**	1.298**
	(0.113)	(0.152)	(0.148)	(0.159)	(0.161)	(0.125)	(0.118)
collsh89	2.980**	2.298**	2.418**	2.883**	2.938**	2.849**	2.706**
	(0.413)	(0.481)	(0.446)	(0.570)	(0.701)	(0.664)	(0.557)
schn	0.053**	0.055**	0.056**	0.056**	0.056**	0.060**	0.065**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)
expp	0.018**	0.019**	0.018**	0.018**	0.018**	0.019**	0.019**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
expp2	-0.046**	-0.047**	-0.046**	-0.047**	-0.047**	-0.048**	-0.047**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)
sex	0.271**	0.280**	0.280**	0.280**	0.280**	0.294**	0.315**
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.034)	(0.029)
oil			0.802**	0.771**	0.770**	0.758**	0.713**
			(0.140)	(0.151)	(0.151)	(0.154)	(0.167)
captlr				-0.052	-0.053	-0.084	-0.193*
				(0.060)	(0.062)	(0.057)	(0.077)
pop89					-0.003	0.050**	0.249**
					(0.012)	(0.014)	(0.067)
Regg		Yes	Yes	Yes	Yes	Yes	Yes
		---	---	---	---	---	---
R-sq	0.162	0.202	0.229	0.229	0.229	0.203	0.198
# obs :	3101	3101	3101	3101	3101	2640	2421

Notes:
Standard errors are in parentheses with p<0.05 = *, p<0.01 = ** and obtained using cluster robust estimator of variance. Model 1 is the basic specification without additional controls.

8. References

Acemoglu, D. (1996) A Microfoundation for Social Increasing Returns in Human Capital Accumulation, *Quarterly Journal of Economics*, p. 779-804.

Acemoglu, D. and Angrist, J. (2000) *How Large are Human Capital Externalities? Evidence from Compulsory Schooling Laws*, unpublished manuscript, MIT.

Andrienko, Y., Guriev, S. (2002) *Determinants of Interregional Mobility in Russia: Evidence from Panel Data*, unpublished manuscript, CEFiR, Moscow.

Buckley, G. (1995) The Myth of Managed Migration: Migration Control and Market in the Soviet Period, *Slavic Review*, Vol. 54, No. 4, p. 896-916.

Card, D. (1999) The Casual Effect of Education on Earnings, in: Ashenfelter, O., Card, D. *Handbook of Labour Economics*, Vol. 3A, North Holland.

Ciccone, A., Peri, G. (2002) Identifying Human Capital Externalities: Theory with an Application to US Cities, *IZA Discussion Paper* No. 488.

Clarke, S. (1999) *The Formation of a Labour Market in Russia*, Edward Elgar, Cheltenham, UK.

Duranton, G. (2004) Human Capital Externalities in Cities: Identification and Policy Issues, forthcoming in: Arnott, R. and D. McMillen (eds.), *Blackwell Companion to Urban Economics.*

Earle, J. and K. Sabirianova (1998) Understanding Wage Arrears in Russia, Stockholm School of Economics, *SITE Working Paper* No. 139.

Gang, I., Stuart, R. (1995) Mobility where Mobility is Illegal: Internal Migration and City Growth in the Soviet Union, *Journal of Population Economics*, Vol. 12, pp. 117-134.

Geishecker, I. and Haisken-DeNew, J. P. (2002) Riding the Transition Roller-Coaster: Flexibility and the Inter-Industry Wage Structure in Russia, *German Institute for Economic Research Discussion Paper* No. 280.

Giannetti, M. (2003) On the Mechanics of Migration Decisions: Skill Complementarities and Endogenous Price Differentials, *Journal of Development Economics*, Vol. 71 No. 2 pp. 329–349.

Heleniak, T. (1997) Internal Migration in Russia During the Economic Transition, *Post-Soviet Geography and Economics*, Vol. 38, No. 2, pp. 81-104.

Hill, F. (2004) *Eurasia on the Move: The Regional Implications of Mass Labour Migration from Central Asia to Russia*, Brookings Institution.

Jurajda, S. (2004) *Are There Increasing Returns to Local Concentration of Skills? Evidence on Wages and Returns to Education in Transition,* CERGE-EI working paper.

Katz, L. F. and Murphy, K. M. (1992) Changes in Relative Wages, 1963-1987: Supply and Demand Factors, *Quarterly Journal of Economics*, Vol. 107, No. 1, pp. 35-78.

Konstantinova Vernon, V. (2002) *Returns to Human Capital in Transitional Russia*, Department of Economics, The University of Texas at Austin.

Krueger, A., Lindahl, M. (2001) Education for Growth: Why and for Whom? *Journal of Economic Literature*, Vol. XXXIX, pp. 1101-1136.

Lucas, R. (1988) On the Mechanics of Economic Development, *Journal of Monetary Economics*, Vol. 22, pp. 3-42.

Moretti, E. (2004a) Estimating the Social Returns to Higher Education: Evidence from Longitudinal and Repeated Cross-Section Data, *Journal of Econometrics* 121: 175-212.

Moretti, E. (2004b), Human Capital Externalities in Cities, in: Henderson, V. and J.-F. Thisse (eds.), *Handbook of Urban and Regional Economics*, Amsterdam: Elsevier-North Holland.

Munich, D., Svejnar, J., Terrel, K. (2002) *Returns to Human Capital under the Communist Wage Grid and during the Transition to a Market Economy*, CERGE-EI working paper.

Murphy, K., Shleifer, A., Vishny, R. (1991) The Allocation for Talent: The Implications for Growth, *Quarterly Journal of Economics*, Vol. 106, pp. 503-530.

Psacharopoulos, G., Patrinos, H. (2002) *Returns to Investment to Education: A Further Update*, World Bank Policy Research Working Paper No. 2881.

Rauch, J. (1993) Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities, *Journal of Urban Economics*, Vol. 34, pp. 380-400.

Rudd, J. (2001) Empirical Evidence on Human Capital Spillovers, unpublished manuscript, Federal Reserve Board.

Sianesi, B., Van Reenen, J. (2003) *The Returns to Education: A Review of the Empirical Macro-Economic Literature*, Journal of Economic Surveys, Vol. 17, No.2, pp. 157-200.

Silva, J. (2002) *Human Capital Externalities or Sorting? Evidence from a Displaced Workers Sample*, unpublished manuscript, IUE.

Spagat, M. (2002) Human Capital, Growth and Inequality in Transition Economies, *CEPR Discussion Paper* No. 3556.

Venniker, R. (2000) *Social Returns to Education: A Survey of Recent Literature on Human Capital Externalities*, CPB report 00/1.

Appendix 1: Definition of variables

Individual characteristics:

wage – hourly earnings from primary and secondary jobs in the reference month. The variable refers to money actually received rather than to the contractual wage. A correction for wage arrears is applied (if a person received no wage in the reference month, the wage is approximated by the value of arrears divided by the length of the period over which these arrears had been accumulated. Outliers corresponding to 2,5% of observations in the lower and upper tails of the wage distribution are excluded from the dataset. Earnings data are reported in Rubles in December 2001 prices (1 US Dollar ≈ 30 Russian Rubles).

schl – individual educational attainment (years of schooling). The variable accounts for years of schooling and is imputed on the basis of highest degree obtained, in particular, 9 years for incomplete secondary education, 11 years for ordinary secondary, 12 years for vocational, 13 years for specialized secondary, 16 years for college, 19 years for a graduate degree. The imputation is the same as in Konstantinova-Vernon (2002).

expp – working experience (potential, calculated as age minus schooling minus seven).

expp2 – working experience squared.

sex – gender, 1 refers to males and 0 to females.

City characteristics:

collsh – city college share in 1994/2002. Calculated for adults (15 years and older) from RLMS.

collsh89 – city college share in 1989. Census data from 1989, adult population (15 years and older).

noinst – number of higher education establishments in a city at the end of the Soviet time.

oil – dummy variable for the cities whose economies are centred around the oil extraction industry.

captlr – dummy variable for cities which are administrative centres of the regions.

Regg1 – Regg8 – dummy variables for regions (RLMS classification: Moscow and St. Petersburg, Northern and North Western, Central and Central Black-Earth, Volga-Vyatka and Volga Basin, North Caucasian, Ural, Western Siberian, Eastern Siberian and Far Eastern).

pop89 – city size, million inhabitants, 1989 census data.