NATIONAL RESEARCH UNIVERSITY

Institutions and the Allocation of Talent: Evidence from Russian Regions

Michael Alexeev (Indiana University)
Timur Natkhov (Higher School of Economics)
Leonid Polishchuk (Higher School of Economics)

Stylized Facts

- Institutions affect economic outcomes (growth, welfare etc.) via the allocation of resources between (directly) productive activities and rentp-seeking
- Private payoff to education (educational wage premium) is observed consistently across the world, but public payoff is elusive ("micro-macro paradox")
- North-Pritchett's "chemical engineering vs. piracy": human capital can be deployed for socially unproductive purposes

Murphy et al., 1991

- Murphy, Shleifer ad Vishny (1991) suggested that the allocation of talent between productive and non-productive purposes serves as a mediator between institutions and economic outcomes
- They proposed to proxy the deployment of talent to productive purposes by obtaining education in sciences (STEM) and engineering disciplines, and the deployment of talent to rent-seeking by pursuing law degrees
- They hypothesized greater sensitivity of the allocation of top talents to institutional quality, and greater significance of such allocation for economic growth
- They observed a negative cross-country correlation between graduation in law and growth rates, but never tested the rest of their hypotheses

Natkhov and Polishchuk 2018

- … have shown that graduation in sciences is strongly positively correlated with institutional quality, whereas for graduation in law an even stronger negative correlation is observed ("law is more popular in lawless countries")
- Such correlations are remarkably robust to data models, estimation techniques, measures of institutional quality, sub-samples of nations etc.
- Allocation of talent solves the micro-macro paradox: in a sub-sample of countries with higher difference between graduation in law and sciences higher educational attainments increase growth rates , whereas in the rest f the sample such correlation is absent

Limitations of Cross-Country Analysis

- Omitted variable bias
- Uneven occupational and educational standards and admission and graduation rules across the world
- Inability to assess the impact of institutions on the allocation of talent, lack of individual data

Advantages of Russian Data

We treat Russian regions as jurisdictional units and make use of profound variations of institutional environments between Russian regions, which are still parts of a single economy and polity.

Interregional institutional diversity in Russia is an outcome of largely exogenous variations of historical, geographic etc. nature

We use a unique data set of enrollment over the 2011-2014 period of nearly all of Russian freshmen students pursuing post-secondary degrees (a total of about 1,300,000 individuals), specifying the chosen field of study, university (region), and Unified State Examination (USE) score, serving as an ability measure

The Model (stylized description)

Individual characteristics: ability (effort multiplier) and idiosyncratic preferences for particular activity

$$u_i(y;\alpha) \equiv u(y,i;\alpha), \qquad i = 1,2$$

Involvement in re-distribution (as opposed to productive efforts) includes offensive and defensive (on behalf of value-creating agents) activities. In equilibrium, both types of activities earn to re-distributors the same rate of return, which is the payoff to redistribution:

$$w = \frac{\Theta(1-\sigma)(1-f\big(x^*(w,\sigma)\big))}{1-\Theta-\Theta x^*(w,\sigma)}$$

Payoff to production:

$$d(w,\sigma) \equiv \sigma + (1-\sigma)f\big(x^*(w,\sigma)\big) - wx^*(w,\sigma)$$

Main Theoretical Results

(i) Allocation of human resources to productive activities increases in institutional quality (property rights protection) (*almost* obvious …)

(ii) Higher (but not necessarily top) talents exhibit greater elasticity in their occupational choices to the quality of institutions: marginal return to intuitional quality increases when talent rises from average too higher level

(iii) Inter-jurisdictional mobility weakens the impact of local institutions on the allocation of talent

Data

- USE scores and “major” for almost all matriculating students from Russia’s regions (about 1.3 million observations) [Major at enrollment determines major at graduation]
- Institutional quality measures for regions (informal employment share, investment climate index, and FOM (2011))
- Other regional characteristics (structure of economy, PC GRP, population, January temperature, mobility)

Years: 2011-2014

Aggregate vs. individual data

We have both aggregate (by region) and individual-level data on the choice of discipline

Aggregate data are comparable with what has been used in the literature, but it is difficult to get at the effect of USE on the choice of discipline; all we can do is look at the entire sample vs. top 25% and top 10%

The results for aggregate data are significant and consistent with our theory but only for between-effects estimation. Fixed-effects results are mostly statistically insignificant

Hence our focus on individual-level data

Aggregate Data

WB estimator (time fixed effects; errors clustered by region)

	$STEM - LAW$	$\frac{STEM - LAW}{STEM + LAW}$	$STEM - LAW$	$\frac{STEM - LAW}{STEM + LAW}$	$STEM - LAW$	$\frac{STEM - LAW}{STEM + LAW}$
	Entire sample	Entire sample	Top 25%	Top 25%	Top 10%	Top 10%
	(1)	(2)	(3)	(4)	(5)	(6)
Inverse invest. risk index (mean)	0.547* (0.301)	1.431* (0.778)	0.769** (0.325)	2.111** (0.914)	0.842*** (0.280)	2.752*** (1.001)
Inverse invest. risk index (dev.)	0.196 (0.185)	0.043 (0.426)	-0.006 (0.234)	0.262 (0.678)	-0.220 (0.272)	0.274 (0.913)
No. of obs.	309	309	308	308	308	308
No. of regions	78	78	78	78	78	78
R-squared (overall)	.167	.186	.209	.190	.235	.202

Regression Models for Individual Data

Dependent variables (AoT_i): dummy variables for the choice of field of study: STEM (science, technology, engineering and mathematics), law (law and public administration), and health.

Main specification:

$$\boldsymbol{AoT_i = \beta_o + \beta_1 USE_i + \beta_2 IQ_j + \beta_3 USE_i \times IQ_j + \gamma X_{tj} + \varepsilon_{ti}}$$

where USE_i is proportion individual USE score, IQ_j is a measure of institutional quality of region j, and X_{tj} is a vector of regional controls, including regional and year fixed effects. Errors are clustered by region

Specifications

In addition, we run regressions with $regions \times year$ fixed effects, although in these regressions we cannot calculate marginal effects of regional quality, because part of it is subsumed in these fixed effects

We run both LPM and Probit regressions

Probit does not allow for regional fixed effects due to incidental parameters problem

Regressions with $region \times year$ fixed effects

	Institutional quality: inverse of informal employment share		
Dependent variable:	$STEM$	LAW	$STEM\_LAW$
USE_i score	-0.014*** (0.004)	-0.004** (0.001)	-0.020*** (0.005)
$USE_i \times IQ_j$	0.015*** (0.005)	-0.004** (0.002)	0.022*** (0.006)
R-squared	0.037	0.017	0.062
Observations	1296900	1296900	554822
Dependent variable:	$HEALTH$	$LAW + HEALTH$	$STEM\_LAW\_HEALTH$
USE_i score	0.020*** (0.004)	-0.024*** (0.004)	-0.033*** (0.005)
$USE_i \times IQ_j$	-0.018*** (0.005)	-0.022*** (0.005)	0.029*** (0.006)
R-squared	0.097	0.075	0.137
Observations	1297000	1297000	671626
Number of regions	77	77	77

Regressions with $region \times year$ fixed effects

	Institutional quality: inverse of investment risk index		
Dependent variable:	$STEM$	LAW	$STEM\_LAW$
USE_i score	-0.015*** (0.003)	-0.004** (0.002)	-0.020*** (0.005)
$USE_i \times IQ_j$	0.016*** (0.004)	-0.004* (0.003)	0.024*** (0.008)
R-squared	0.037	0.017	0.062
Observations	1294019	1294019	554018
Dependent variable:	$HEALTH$	$LAW + HEALTH$	$STEM\_LAW\_HEALTH$
USE_i score	0.016*** (0.006)	-0.020*** (0.006)	-0.031*** (0.006)
$USE_i \times IQ_j$	-0.013* (0.008)	-0.017** (0.008)	0.027*** (0.009)
R-squared	0.095	0.074	0.136
Observations	1294119	1294119	670478
Number of regions	77	77	77

Individual data regressions (LPM; informal employment)

	Fixed effects OLS					
Dependent variable:	$STEM$	LAW	$STEM\_LAW$	$HEALTH$	LAW_HEALTH	$STEM\_LAW\_HEALTH$
	(1)	(2)	(3)	(4)	(5)	(6)
USE_i score	-0.013***	0.004***	-0.019***	0.019***	0.022***	-0.031***
	(0.003)	(0.001)	(0.004)	(0.004)	(0.004)	(0.005)
IQ_j	-0.559*	0.223**	-1.057***	0.885***	1.108***	-1.353***
	(0.333)	(0.100)	(0.289)	(0.277)	(0.287)	(0.395)
$USE_i \times IQ_j$	0.013***	-0.004**	0.021***	-0.016***	-0.020***	0.027***
	(0.005)	(0.002)	(0.006)	(0.005)	(0.005)	(0.006)
Marginal effect of	0.368**	-0.070	0.380**	-0.238**	-0.307***	0.513***
IQ_j at $USE_i = 70$	(0.115)	(0.061)	(0.164)	(0.094)	(0.107)	(0.160)
Marginal effect of	0.501***	-0.112#	0.586***	-0.398***	-0.509***	0.779***
IQ_j at $USE_i = 80$	(0.134)	(0.071)	(0.208)	(0.127)	(0.144)	(0.194)
Marginal effect of	0.633***	-0.153*	0.791***	-0.558***	-0.711***	1.046***
IQ_j at $USE_i = 90$	(0.167)	(0.084)	(0.256)	(0.166)	(0.188)	(0.239)
R-squared (within)	.007	.002	.068	0.013	0.051	0.079
No. of obs.	1,296,900	1,296,900	554,822	1,297,000	1,297,000	671,626
No. of regions	77	77	77	77	77	77

Individual data regressions (LPM; investment risk index)

	Fixed effects OLS					
Dependent variable:	$STEM$	LAW	$STEM\_LAW$	$HEALTH$	LAW_HEALTH	$STEM\_LAW\_HEALTH$
	(1)	(2)	(3)	(4)	(5)	(6)
USE_i score	-0.014***	0.004**	-0.020***	0.015***	0.019***	-0.029***
	(0.003)	(0.002)	(0.005)	(0.005)	(0.006)	(0.006)
IQ_j	-0.746**	0.237	-1.278***	0.556	0.793	-1.212**
	(0.309)	(0.155)	(0.454)	(0.424)	(0.477)	(0.561)
$USE_i \times IQ_j$	0.015***	-0.005*	0.023***	-0.011	-0.016**	0.025***
	(0.004)	(0.002)	(0.007)	(0.007)	(0.008)	(0.008)
Marginal effect of	0.292#	-0.081	0.352*	-0.248*	-0.329**	0.564**
IQ_j at $USE_i = 70$	(0.181)	(0.068)	(0.196)	(0.128)	(0.157)	(0.247)
Marginal effect of	0.441**	-0.127#	0.585**	-0.363**	-0.490**	0.817***
IQ_j at $USE_i = 80$	(0.197)	(0.079)	(0.238)	(0.180)	(0.210)	(0.283)
Marginal effect of	0.590***	-0.172*	0.818***	-0.478**	-0.651**	1.071***
IQ_j at $USE_i = 90$	(0.220)	(0.095)	(0.293)	(0.242)	(0.274)	(0.337)
R-squared (within)	.007	.002	.013	0.067	0.050	0.078
No. of obs.	1,294,019	1,294,019	554,018	1,294,119	1,294,119	670.478
No. of regions	77	77	77	77	77	77

Individual data regression (Probit; investment risk index)

Dependent variable:	$STEM$	LAW	$STEM\_LAW$	$HEALTH$	LAW_HEALTH	$STEM\_LAW\_HEALTH$
	(1)	(2)	(3)	(4)	(5)	(6)
USE_i score	-0.043***	0.025***	-0.062***	0.073***	0.068***	-0.088***
	(0.009)	(0.009)	(0.017)	(0.026)	(0.018)	(0.018)
IQ_j	-2.236**	1.497	-3.713**	2.263	2.868	-4.018**
	(0.926)	(0.957)	(1.702)	(2.628)	(1.813)	(1.939)
$USE_i \times IQ_j$	0.047***	-0.028**	0.071***	-0.044	-0.054**	0.077***
	(0.013)	(0.013)	(0.024)	(0.035)	(0.025)	(0.025)
Marginal effect of IQ_j	0.370**	-0.064*	0.339**	-0.177*	-0.267***	0.519**
at $USE_i = 70$	(0.156)	(0.037)	(0.140)	(0.103)	(0.102)	(0.211)
Marginal effect of IQ_j	0.515***	-0.110***	0.570***	-0.388**	-0.506***	0.833***
at $USE_i = 80$	(0.165)	(0.042)	(0.173)	(0.154)	(0.138)	(0.226)
Marginal effect of IQ_j	0.644***	-0.160***	0.826***	-0.636**	-0.763***	1.064***
at $USE_i = 90$	(0.181)	(0.054)	(0.227)	(0.0.265)	(0.211)	(0.258)
Pseudo R-squared	0.009	0.009	0.022	0.112	0.058	0.071
Number of observations	1,294,019	1,294,019	554,018	1,294,119	1,294,119	670,478

Individual data, (Probit; informal employment)

Dependent variable:	*STEM*	*LAW*	*STEM_LAW*	*HEALTH*	*LAW_HEALTH*	*STEM_LAW_HEALTH*
	(1)	(2)	(3)	(4)	(5)	(6)
USE_i score	-0.041***	0.021***	-0.059***	0.090***	0.075***	-0.088***
	(0.011)	(0.008)	(0.015)	(0.023)	(0.015)	(0.017)
IQ_j	-2.056**	1.344**	-3.472***	3.937*	3.598***	-4.171***
	(0.858)	(0.605)	(1.078)	(2.047)	(1.238)	(1.377)
$USE_i \times IQ_j$	0.042***	-0.023**	0.064***	-0.065**	-0.061***	0.074***
	(0.015)	(0.010)	(0.020)	(0.029)	(0.020)	(0.023)
Marginal effect of IQ_j	0.320**	-0.035	0.259*	-0.129#	-0.195*	0.391**
at $USE_i = 70$	(0.127)	(0.040)	(0.140)	(0.061)	(0.104)	(0.198)
Marginal effect of IQ_j	0.452***	-0.070	0.461**	-0.381**	-0.442***	0.690***
at $USE_i = 80$	(0.160)	(0.052)	(0.193)	(0.161)	(0.165)	(0.263)
Marginal effect of IQ_j	0.569***	-0.110#	0.683***	-0.702**	-0.717***	0.920***
at $USE_i = 90$	(0.197)	(0.067)	(0.254)	(0.0.274)	(0.239)	(0.323)
Pseudo R-squared	0.009	0.009	0.021	0.113	0.058	0.070
Number of observations	1,296,900	1,296,900	554,018	1,297,000	1,297,000	671,626

LPM regressions for FOM measure of corruption (2011; individual-level data)

	LPM (OLS)					
Dependent variable:	$STEM$	LAW	$STEM\_LAW$	$HEALTH$	LAW_HEALTH	$STEM\_LAW\_HEALTH$
	(1)	(2)	(3)	(4)	(5)	(6)
USE_i score	-.0003 (.0018)	-.0001 (.0005)	.0012 (.0015)	.002# (.001)	.002# (.001)	-.003# (.002)
IQ_j	.006 (.006)	-.003# (.002)	.013** (.005)	-.015*** (.004)	-.018*** (.004)	.025*** (.006)
$USE_i \times IQ_j$	-.0002* (.0001)	.0001* (.0000)	-.0002*** (.0001)	.0003*** (.0001)	.0003*** (.0001)	-.0005*** (.0001)
Marginal effect of IQ_j at $USE_i = 70$	-.005** (.002)	.0013# (.0009)	-.006** (.003)	.003* (.002)	.004** (.002)	-.007** (.003)
Marginal effect of IQ_j at $USE_i = 80$	-.007** (.003)	.0019* (.0011)	-.009*** (.003)	.005*** (.002)	.007*** (.002)	-.011*** (.004)
Marginal effect of IQ_j at $USE_i = 90$	-.008** (.004)	.0025* (.0014)	-.011*** (.004)	.008*** (.002)	.010*** (.003)	-.016*** (.005)
R-squared	.018	.007	.031	.067	.063	.105
Number of obs.	274,585	274,585	118,075	274,585	274,585	142,019
Number of regions	70	70	70	70	70	70

Probit regressions for FOM measure of corruption (2011; individual-level data)

	Probit					
Dependent variable:	$STEM$	LAW	$STEM\_LAW$	$HEALTH$	LAW_HEALTH	$STEM\_LAW\_HEALTH$
	(1)	(2)	(3)	(4)	(5)	(6)
USE_i score	.0002 (.0049)	.003 (.004)	-.0005 (.0068)	.016*** (.006)	.013*** (.004)	-.014** (.007)
IQ_j	.019 (.015)	-.013 (.015)	.037# (.023)	-.089*** (.023)	-.062*** (.016)	.067*** (.023)
$USE_i \times IQ_j$	-.0005** (.0002)	.0003 (.0002)	-.0008** (.0004)	.0013*** (.0003)	.001*** (.0003)	-.001*** (.0004)
Marginal effect of IQ_j at $USE_i = 70$	-.005** (.002)	.001 (.001)	-.005* (.003)	.001 (.002)	.003 (.002)	-.006# (.003)
Marginal effect of IQ_j at $USE_i = 80$	-.007** (.003)	.0016 (.0012)	-.007** (.004)	.005* (.003)	.007** (.003)	-.011** (.005)
Marginal effect of IQ_j at $USE_i = 90$	-.008** (.003)	.0023 (.0017)	-.011** (.005)	.011*** (.004)	.011*** (.004)	-.014*** (.006)
Pseudo R-sq	.014	.015	.037	.108	.072	.089
Number of obs.	274,585	274,585	118,075	274,585	274,585	142,019
Number of regions	70	70	70	70	70	70

Accounting for migration; marginal effects (institutional quality: investment risk, 2014, LPM)

Share of graduates staying in region: / USE score	65	75	85
70	0.352	0.534**	0.716#
	(0.349)	(0.261)	(0.452)
80	0.418	0.693**	0.968*
	(0.427)	(0.328)	(0.584)
90	0.483	0.852**	1.220*
	(0.514)	(0.403)	(0.728)

Accounting for migration; marginal effects (institutional quality: investment risk, 2014, Probit)

Share of graduates staying in region: / USE score	65	75	85
70	0.374	0.528**	0.699#
	(0.306)	(0.242)	(0.432)
80	0.436*	0.690**	0.980*
	(0.376)	(0.307)	(0.579)
90	0.497	0.860**	1.279*
	(0.455)	(0.382)	(0.741)

Accounting for migration; marginal effects (institutional quality: informal employment, 2014, LPM)

Share of graduates staying in region: / USE score	65	75	85
70	0.301*	0.447**	0.593*
	(0.179)	(0.207)	(0.351)
80	0.374*	0.617**	0.859*
	(0.221)	(0.257)	(0.446)
90	0.448*	0.786**	1.125**
	(0.270)	(0.311)	(0.544)

Accounting for migration; marginal effects (institutional quality: informal employment, 2014, Probit)

Share of graduates staying in region: / USE score	65	75	85
70	0.290*	0.411**	0.548
	(0.176)	(0.201)	(0.386)
80	0.356#	0.592**	0.864*
	(0.219)	(0.257)	(0.518)
90	0.423#	0.783**	1.198*
	(0.270)	(0.319)	(0.649)

Placebo tests

- We ran regressions similar for those for STEM and Law&Public Admin and Health for all other disciplines with more than 100,000 matriculants.
- Disciplines: Agricultural Studies, Economics and Management, Education, Humanities and Social Sciences.
- None of these disciplines exhibits more or less consistent statistically significant marginal effects of institutional quality

Conclusions

(i) Higher quality institutions and policies are essential for making proper use of factors of production, including investments in human capital

(ii) Higher ability individuals are more sensitive to the quality of institutions, at least within certain range of ability

(iii) The possibility of migration reduces responsiveness of the allocation of talent to institutional quality

(iv) Russia's apparent comparative advantage in terms of quality of human capital would not be useful for diversifying the economy and generating innovation-based growth without improving institutions