Preliminary and incomplete

How Does the Expansion of Higher Education Change the Returns to College Quality? Insights from Sixty Years of Russian History

April 20, 2013

Olga Belskaya
University of North Carolina-Chapel Hill
belskaya@live.unc.edu

Klara Sabirianova Peter
University of North Carolina-Chapel Hill
IZA and CEPR
kpeter@unc.edu

Abstract

In the ten years between 1997 and 2008 the number of college students in Russia surged from 3 to 7.5 million, the number of universities has more than doubled, and the college enrollment rate increased by more than 30 percentage points. We study the effect of this remarkable expansion of higher education on changes in college quality and labor market returns to college quality. We use a unique linked dataset that combines four different sources: a 17-year nationally representative panel of individuals; the university database; official regional and national statistics from 1950 to 2011; and the archives of Soviet and Russian laws. The majority of previous studies on college quality rely on time-constant quality measures for a given cohort of students in a cohort-based panel such as NLSY. Our study makes a step forward compared to the existing literature by constructing a number of time-varying proxies for college quality over a 60-year period and identifying the cross-cohort shifts in returns to college quality. After controlling for age-time effects and constant individual heterogeneity, we find evidence of declining college quality and increasing returns to college quality for the cohorts that obtained education during the period of expansion. We also model the selection into college using the exogenous changes in the rules of military conscription of students after WWII, changes in compulsory schooling laws, and unexpected shifts in the wage gap between manual and non-manual workers. The paper shows an increase in the heterogeneity of the returns to college quality and flattening the "earnings-quality" profile.

Keywords: college expansion, returns to college quality, Russia

Acknowledgement
We are thankful to Saraswata Chaudhuri, Robert Gonzalez, Lutz Hendriks, and Christian Posso for useful comments.

1. Introduction

The last forty years were the years of the worldwide expansion of post-secondary education. The number of students pursuing tertiary education worldwide has skyrocketed from 28.6 million in 1970 to 152.5 million in 2007 (UNESCO, 2009). While the tertiary education is often deemed to increase individual earnings and labor productivity, its rapid growth does not come without the costs which are frequently overlooked. Unless additional resources are simultaneously allocated to the educational system, the quality of education deteriorates, as the number of students rises (Bound *et al*., 2010). The higher number of students increases the average class size, puts a downward pressure on the number of faculty per student, limits spaces in the labs and libraries, and, in principle, may adversely affect student learning. The expansion of higher education also changes the ability composition of college graduates by bringing new entrants from a lower part of the distribution of abilities. As argued in Carneiro and Lee (2009) and Carneiro *et al*. (2011), expansions in college enrollment at the margin attract students with lower ability and lower returns to schooling, thus affecting the within-group inequality and the average returns to schooling.

In this study, we highlight another important channel of how rising college enrollment may influence the wage distribution of college graduates – namely, via changes in the returns to college quality. We show that labor markets can respond to an expansion-related decline in average college quality by penalizing graduates from low-quality programs, by awarding graduates from high-quality programs, and by increasing the overall price heterogeneity. We also show that the expansion of higher education is likely to increase the observed returns to college quality by improving the assortative matching between the college quality and student abilities. Both price heterogeneity and assortative matching effects are found to be positively associated with college expansion.

To study the labor market outcomes of college expansion, we consider the case of Russia. Having moved away from centrally controlled to market-oriented, the Russian higher education expanded tremendously over a short period of time. Public universities grew and over 400 new private colleges emerged in response to growing demand and liberalization reforms. In just ten years between 1998 and 2008, the number of students has more than doubled (Goskomstat, 2012) and Russia emerged among countries with the highest share of young adults with tertiary education (OECD, 2012). Although Russia's tertiary enrollment is not the fastest growing in the world (this title would belong to South Korea), Figure 1 shows that within a very short time Russia has narrowed the gap with the U.S. and pulled ahead of most European countries in terms of population enrollment in tertiary education. Furthermore, Russia's high completion rates (almost 80 percent compared to 45 percent in the U.S.) translate the enrollment growth into much higher supply of college educated workers (OECD 2010).

Given these tremendous changes, we examine how the expansion of higher education in Russia affected the distribution of college quality and the distribution of the returns to college quality. The large body of economic literature analyzes the relationship between the college quality and individual earnings; see Zhang (2005) for a review. Many of these studies show that there are indeed significant gains associated with going to a better college (Hoekstra 2009; Loury and Garman 1995; Monks 2000 among others). However, previous studies have not yet linked the dynamic changes in the returns to college quality to rising college enrollment.

We contribute to the literature by providing the first estimate of the effect of college expansion on the quality and wages of college graduates. We find an evidence of a declining quality of college education and increasing returns to college quality for the cohorts that obtained education during the period of expansion. Our findings suggest that the returns to high-quality colleges have grown over time while the returns to low-quality colleges diminished. To our knowledge, no existing study has been able to utilize the data covering multiple cohorts of individuals, the quality of their education, and labor market outcomes simultaneously. To the extent that these individuals were exposed to dramatically different conditions when in college, our results provide a picture that unifies changes over a long period of time and a large number of people who were affected.

Our paper fits in the current heated debate in Russia on the restructuring of ineffective colleges and their branches. The current Russian government is clearly favoring more administrative measures to ensure the quality of educational institutions and to prevent further decline in college quality, with the government measures ranging from direct monitoring to the shutting down of ineffective low-performing universities. Our study shows that the market forces in Russia are strong enough to counter-balance the negative consequences of college expansion, and that the labor market can correct the over-supply of graduates from lower-quality programs by lowering their wage premium.

Our study focuses on higher professional education (HPE) institutions as part of tertiary education that offers 4-6 year programs after completion of general secondary education.[1] Russian HPE institutions are referred to by a variety of names, including “Universities”, “Institutes”, “Academies”, and “Higher Schools”. For compatibility of terminology with international literature on this topic, this paper uses the terms “Universities” and “Colleges” to denote HPE institutions, even though the latter term is not technically accurate from the Russian language point of view. Thus, such terms as “College”, “University”, and “HPE institution” are used interchangeably throughout this paper.

The paper is organized as follows. Section 2 provides background on the variation in college quality during the Soviet era and describes key sources behind the recent expansion of higher education in Russia. Section 3 introduces data sources. Section 4 discusses the

[1] The UNESCO definition of tertiary education also includes Russian institutions of secondary professional education (SPE) that train technicians and other associate professionals such as nurses, bank tellers and elementary school teachers. SPE institutions are called *colleges* or *tehnikums* and offer 4-year programs after 9 years of secondary schools or 2-3-year programs after 11 years of complete general secondary education.

measurement of college quality and shows shifts the distribution of college quality during the expansion period. Section 5 describes the empirical model, which predicts changes in the returns to college quality in response to college expansion. Section 6 performs decomposition of wage variance in quality vs. price effects. Section 7 summarizes our findings and discusses their implications.

2. Relevant Institutional Details

In this section, we provide an institutional context for understanding the expansion of higher education in Russia and related changes in college quality. Since some individuals in our sample earned their college degree in the 1950s, we also offer a historical perspective on the sources of variation in college quality in the Soviet era (1950-1991).[2] In particular, we highlight three main sources: demographic fluctuations, part-time programs, and resource allocation.

First, historic demographic fluctuations played an important role in the evolution of college quality post WWII. As Figures 2A and 2B show, the growth in the number of students has been steady through most of the Soviet era prior to 1991, but the enrollment rates (number of students divided by population 18-22) almost quadrupled during the first five years of the 1960s. The spike in enrollment rates in the 1960s was caused by the dwindling college age population resulting from low birth rates during the WWII period. The enrollment consequences of WWII can still be seen in the 1980s for the cohort of grandchildren of the WWII generation (Figures 2A and 2B). Since the college quality (i.e., mean ability of students and resources per student) tends to vary with enrollment rates, the cross-cohort differences in quality are likely to be substantial even prior to the most recent college expansion in Russia.

Second, widespread programs of correspondence and part-time evening education have also contributed to both between– and within–cohort differences in college quality in the Soviet

[2] The key features of the Soviet system of higher education are summarized in accompanied web appendix W1.

era. The correspondence education was the predecessor of modern distance learning programs and it offered a relatively inexpensive way of getting an education while working.[3] Russia claims to be the first country in the world to introduce a mass-scale correspondence education in the 1920s.[4] Several HPE institutions focused exclusively on correspondence education. At their peak during the "space race" in the 1960s, correspondence programs enrolled 44 percent of all students. Another 16 percent of students studied in evening programs, with total of 60 percent enrolled in part-time programs (as shown in Figure 3A). The quality of correspondence and evening programs was lower than that of day-time programs due to the lack of interactions between students and professors, limited peer effects, inadequate lab time, lower quality of students, and the allocation of left-over resources and "left-over" faculty to part-time programs by the university administration (de Witt, 1961; OECD, 1999).

Third, the differences in college quality existed due to the uneven allotment of resources across universities of different types by the Soviet government. More resources and better quality faculty were assigned to main campuses as compared to branches, to leading Moscow universities and regional centers of education, and to large polytechnic universities as opposed to smaller specialized institutes. It is interesting that the 71 prominent Russian universities honored by the Soviet State a few decades ago still continue to significantly outperform other universities and attract better students today.[5]

---

[3] A typical semester in the Soviet correspondence program consisted of 3 parts: (1) intensive start-up on-campus session where professors give an overview of the course concepts; (2) 2-3 months of self-education at home and sending assignments by correspondence; and (3) lab and testing session on-campus.

[4] In 1926, several leading universities opened correspondence programs that accepted 37,000 students; see http://encyclop.ru/26515.

[5] These universities were recognized for outstanding past services and high-quality training and awarded with special Soviet honors such as the Order of Lenin and the Order of the Red Banner of Labor; the list is published in Rosen (1980). According to the 2012 efficiency assessment by the Ministry of Education, the main campuses of the universities that were honored by the Soviet regime enroll students with 10 percent higher test scores, have almost triple the share of foreign students, spend more than double on research and development per faculty, and have significantly larger instructional space per student compared to other main campuses (authors' calculations).

The college attainment in post-WWII Russia has always been high by international standards for the level of Russia's economic development. However, the latest expansion in the 2000s is more than remarkable. After a short period of declining student population in the 1990s, the number of students in 1997 exceeded the 1980 level of 3 million and then surged to 7.5 million over the next ten years (Figure 2B). The number of universities has more than doubled between 1992 and 2008, and the college enrollment rate increased from 27 percent to 62 percent of college age population (Figure 2A). Around the year 1999, the labor markets started experiencing a huge influx of new college graduates. To date there are no systematic micro-level studies of the reasons for the massive college expansion in Russia. We can only conjecture that the increasing returns to skills and the liberalization of college entry were the two most likely forces behind the surge of college attainment. The transition to a market economy moved Russia away from the highly compressed wage setting that favored manual industrial workers. Figure 4A documents a sizeable drop in the wages of manual workers relative to the wages of non-manual workers during the transition to a market economy in the 1990s. Figure 4B tells a similar story by showing a rise in college wage premium in the early 1990s right before the college expansion.[6]

If the rising returns to college created incentives for individuals to obtain higher education, the liberalization of entry on the supply side expanded individual opportunities for receiving higher education. The 1992 Law on Education authorized the opening of the first private HPE institutions, and their number started growing rapidly (Federal Law, 1992). In 2008, as many as 474 new private HPE institutions enrolled 17 percent of all students (Figure 3B). The Constitution of the Russian Federation (1993) also legalized tuition charges at public universities for those students

[6] A rapid increase in college premium during the first transition years was likely instigated by the market adjustment of wages to reflect the true marginal products of the different types of labor. The systemic transformation and extreme volatility may have also generated an additional premium for the ability of highly-educated individuals to respond to changing opportunities in a disequilibrium situation, as conjectured by Schultz (1975) and empirically tested by Fleisher et al. (2005). We also do not exclude a potential role of demand shifters (such as skill-biased restructuring of privatized companies and foreign direct investment) in the rise of private returns to higher education in the 1990s.

who fell short of the competition for publicly-financed or so called budget slots. Although the Russian government has been regulating the size of tuition-charged enrollments at public institutions by imposing quotas for each subject major, the share of students paying tuition increased quickly – reaching 55 percent in 2010 (Figure 3A).

Russian universities responded to the growing demand for higher education and to the tight public budget constraint by lowering admission criteria and offering tuition-based programs to those applicants who would not have passed old admission standards (OECD, 1999). While facing a resource/infrastructure constraint at the headquarter campuses, university administrations began opening branches in remote locations and offering less resource-demanding part-time programs. The number of branches exploded since the late 1990s, peaking at 2201 branches in 2004 (Goskomstat, Regions 2006). At the same time correspondence and evening programs became a dominant form of higher education (Figure 3A). Universities also responded to public demand by changing the composition of majors subject to the government approval. The share of public HPE graduates in business-related fields increased from 13 percent in the 1980s to 37 percent in 2010. Practically all students at private HPE institutions (95 percent) specialize in the fields of business and humanities (see web appendix W1 for detail on majors).

We note that the absolute number of students in regular (publicly-financed full-time) programs experienced very little fluctuations from 1.5 to 2 million of students annually over the 1975-2010 period. However, their share dropped from 63 percent in 1992 to 26 percent in 2008 percent due to the inflow of admissions in private schools and tuition-paying students at public universities.

The federal government recognized that the expansion may come at the price of lower quality and undertook a number of measures to ensure quality. For example, since 1992 all private HPE institutions are required to obtain a government license to operate legally. Periodically, the Russian Ministry of Education and Science performs an assessment of the curricula, faculty, and material resources and then accredits HPE institutions and local branches if their educational programs meet the state educational benchmarks. In the late 2000s Russia instituted a standardized college entrance test similar to the SAT in the U.S. with the goal of implementing identical admission standards across HPE institutions. The test is called EGE, which stands for the unified state exam. The EGE replaced the old Soviet system of admission, which was based on 3-4 subject exams determined and administered by each university separately.

More recently, several political leaders called for the closing of the correspondence programs after a heavily publicized incident of 83 students-lawyers from St. Petersburg not passing the state qualifying exam in 2008. The two leading Russian universities, Moscow State and St. Petersburg State, terminated their correspondence programs in 2010 and justified their decision by joining the European "education space" and being recognized by the international community.[7] In the campaign for quality the federal government decreed an overhaul of higher education.[8] An external audit commissioned by the Ministry of Education and Science in 2012 identified 136 out of 502 public universities and 450 out of 930 public branches as ineffective in terms of the quality of students, teaching space, and research expenditures per faculty member (Monitoring of Higher Professional Education, 2012).[9] Based on the results of the audit, the government proposed to shut down almost 300 campuses of low-quality public universities or merge them with effective ones. In response to the proposed measures, students and faculty at several Russian universities in jeopardy went on strike.[10]

---

[7] See http://ria.ru/edu_news/20100209/208421810.html and http://www.memoid.ru/node/Polemika_ob_otkaze_vuzov_Rossii_ot_zaochnogo_obucheniya (in Russian). There is a significant resistance to this movement in the media and among lower ranked universities, where tuition correspondence programs provide a major source of financing.

[8] See The New York Times, March 25, 2012; The Moscow Times, August 31, 2012; Nature, December 17, 2012.

[9] Private HPE institutions and their branches participated in the audit voluntarily, but 85 percent of them declined to participate. 41 out of 70 participating private HPE institutions and 55 out of 97 private branches were found to be ineffective (Monitoring of Higher Professional Education, 2012). The result is not surprising given that many newly established private HPE institutions, without state funds and reputation, face challenges of hiring permanent faculty, investing into libraries and facilities, and attracting high-quality students. A compulsory audit of all private universities and their branches is scheduled for 2013.

[10] See http://en.ria.ru/russia/20121226/178417137.html and

Perhaps it is not coincidental that the expansion of higher education brought the issue of college quality to the center of heated national debate, which drew the attention of international media. While we do not take a stance in this debate, we hope that our research will shed light on the labor market response to the expansion of college education and associated changes in the returns to college quality.

3. Data

This study relies on the four distinct data sources. The primary micro data are drawn from RLMS, a multi-purpose annual household panel survey.[11] The number of surveyed individuals varied from 10,465 to 14,690 per year until 2009. In 2010-2011, the sample expanded to 21,300-22,000 respondents. Compared to similar panel surveys in other countries, RLMS has relatively high response rate[12] and low sample attrition due to compensation for the survey participation, lower mobility, and infrequent changes of residences. The RLMS sample covers 160+ cities and villages (municipalities or locations thereafter), 38 primary sample units (PSU thereafter), 32 state-sized federal subjects (regions thereafter), and all 7 federal districts of the Russian Federation. The sample represents the geographic and economic heterogeneity of Russia. For example, in terms of the regional domestic product per capita, the maximum-to-minimum ratio across RLMS regions was 11 in 2010.

The survey provides a variety of information on college graduates, including demographics, earnings, current employment status, whether they received a college degree, year of graduation, and how many years they studied.[13] These characteristics are available beginning with the 1995 survey. Subsequent rounds of RLMS added questions on college major (1998-present), the year of

---

http://www.universityworldnews.com/article.php?story=20130411163909607

[11] RLMS is organized by the National Research University Higher School of Economics, Moscow together with the Carolina Population Center at the University of North Carolina at Chapel Hill and the Institute of Sociology at the Russian Academy of Sciences. The panel started in 1994. The survey was not administered in 1997 and 1999.

[12] The response rate exceeds 80 percent for households and is about 97 percent for individuals within the households.

[13] The description of all the individual-level variables employed in this study is presented in Appendix A1.

college exit for dropouts (2006-present), tuition for current students (2006-2008), full-time program (2009-present), and name and location of attended university (2004-2005, 2010-2011). Because of the panel structure of the data, we were able to fill in college information in all other years (1995-2011) for respondents with at least one valid data point (see details of imputation procedure in Appendix A1). For some respondents, the college information cannot be retrieved due to panel attrition and also due to the inconsistencies in majors across years when the imputation was not performed.[14] All individual responses on college majors were sorted into 58 groups based on the Russian Classification of the Field of Study (see web appendix W2 on the coding of majors in RLMS).

Although RLMS respondents reside in 32 regions at the time of the survey, they graduated from universities located in 72 of 83 Russian regions and all 15 former USSR republics. We identified and confirmed via Internet sources 1073 Russian universities and branches, which were attended by RLMS respondents. Using university websites and the Federal Portal "Russian Education" (www.edu.ru), we created indicators for public vs. private ownership, main campus vs. local branch, university region and municipality, type of municipalities (Moscow city, other central city, non-central city/township), one of the 11 university types (e.g., classical university, polytechnic, pedagogy, etc.), and founding date (see Appendix A1 for definitions of university-level variables).[15] 799 of 1073 universities are publicly-owned and 360 are local branches.

The summary statistics presented in Table 1 for the RLMS sample of individuals age 20-60 is consistent with the established general population trends reported by official statistics. Table 1 shows that the share of college graduates increased from 19 percent in 1995 to 26 percent in 2010, that years of schooling have been gradually rising, that the share of females among college

[14] We lose about 6 percent of college graduates due to missing major and 13 percent due to missing university location in 1995-2011. We employ the inverse propensity weighting (IPW) to account for attrition and other sources of missing data.

[15] Whenever possible, we assigned these characteristics to foreign universities and to unidentified Russian universities (e.g., when the name is abbreviated or not clear, but location and ownership are stated by a respondent).

graduates grew substantially and reached 62 percent in 2010, and that the composition of college graduates shifted away from engineering and pedagogy fields towards college majors in economics, business, law, and humanities. Table 1 also confirms well-known facts about college graduates having much higher wages, more years of schooling, and greater likelihood of being employed compared to a randomly sampled individual. Between 1995 and 2010, we see a noticeable increase in the share of college graduates who earned their degree in local branches (from 9 to 15 percent), private universities (from 0.2 to 7.4 percent), and universities in non-central cities (from 11 to 16 percent).

Besides the university characteristics identified in the RLMS sample, we also use (i) university performance indicators for 1432 public universities and local branches from the 2012 Ministry of Education and Science assessment report (R&D expenditures per faculty, available instructional space per student, revenues per faculty, and the share of foreign students); and (ii) annual information on EGE test scores and tuition for enrolled freshmen in 2010-2012 collected by the Higher School of Economics Monitoring of Quality (HSE-EGE database thereafter).[16] The EGE test scores are averaged by university, college major, program of study, and year. To match the HSE-EGE data with our theoretical and empirical models, all observations are split into two categories: publicly-financed full-time programs (program A) and other programs including tuition-based and part-time programs (program B). For tuition-paid programs, we also use annual tuition by university, college major, and year in 2010 prices.

In total, from all sources including RLMS, we identified and classified 2387 former and current universities, which we refer to as the Russian University Database. All universities in the database are categorized in the same way as RLMS universities in terms of their ownership, location, university type, and founding date. Appendix A3 indicates a substantial amount of

[16] Using on-line enrollment reports, we manually added about 300 observations for the underrepresented categories of correspondence programs, branches, and military majors. The final HSE-EGE database includes 1014 main campuses and branches (506 public universities, 81 private universities, and 427 loca l branches).

heterogeneity across Russian universities. We observe much higher test scores and better performance results in main campuses, Moscow universities, publicly-financed programs, and older schools compared to branches, universities in non-central cities, tuition-financed programs, and newer schools.

Our third data source is the Russian Statistical Office (Goskomstat thereafter) and its USSR predecessor, which provide both regional and cohort-level variables going back to the year 1950.[17] Some examples of these variables include the number of students by program and sources of financing, population size by age, wages and employment by industry and worker skills, the number of faculty, etc. Definitions and sources for each variable are presented in Appendix A1. Regional quality measures and enrollment rates are linked to RLMS respondents using the region of attended university and the year when individuals reached 20 years old (medium college age). The national averages of quality measures are linked to the year subjects turned 20 (i.e., the 1955 values are used for a 60-year old college graduate surveyed in 1995).[18] Finally, we rely on the extensive archives of Soviet and Russian legal documents to trace changes in compulsory schooling laws, military conscription laws related to students, admission rules, and general education policies (e.g., www.libussr.ru; www.consultant.ru).

The ultimate linked dataset from the four different sources – a 17-year nationally representative panel of individuals, the university database, official regional and national statistics from 1950 to 2011, and the archives of Soviet and Russian laws – presents a unique opportunity to

[17] A vast repository of scanned Soviet statistical books has been created by public users to preserve historical data of the lost Soviet empire; see http://istmat.info/. There could be a concern that the aggregate data from the Soviet times might be distorted (Ofer, 1987; Rosefielde, 2003). However, compared to many other economic indicators, official Soviet reports regarding higher education have been fairly reliable, as the number of student slots, the faculty-student ratio and other education indicators were centrally set, i.e., not collected from the bottom up and linked to plan targets like production numbers. Hence, there was no direct material incentive for local universities to falsify this information. We did not find evidence of the intentionally falsified Soviet statistics on education in previous Western or Russian studies.

[18] Some national averages have to be imputed for a few years in the 1950-1960s (see Appendix A1 for details on each imputed measure). The share of our estimation sample with imputed measures is insignificant (less than 5 percent). These are mostly older individuals surveyed in the 1990s.

study changes in college quality since the Soviet era and to estimate the time-varying returns to college quality in response to the recent expansion of higher education.

4. Measurement and Distribution of College Quality

Our ultimate objective is to determine whether and how the expansion of higher education shifts the distribution of wages of college graduates and what part of this shift can be attributed to (i) changes in the distribution of college quality (composition effect) and (ii) changes in the distribution of market returns to quality (price effect). In this section, we focus on the composition effect of college expansion, with a special emphasis on measuring college quality over time.

4.1 Approaches to Measuring College Quality

The college quality Q^* is a latent variable that can be approximated by many observable indicators, but any single indicator q_k can only partially capture the true quality. The choice of indicators is primarily driven by available data, rather than by a theory; although, there is a long tradition in the literature that goes back to Solmon (1975) to isolate and measure the two sets of college attributes: student quality (e.g., SAT scores of entering freshmen, high school GPA, percent of accepted applicants, etc.) and instructional quality (e.g., faculty-student ratio, faculty wages, expenditures per student, library size, etc.). As noted by Solmon (1975) and many others, different college attributes are highly collinear, and thus including many of them in one regression would reduce the estimation efficiency and complicate sensible interpretation of estimated parameters. Attempts to reduce the multicollinearity problem by either using each proxy one at a time or combining them into some kind of a weighted average index[19] do not come without the cost. The concern here is not only the arbitrariness in the choice of indicators and weights, but also the

[19] The most commonly used index in the U.S. literature on college quality is the Barron's selectivity index that designates each college into one of the six categories from non-competitive to most competitive based on the entering class's SAT and ACT scores, class rank, high school GPA, and the proportion of accepted applicants (Brewer, 1999; Light and Strayer, 2000; Monks, 2000). The discrete categorization of colleges can be found in other countries; e.g., elite schools in China (Hongbin *et al*, 2011) or prestigious schools in UK (Chevalier and Conlon, 2003). While between-group comparisons can be informative, they discard within-group variation in quality.

attenuation bias in the estimated effect of q_k on some outcome due to reduced variance and an increased noise-to-signal ratio. Black and Smith (2006) derive that, in addition to the attenuation bias coming from the measurement error u_k, the deviation of a scaling factor α_k from 1 in $q_k = \alpha_k Q^* + u_k$ also generates the bias of unknown direction; hence, using multiple indicators in some aggregated way without the loss of variation is generally preferred.

In this study, we use the composite index of student quality based on the test scores of entering students and then calculate the first principal component of multiple proxies for college quality, including measures of instructional resources (also used in Black and Smith, 2004; Long, 2008). The latter approach is in line with the most recent literature on using multiple proxies for measuring college quality. However, we depart from the existing literature in that we disaggregate entering test scores by university-major. Students in Russia declare their major at the time of application and follow major-specific curricula from the first semester onward, without mingling with students from other majors in their class work. Universities appoint independent admission committees and set different test score thresholds for each major separately. The aggregation of test scores at the level of university, as it is done in the U.S. literature, would result in a substantial loss of within-college variation in average ability of entering students across different majors.[20] Our data allow us to exploit both types of variation.

4.2 Measuring Time-Varying College Quality

College quality is difficult to measure for one cohort of students. It is even more challenging to get consistent measures of college quality for multiple cohorts over a long period of time. Not surprisingly, studies of the returns to time-varying quality measures are practically non-existent, with the exception of multiple cohort samples (e.g., Brewer *et al*, 1999; Long, 2010) and aggregate measures (e.g., cohort-specific college enrollment by region of birth as in Carneiro and Lee, 2011).

[20] In our data of EGE scores averaged by university-major for a given program and year, the within-college variation in test scores between subject majors accounts for 40-50 percent of total variation in test scores; the remaining 50-60 percent is due to between-college variation.

The limitation of the former approach is a small number (up to three) of independently drawn cohort samples with differences in definitions across databases; while the limitation of the latter approach is that it removes all within-regional variation in student ability across universities and majors and omits the resource component of the college quality. The majority of college quality literature relies on time-constant, but university-varying, quality measures for a given cohort of students in a cohort-based panel (such as NLSY).

Given the mass expansion and transformation of the Russian educational system in the last two decades, having time variation in quality measures is essential for estimating the cross-cohort shifts in the returns to college quality. The usual trade-off is that lengthier time period and larger number of cohorts imply a smaller set of available quality measures, especially if we go back to the Soviet era. Still, our study takes a step forward, compared to the U.S. literature, in identifying a number of time-varying proxies for college quality over the sixty-year period. First, in the spirit of Carneiro and Lee (2011), we use region-specific enrollment rates as a proxy for student ability from 1950 to 2011. Second, we supplement enrollment data with the aggregate measures of resources such as faculty-student ratio and wages in education industry. Third, based on the name of university, we identify several observed correlates of quality, such as public vs. private university, classical vs. specialized university, college location, main campus vs. branch, university founding date, etc. While these characteristics are constant for an individual in the panel database, they vary across cohorts, as the composition of graduates by university type changes over time. Finally, we employ tuition and average test scores by university-major-program. Because test scores are only available for college entrants in 2010-2012 years, we use the following decomposition procedure to introduce time variation.

Let's suppose we have two programs of study, a budget full-time program A and a part-time and tuition-based program B.[21] The size of program A changed little, but program B expanded considerably over time. The average student quality, θ_{jmc}^{P}, in university j, subject major m, program P = {A, B}, and cohort c can be presented as a deviation from the average test score of entering freshmen in region r, i.e., $\theta_{jmc}^{P} = \bar{\theta}_{rc} + \Delta\theta_{jmc}^{P}$, with $\Delta\theta_{jmc}^{A} > \Delta\theta_{jmc}^{B}$. Hence, the average test score of all admitted students can be written as a weighted average of student quality in both programs:

$$\theta_{jmc} = \bar{\theta}_{rc} + p_{jmc}^{A}\Delta\theta_{jmc}^{A} + (1 - p_{jmc}^{A})\Delta\theta_{jmc}^{B}, \quad (1)$$

where p_{jmc}^{A} is the share of students in program A. Let $\kappa_{jmc} = \frac{\Delta\theta_{jmc}^{B}}{\Delta\theta_{jmc}^{A}} < 1$ be the quality ratio between the two programs and re-write equation (1) as

$$\theta_{jmc} = \bar{\theta}_{rc} + \Delta\theta_{jmc}^{A}\left(\kappa_{jmc} + \left(1 - \kappa_{jmc}\right)p_{jmc}^{A}\right) \quad (2)$$

Therefore, changes in the average quality of entering students over time (across cohorts) depend upon changes in the regional average student quality, the university-major deviations from the region-cohort mean value, the quality differences between the two programs, and the share of students enrolled in a higher-quality program A. Most of these components are not observed in actual data, as test scores are only available for the last cohort $\bar{c}$, $\Delta\theta_{jm\bar{c}}^{A}$ and $\kappa_{jm\bar{c}}$. However, we can make reasonable simplifications/aggregations to capture at least partial dynamics in test scores.

First, $\bar{\theta}_{rc}$ is negatively related to the enrollment rate, e_{rc}, that we observe for every region and every cohort. As shown in Appendix A2, if test scores are uniformly distributed on a scale from 0 to 100, then the relationship between the average test score of admitted students and enrollment rate is conveniently linear, that is, $\bar{\theta}_{rc} = 100 - 50e_{rc}$. For test scores with truncated normal distribution, the functional form is more complicated, but the numerical solution in Appendix A2 shows that under reasonable assumptions $\bar{\theta}_{rc}$ is inversely related to e_{rc}, with a negative correlation of -0.95 due to non-linearity. Second, while we do not observe a time-varying share of students in

[21] In Russian case, program B includes all tuition-based programs at both public and private universities as well as publicly-financed programs of part-time education.

program A for each university-major, we know how this share changed for each major over time, so we can capture p^A_{jmc} via $\bar{p}^A_{mc}$. An observed decline in $\bar{p}^A_{mc}$ could be partly responsible for the declining quality of students. Third, the average (by university-major) deviations of test scores for past cohorts can be predicted based on an individual's major $M^{P=A,B}_{mc}$ and a broad range of observed university characteristics $X^{P=A,B}_{jc}$ using the last cohort distribution of scores.

$$\Delta\theta^P_{jm\bar{c}} = \beta^P_{\bar{c}} X^P_{jc} + \gamma^P_{\bar{c}} M^P_{mc} + \varepsilon^P_{jm\bar{c}} \ \rightarrow\ \widehat{\Delta\theta}^P_{jmc} = \hat{\beta}^P_{\bar{c}} X^P_{jc} + \hat{\gamma}^P_{\bar{c}} M^P_{mc} \quad (3)$$

By substituting e_{rc}, $\widehat{\Delta\theta}^P_{jmc}$ and $\bar{p}^A_{mc}$ into (1), we obtain a "composite index of student quality":

$$\tilde{\theta}_{jmc} = f(e_{rc}) + \bar{p}^A_{mc}\widehat{\Delta\theta}^A_{jmc} + (1 - \bar{p}^A_{mc})\widehat{\Delta\theta}^B_{jmc} + u_{jmc}, \quad (4)$$

where $f(e_{rc})$ is a numerical solution for $\bar{\theta}_{rc}$, assuming the truncated normal distribution of test scores on a scale from 0 to 100, and u_{jmc} is the measurement error that is assumed to be independent of the true average student quality, with zero mean. With this approach, the cross-cohort variation in test scores comes from changes in (i) enrollment rate, e_{rc}, (ii) relative size of programs, $\bar{p}^A_{mc}$, and (iii) composition of majors M^P_{mc} and university characteristics X^P_{jc} over time, holding the estimated weights $\hat{\beta}^P_{\bar{c}}$ and $\hat{\gamma}^P_{\bar{c}}$ constant. Despite assumptions of constant β's and γ's, this measure is an improvement over other compositional indicators used in the literature such as region-specific enrollment rates, because it also accounts for quality-related changes in the distribution of students across university characteristics, majors, and programs of study. We interpret the estimated labor market effects of predicted test scores as compositional effects.

4.3 Composite Index of College Quality: Trends and Distribution

For the RLMS sample of college graduates, we first show the trends in each of the four elements of our aggregate student ability measure: $\overline{\theta}_{rc}$, $\bar{p}^A_{mc}$, $\hat{\beta}^P_{\bar{c}} X^P_{jc}$, and $\hat{\gamma}^P_{\bar{c}} M^P_{mc}$. The first element, $\bar{\theta}_{rc} = f(e_{rc})$, is calculated based on region-cohort enrollment rates, as described in Appendix A2. The trends in the regional average test score in Figure 5A mirror changes in the number of students and the size of cohort. A sizeable drop in $\overline{\theta}_{rc}$ in the mid-1960s is attributed to the small size of population aged 18-22 as a consequence of low birth rates during the World War II. College expansion since the mid-1990s resulted in a more recent decline in $\overline{\theta}_{rc}$; although the shrinking size of college-age population after 2005 started contributing to the latest decline as well.

The second element, $\bar{p}^{A}_{mc}$, – the share of college graduates who completed a publicly-financed full-time program by major – is depicted in Figure 5B. The share varies substantially by major: e.g., armament, math and physics have predominantly full-time budget programs, while the share of full-time budget programs in economics, business, and management was only 12 percent in 2008. The $\bar{p}^{A}_{mc}$ had a U-shaped trend during the Soviet era dropping to almost 40 percent at the peak of correspondence education in the mid-1960s, but subsequently recovering to about 55-60 percent. However, since the beginning of college expansion in the 1990s, the trend in $\bar{p}^{A}_{mc}$ across all majors has been downward.

To calculate the last two elements of average student ability measure, we estimate equation (3) by regressing deviations from $\overline{\theta}_{rc}$ on university characteristics and college majors for programs A and B separately and report results in columns 1 and 2 of Table 2. We also report results for average EGE, without subtracting from the regional mean, in columns 3 and 4. The sign of the estimated effects of various university characteristics on test scores is fairly similar between the two programs. A lower average EGE is found to be in branches and universities located in non-central cities. Medicine, law, and governance majors attract students with higher marks, while agriculture and pedagogy enroll lower-scoring students. High-ability students tend to choose universities with time-honored traditions going back to 18th and 19th centuries and the early Soviet era. These specifications have a relatively high predictive power, as observed college-major characteristics explain almost three quarters of variation in test score deviation from the regional mean.

We use the estimates in columns 1 and 2 of Table 2 to predict the remaining two elements of student quality shown in Figures 5C and 5D. As expected, the predicted test scores are significantly higher in program A than in program B. An overall downward trend in $\hat{\beta}_{\bar{c}}^{P} X_{jc}^{P}$ reflects an increased share of graduates from colleges with lower $\hat{\beta}_{\bar{c}}^{P}$'s (e.g., branches and emerging private schools), whereas an upward trend in $\hat{\gamma}_{\bar{c}}^{P} M_{mc}^{P}$ reflects a long-term shift towards majors with higher contemporary $\hat{\gamma}_{\bar{c}}^{P}$'s (e.g., business and law).

Using the above four elements, we construct a composite index of student quality ($SQ1$) according to equation (4) and plot its density and time trend in Figure 6. Panel A shows the distributional shift in the student quality between 1995 and 2010, while Panel B plots the kernel densities of $SQ1$ for the cohorts educated in 1986-1998 and 1999-2011. We observe a significantly lower mean level and a much higher variance of average test scores in 2010 compared to 1995 and among recent graduates. Panel C, where $SQ1$ is plotted over the entire 60-year period, also shows a clear downward trend in the student quality measure for more recent cohorts, which is an expected outcome of massive college expansion and associated compositional changes. Panel D displays an additional measure of student quality to be used in sensitivity analysis. $SQ2$ is a linear prediction of test scores based on university characteristics, major and program of study using estimates in columns 3 and 4 of Table 2. It captures compositional changes across university characteristics, majors and programs of study, but it does not account for rising enrollment rates, as $SQ1$ does. As a result, $SQ2$ is trending downward with much less pronounced rate of descent compared to the $SQ1$ measure.

Finally, we construct the composite index of college quality by retaining the first principal component of the aggregate student ability measure, $SQ1$, and six other proxies of college quality. Two of the proxies are cohort-specific measures with partial variation across regions, such as wages in education industry and faculty-student ratio in HPE. They are available for the entire 1950-2010 period with minimal imputations (see Appendix A1 for details). The other four proxies provide a richer cross-sectional variation by university and sometimes by university-major but do not have time variation, such as tuition for each major, available instructional space per student, revenues per faculty, and the share of foreign students.[22] The cross-cohort variation in these measures comes from changes in the composition of college graduates by university characteristics. Table 3 reports the loadings of the 7-factor index of college quality ($PC7$), where loadings are the weights of the linear combination of quality proxies that maximizes the overall variance (the first principal component accounts for about 30 percent of the variance). For the robustness checks, we also employ the 4-factor composite index ($PC4$) that extracts the first principal component of $SQ1$, tuition, wages in education, and faculty-student ratio. The mean of both composite indices is trending downward in the 2000s, while their cross-sectional variance is rising (see Figure 7). We also observe a relative increase in the supply of college graduates from a lower end of the quality distribution.

To sum up, the descriptive analysis in this section draws a relatively consistent picture of the overall decline in average college quality and its various components during the period of mass expansion of higher education. This result is consistent with a scarce U.S. literature showing a decrease in the average test scores of admitted students in the regions with higher college enrollment (Carneiro and Lee, 2011) and a decline in resources per student outside the top universities (Bound and Turner, 2007; Bound *et al.*, 2010). Our evidence of the increased variance in quality as well as the increased relative supply of college graduates from lower quality programs will play an important role in understanding how the expansion of higher education changes the returns to college quality.

---

[22] Table S1 of Appendix A3 shows that Moscow universities and older public universities, on average, have advantage in practically every available measure of quality. The results by university type are less consistent across measures, but the sign of coefficients is plausible. For example, tuition charges are much higher in schools that specialize in economics, business, and management, followed by law and humanities. Medical schools attract more foreign students. Engineering and agricultural space provide more instructional space per student. Controlling for the type of schools, some of the most expensive majors include health care, international relations, information security, and architecture.

5. Returns to College Quality Before and During the Expansion

5.1 Conceptual Framework

Now we turn to the effect of college expansion on equilibrium prices or monetary returns to quality (RTQ). The identification of the price effect is challenging, since the related issues of non-random selection into college, the endogeneity of college quality, consistency over time, and price heterogeneity are each individually complicated and need to be addressed all at once. It is useful to first discuss what effects can be captured with traditional estimators of RTQ and then identify some of the channels of how college expansion can affect RTQ. For the moment, we ignore the time dimension and assume that the logarithmic hourly wage of college graduates (Y_i) varies with college quality (θ_i), other observed individual characteristics (X_i), sample selection factors (λ_i), unobserved individual ability (a_i), and *iid* transitory shocks and statistical noise (ϵ_i). The monetary returns to quality (b_i) are individual-specific and allowed to correlate with college quality.

$$Y_i = a_0 + b_i\theta_i + a_1 X_i + \lambda_i(X_i, W_i) + a_i + \epsilon_i, \tag{5.1}$$

$$D_i = \mathbf{1}[\delta_0 + \delta_1 X_i + \delta_2 W_i + \zeta_i > 0],\ \zeta \sim N(0,1), \tag{5.2}$$

$$a_i = \rho\theta_i + u_i, \tag{5.3}$$

$$b_i - \bar{b} = \psi(\theta_i - \bar{\theta}) + v_i. \tag{5.4}$$

The way how individual heterogeneity is modeled here is similar to generalized earnings functions with heterogeneous returns to years of schooling and individual-specific intercept (Card, 2001). The parameter a_0 is determined such that the mean of unobserved ability is zero, $E[a_i] = 0$. The parameter ρ is expected to be positive due to assortative matching, i.e., high-quality colleges selecting high-ability students. The parameter ψ is likely to be negative due to diminishing marginal returns to each additional unit of quality. A selection term (inverse Mills ratio) λ_i is included in the model to account for the non-random selection of individuals into college, and it is obtained from the probit estimation of (5.2) with a valid exclusion restriction W_i; $D_i = 1$ if an individual graduates from college. Alternatively, inverse propensity weighting (IPW) can be employed, which could also account for other sources of missing data.

Other assumptions:

5a) (Y, X, W) are always observed, θ is observed only when $D = 1$,

5b) (ϵ, ζ) are independent of (X, W, θ),

5c) (u, v) are independent of θ, $E[\theta u] = E[\theta v] = 0$,

5d) $E[\epsilon] = E[\ \zeta] = E[u] = E[v] = 0$,

5e) λ and ϵ are additively separable and linear in ζ[23]

By estimating $\hat{\lambda}_i$ from (5.2) and substituting equations (5.3) and (5.4) into (5.1), we obtain earnings function (6) with linear and quadratic college quality terms.

$$Y_i = a_0 + a_1 X_i + \left(\bar{b} + \rho - \psi\bar{\theta}\right)\theta_i + \psi\theta_i^2 + \hat{\lambda}_i + \varepsilon_i, \tag{6}$$

where $\varepsilon_i = \theta_i v_i + u_i + \epsilon_i$ with $E[\varepsilon] = 0$ under assumptions (5c) and (5d). We also assume that $E[\theta\varepsilon] = 0$.[24]

Under these assumptions, the probability limit of the OLS coefficient on the linear θ_i term is $plim\ \hat{b}_{OLS} = \bar{b} + \rho - \psi\bar{\theta}$. Given the expectations regarding $\rho > 0$ and $\psi < 0$, an OLS estimate of $\hat{b}_{OLS}$ is likely to be larger than population-average RTQ ($\bar{b}$) due to ability bias (ρ) and heterogeneity bias ($-\psi\bar{\theta}$).

Now let's consider two cohorts (o and y) and estimate equation (6) by OLS for each cohort separately at the same point in their life cycle. The between-cohort difference in RTQ (evaluated at the mean quality $\bar{\theta}$), can be decomposed into four different sources:

$$RTQ|_{\bar{\theta}}^{y} - RTQ|_{\bar{\theta}}^{o} = \left(\bar{b}^y + \rho^y - \psi^y\bar{\theta}^y + 2 \cdot \psi^y\bar{\theta}^y\right) - \left(\bar{b}^o + \rho^o - \psi^o\bar{\theta}^o + 2 \cdot \psi^o\bar{\theta}^o\right)$$

[23] The linearity assumption is weaker than a standard assumption of joint multivariate normal distribution of errors, as shown by Wooldridge (2002, chapter 17).

[24] The assumption $E[\theta\varepsilon] = 0$ requires either the orthogonality between (ϵ, u, v) and (θ, θ^2) or zero conditional expectation assumption $E[\epsilon|\theta] = E[u|\theta] = E[v|\theta] = 0$.

$$= \underbrace{(\bar{b}^y - \bar{b}^o)}_{\text{"average price effect"}} + \underbrace{(\rho^y - \rho^o)}_{\text{"assortative matching effect"}} + \underbrace{(\psi^y - \psi^o)\bar{\theta}^o}_{\text{"price heterogeneity effect"}} + \underbrace{(\bar{\theta}^y - \bar{\theta}^o)\psi^y}_{\text{"change in quality effect"}}, \quad (7)$$

The average price effect would generally be considered as the only relevant parameter of interest, while the rest are the biases that need to be eliminated. We are going to argue, however, that the other three effects are also relevant margins of price adjustment in response to changes in the quantity and quality of new college graduates, and that the contribution of these effects amplifies when the college enrollment is on the rise.

Suppose o and y stand for the cohorts graduated before and during the expansion period, respectively. Based on the facts established in Section 4, we can predict that as a result of college expansion:

(i) The "change in quality effect" is positive due to a decline in average quality.

(ii) The "price heterogeneity effect" is likely to be positive. A greater supply of college graduates from the lower tail of quality distribution and a reduced supply of graduates from the upper tail would presumably lower RTQ for the former group and increase RTQ for the latter group, which implies that the "price-quality" profile gets flatter (i.e., ψ decreases in absolute value).

(iii) The "assortative matching effect" is also expected to be positive. A relative price change favoring high-quality education would raise the marginal benefits for high-ability students to enter high-quality schools, thus improving the "quality-ability" match and increasing ρ.

Thus, we can formulate the following testable propositions:

1. *The expansion of higher education is expected to flatten the "price-quality" profile by increasing the returns to higher quality programs relative to lower-quality programs.*

2. *The expansion is likely to improve the assortative matching between students and colleges.*

3. *The deviation of cross-cohort difference in RTQ* $(\Delta RTQ|_{\bar{\theta}})$ *from the average price effect is predicted to be greater when a lower tail of quality distribution expands and the average quality falls.*

5.2 Exclusion Restrictions

To test the above propositions, we can take advantage of the Russian multi-cohort panel with earnings at various points of individual life cycle and exploit some of the exogenous shocks from the turbulent Soviet/Russian history. The latter shocks are useful in finding credible exclusion restrictions W in selection equation (6.3).[25] Exclusion restrictions are chosen for age 17 when individuals typically make their decision to go to college in Russia. They are analogous to an instrumental variable for schooling/college in a standard Mincerian earnings function.[26]

First, we use the wage ratio between manual and non-manual industrial workers at age 17. Ideologically-justified wages of manual workers were always set high by the Soviet government, sometimes exceeding wages of engineers and technicians (see Figure 4A). However, the transition to a market economy practically sent wages of manual workers to a free fall and created a colossal wage gap of 55-60 percent. These large movements in the wage ratio were likely unanticipated by most secondary school students.

The second exclusion restriction is common in the human capital literature and relies on cohort-specific changes in the laws related to secondary education (e.g., Acemoglu and Angrist, 2000; Oreopoulos, 2006). In December 1958, the Soviet government increased the length of general secondary education required for college admission from 10 to 11 years (Soviet Law, 1958). The first cohort that had to study eleven years was born between September 1, 1944 and August 31, 1945 (7th graders in 1958-59). When the first "treated" cohort was in the 11th grade, the cohort born one year earlier (1943-44) had an extra year to apply to college without facing much

[25] In the existing RTQ literature, the decision to attend college is either omitted by restricting analysis to the sample of college graduates, or it is added as an alternative category (no college attendance) to a few broad quality categories (e.g., Light and Strayer, 2000). The latter approach is not suitable in the case of continuous and time-varying quality measures, unless there is a valid reason to suppress an individual variation in college quality both within the quality category and especially over time. The possibility that the decision to attend college might be governed by a different set of factors than the decision to apply for a college of certain quality is another motivating factor to have a selection equation for college participation.

[26] Some common IVs are not applicable to our case because of their irrelevance to the Soviet period (e.g., unemployment rates were zero or even negative due to labor shortages) or because they serve as a proxy for college quality such as tuition.

competition. The competition pressure for neighboring cohorts has also reduced as a result of this change. But in 1964, the Soviet Union returned to a 10-year secondary education, hence effectively forcing the two cohorts of secondary school students (born between September 1, 1946 and August 31, 1948) to graduate in 1965 and compete against each other for the same college slots (Soviet Law, 1964). In response to these changes in schooling laws, we expect the college graduation rates to be higher in the 1942-45 cohort and lower in the 1946-48 cohort, controlling for the number of approved new slots in a given year.[27]

The final set of exclusion restrictions includes changes in the rules for the military conscription of students.[28] In Russia, all males over age 18 are required to undergo military training. The compulsory length of military service varied from 12 to 36 months (see Figure 8). During certain periods, full-time students were either exempt from military service (1944-1960) or received deferments until the graduation with a shorter length of required service (1968-1983, 1989-present). Twice in the post-WWII Soviet history, college students were obliged to serve on a general basis: in 1961-1967 because of the sharp decline in the male population of young ages and in 1984-1988 because of the war in Afghanistan and unfavorable demographic trends. These conscription rules clearly influence the opportunity cost of attending college, and our data reflect that.

As shown in Table 4, column 2, all three exclusion restrictions are important determinants of college decisions (joint χ-squared = 546.78). College enrollment within a cohort declines when the competition between adjacent cohorts intensifies for limited student slots. It also falls when

[27] In 1984, the Soviet government announced again a gradual transition to 11 years of general secondary education, but the new system was implemented rather sparingly. In most cases, students skipped one grade, thus finishing 11 grades on paper, but 10 grades factually. It was only in 2001, when 11 years became compulsory for the first graders, who graduated in 2012 instead of 2011 (Federal Law, 2001). As a result, universities had insufficient number of applicants in 2011, leaving many authorized slots unfilled. In our 2011 survey, we do not observe yet the labor market consequences of this policy change, but this natural experiment can make an interesting case for future research.

[28] In studies of other countries, wars and military-related policies have been used to address the endogeneity of education (e.g., Ichino and Winter-Ebmer, 2004; Maurin and Xenogiani, 2007; Angrist and Chen, 2011).

earnings of manual workers compared to non-manual workers rise at the time an individual is 17 years old. The likelihood of males to obtain a college degree rises with greater difference in the length of required military service between students and non-students. These results are intuitive and expected.

The exclusion restrictions help to correct for selection bias either via the Heckman control function with inverse Mills ratio or via IPW based on the inverse propensity score. The results do not differ much between the two methods, as shown in Table 4, columns 3 and 4. However, non-random selection into higher education is not the only reason for the selection bias in equation (6). Missing earnings, missing college characteristics, and missing individual *X*s lead to an estimation sample that is likely to be different from the general population. In the case of multiple sources of missing data, the IPW approach is simpler to implement and generally preferred in the recent econometrics literature on selection due to missing data (Wooldridge, 2007).

[Sections 5.3-5.5 and 6 to be completed]

5.3 Price Heterogeneity

The conventional OLS estimation of equation (6) with quadratic quality terms is a useful starting point in testing whether the earnings-quality profile becomes flatter during the expansion period. According to decomposition (7), the coefficient on the quadratic term can capture the "price heterogeneity" effect.

There are many choices in estimating equation (6) with regard to definitions of earnings (primary job vs. all jobs, received last month vs. usual, monthly vs. hourly, raw vs. adjusted for price inflation and outliers, etc.), measures of college quality (see Section 4), and which controls to include in vector *X*. We begin with the baseline specification that includes traditional controls (and retains the largest sample size) and then perform sensitivity analysis by using alternative measures and adding less common and partially missing covariates (such as place of birth, tenure, and college education of parents). The estimation is performed for all college graduates ages 20 to 60 and separately for the two neighboring cohorts – educated in 1986-1998 and 1999-2011. The year 1999 is chosen as a threshold year because it was the first year when the number of college graduates significantly exceeded the 1980s level.

- Column 1 of Table 4 presents the results for our baseline OLS specification. We use *SQ1* as a measure of quality and obtain a coefficient which is positive and statistically significant. A 1-point increase in quality measured by ranking *SQ1* leads to a 0.4 percent increase in the log of hourly wages. We control for the log of the number of freshmen and find that it has negative and statistically significant effect on hourly wages. A 10 percent increase in the size of entering cohort decreases hourly wages by 0.51 percent. This finding is in line with the recent literature that establishes the negative link between college enrollment and wages of college graduates in a given cohort (Carneiro and Lee, 2009, 2011). Panel B of Table 4 reports the results for an OLS specification with a square term for college quality added to the specification. In this model, the effect of quality is higher and the coefficient on quality squared is negative and statistically significant. This coefficient captures "price-heterogeneity effect" and we interpret it as an evidence of wages-quality profile flattening during the period of expansion.
- To control for the nonrandom selection of individuals into college, we estimate a Heckman selection model. The estimates of the selection equation and wage equation are presented in Columns 2 and 3 of Table 4. The selection equation determining the decision to graduate from college includes all the individual level controls used in the wage equation, plus the exclusion restrictions. We use exclusion restrictions that are described in detail in Section 5.2. The effects of quality and log of cohort size on wages are higher in the Heckman model, and imply that selection is in fact a nonrandom process. To test the robustness of these results we report the estimates of IPW and IPW2 models in Columns 4 and 6 respectively. The effect of quality remains of similar magnitude and highly significant across specifications. In Table 5 we report the results of IPW2 model estimated with different measures of college quality.

- Linear mixed model:

$$lnw_{kt} = \beta \cdot \theta_k + \gamma \cdot X_{kt} + u_{k1} + u_{k2} \cdot \theta_k + \varepsilon_{kt}, \quad (8)$$

where u_{k1} is constant individual heterogeneity (ability) and u_{k2} is individual-specific returns to college quality. Change in returns to quality = $\beta + u_{k2}$ varies by individual. The correlation is allowed between u_{k1} and u_{k2} (abilities and RTQ).

Results of the panel mixed model are reported in Table 7 and Figure 9. Main results of the mixed model:

- $Cov(u_{k2}, \theta_k)$ is larger for younger cohort (flatter prices-quality profile)
- $Cov(u_{k2}, u_{k1})$ is larger for younger cohort (stronger relationship between abilities and returns during the expansion period)
- $Cov(\theta_k, u_{k1})$ is larger for younger cohort (stronger assortative matching during the expansion period)

- Thus, results in this section suggest that individuals graduated during the period of expansion have (1) higher price dispersion; (2) flatter price-quality profile; (3) increased RTQ from higher-quality programs and decreased RTQ from lower-quality programs; (4) better assortative matching; and (5) stronger correlation between abilities and RTQ.

5.4 Assortative Matching

- Now the challenge is how to separate $\overline{b}^y and\ \rho^y$. One way is to use control function approach by explicitly modeling ρ. The other way is finding instrumental variables that are not correlated with abilities.
- Potential IVs: average quality in student's location at age 17, the share of freshmen entering program A, and the share of graduates from program A
- Preliminary estimates show that IV estimates of the returns to quality are significantly higher than the OLS estimates.
- IV is likely to generate its own biases

5.5 Prices and College Expansion

By re-estimating the models in subsections 5.4 and 5.5 for cohort c and time period t, we can retrieve the following time-varying parameters B= {$\bar{b}_{ct}^{OLS} + \rho_{ct}^{OLS}$; ψ_{ct}^{OLS}; b_{ct}^{IV}; b_{ct}^{MM}; $ov\left(b_{ict}^{MM}, q_{ict}\right)$, $Cov\left(b_{ict}^{MM}, a_{ict}\right)$, and $Cov(q_{ict}, a_{ict})$}. We can directly test how the estimated parameters are changing with enrollment rates by estimating a simple model:

$$B_{ict} = \vartheta_0 + \vartheta_1 Enroll + \vartheta_c + \vartheta_t + \epsilon_{ict},$$

where B is one of the parameters of interest and *Enroll* is enrollment at the time when individuals were 17-year old. Time periods include the late Soviet period (1985-1990), the transition period (1995-1998), and the modern period (2000-2011), and cohorts are defined based on the year when individuals received college education in the 1950s, 1960s, 1970s, 1980s, 1990s, and the 2000s. The preliminary estimates (not shown yet) suggest that various estimates of the returns to college quality (including price heterogeneity and assortative matching components) are positively associated with college expansions.

6. Decomposition of Wage Variance: Quality vs. Prices

In this section, we use the two cohorts of college graduates described in Section 5.3 to decompose the difference in the variance of their wages into price and quality components. We estimate the re-centered influence function (RIF) to perform the decomposition (Firpo et al. 2009).

- Describe method
- Describe results

7. Conclusions

This paper estimates changes in the returns to college quality in response to the expansion of higher education. We use a unique linked dataset that is constructed based on a 17-year nationally representative panel of individuals, the university database, official regional and national statistics from 1950 to 2011, and the archives of Soviet and Russian laws. We construct measures of quality using the first principal component of multiple proxies for college quality. Returns to quality increase for the cohorts that receive college education during the expansion. Moreover, returns to high quality colleges increase while returns to low quality colleges diminish. The variance of individual wages due to quality increases, and the variance due to prices decrease. Expansion improves assortative matching between students and colleges of different qualities, i.e., high ability students are more likely to enroll into high quality colleges.

We find that the distribution of college quality shifts to the left due to the expansion. Having realized the inefficiency of college system, Russian government has taken unprecedented measures to shut down the ineffective institutions and their branches. However, the arbitrariness of the measures of efficiency which are used for the evaluations poses a threat of misclassification. Colleges defined as "inefficient" by a set of criteria may not in fact be of low quality. We argue that the market of higher education will eventually reach equilibrium by itself. High competition and declining demand for college education based on the shrinking population size will drive inefficient low quality institutions out of the market. From the policy perspective, providing more information for students when they apply to college may improve the allocation of students across colleges. Better information on state accreditation, enrollment, faculty, and employment after graduation, among many others, will help students make a better decision.

References

Almlund, Mathilde, Angela Lee Duckworth, and James Heckman, 2011. Personality Psychology and Economics, *Handbook of the Economics of Education*, Volume 4: 1-181.

Andrews, Rodney J., Jing Li, and Michael F. Lovenheim, 2012. Quantile Treatment Effects of College Quality on Earnings: Evidence from Administrative Data in Texas, NBER Working Paper No. 18068.

Angrist, Joshua, and Stacey Chen, 2011. Schooling and the Vietnam-Era GI Bill: Evidence from the Draft Lottery, *American Economic Journal: Applied Economics* 3(2): 96-118.

Becker, Gary S., 1964. Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. New York: Columbia University Press.

Black, Dan, and Jeffrey Smith, 2004. How Robust Is the Evidence on the Effects of College Quality? Evidence from Matching, *Journal of Econometrics* 121 (1-2): 99-124.

Black, Dan, and Jeffrey Smith, 2006. Estimating the Returns to College Quality with Multiple Proxies for Quality, *Journal of Labor Economics* 24(3): 701-728.

Bound, John, Michael F. Lovenheim, and Sarah Turner, 2010. Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources, American Economic Journal: Applied Economics 2: 129-157.

Bound, John, and Sarah, Turner, 2007. Cohort crowding: How resources affect collegiate attainment, *Journal of Public Economics* 91: 877-899.

Brewer, Dominic, Eric Eide, and Ronald Ehrenberg, 1999. Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings, *Journal of Human Resources* 34(1): 104-123.

Card, David, 2001. Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems, *Econometrica* 69(5): 1127-1160.

Carneiro, Pedro, and Sokbae Lee, 2011. Trends in Quality-Adjusted Skill Premia in the United States, 1960-2000, *American Economic Review* 101(6): 2309-2349.

Carneiro *et al*. (2011)

Chevalier, Arnaud, and Conlon Gavan, 2003. Does It Pay to Attend a Prestigious University? Unpublished manuscript. Dublin: University College.

Constitution of the Russian Federation, 1993. December 12, 1993.
Available at http://www.consultant.ru/online/base/?req=doc;base=LAW;n=2875

Dale, Stacy, and Alan B. Krueger, 2011. Estimating the Return to College Selectivity over the Career Using Administrative Earnings Data, NBER Working Paper No. 17159.

De Witt, Nicholas, 1961. *Education and Professional Employment in the USSR*. Washington: United States Government Printing Office.

Education in Russia-2006, 2007. Statistical bulletin. Moscow: Moscow State University of Tool-making and Information MGUPI

Education in Russia-2008, 2009. Statistical bulletin. Moscow: Moscow State University of Tool-making and Information MGUPI.

Education in Russian Federation: 2010, 2010. Statistical bulletin. Moscow: National Research University "Higher School of Economics".

Education in Russian Federation: 2012, 2012. Statistical bulletin. Moscow: National Research University "Higher School of Economics".

Federal Law, 1992. *Ob Obrazovanii [On Education]*, July 10, 1992.
Available at http://www.consultant.ru/popular/edu/

Federal Law, 1996. "O vnesenii izmeneniy v zakon Rossiyskoy Federatsii "O voinskoy obyazannosti i voennoy sluzhbe" [On changes in Federal Law "On military duty and military service], May 9, 1996.

Federal Law, 2001. "Ob utverzhdenii tipovogo polozheniya ob obscheobrazovatel'nom uchrezhdenii" [On the approval of typical status of educational institution], March 19, 2001. Available at http://www.consultant.ru/online/base/?req=doc;base=LAW;n=85826#p118.

Federal Law, 2006. "O vnesenii izmeneniy v otdel'nye zakonodatel'nye akty Rossiyskoy Federatsii v svyazi s sokrascheniem sroka voennoy sluzhby po prizyvu" [On the introduction of changes to the legislative acts of the Russian Federation in relation to to the reduction of length of required military service], July 6, 2006.

Firpo, Sergio, Nicole Fortin, and Thomas Lemieux, 2009. Unconditional Quantile Regressions, *Econometrica* 77(3): 953-973.

Fleisher, Belton, Klara Sabirianova, and Xiaojun Wang, 2005. Returns to Skills and the Speed of Reforms: Evidence from Central and Eastern Europe, China, and Russia, *Journal of Comparative Economics* 33(2): 351-370.

Gatsko, M.F., 2008. *The Legal Support of Armed Forces of the Russian Federation*. Moscow: Flinta, Nauka. *In Russian.*

Going Global, 2012. The Shape of Things to Come: Higher Education Global Trends and Emerging Opportunities to 2020. British Council.

Goskomstat, 2006. *Regions of Russia*. Moscow: Goskomstat.

Heckman, James, Jora, Stixrud, and Sergio Urzua, 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior, *Journal of Labor Economics* 24(3): 411-482.

Hongbin, Li, Lingsheng Meng, Xinzheng Shi, and Binzhen Wu, 2011. Does Attending Elite Colleges Pay in China? Available at SSRN: http://ssrn.com/abstract=1849523

Ichino, Andrea, and Rudolf Winter-Ebmer, 2004. The Long-Run Educational Cost of World War II, *Journal of Labor Economics* 22(1): 57-87.

James, Estelle, Nabeel Alsalam, Joseph C. Conaty, and Duc-Le To, 1989. College Quality and Future Earnings: Where Should You Send Your Child to College? The American Economic Review 79(2): 247-252.

Light, Audrey, and Wayne Strayer, 2000. Determinants of College Completion: School Quality or Student Ability, *Journal of Human Resources* 25(2): 299-355.

Long, Mark, 2008. College Quality and Early Adult Outcomes, *Economics of Education Review* 27(5): 588-602.

Long, Mark, 2010. Changes in the Returns to Education and College Quality, *Economics of Education Review* 29(3): 338-347.

Loury, Linda Datcher, and David Garman, 1995. College Selectivity and Earnings, *Journal of Labor Economics* 13(2): 289-308.

Maurin, Eric, and Theodora Xenogiani, 2007. Demand for Education and Labor Market Outcomes: Lessons from the Abolition of Compulsory Conscription in France, *Journal of Human Resources* 42(4): 795-819.

Monitoring of Higher Professional Education, 2012. Moscow: Ministry of Education and Science of Russian Federation.

Monks, James, 2000. The Returns to Individual and College Characteristics: Evidence from the National Longitudinal Survey of Youth, *Economics of Education Review* 19(3): 279-289.

OECD, 1999. *Reviews of National Policies for Education: Tertiary Education and Research in the Russian Federation*. Paris, France: Organization for Economic Co-operation and Development.

OECD, 2012. *Education at a Glance: OECD Indicators 2012. Country Note: Russian Federation*. OECD Publishing.

Ofer, Gur, 1987. Soviet Economic Growth: 1928-1985, *Journal of Economic Literature* 25(4): 1767-1833.

Oreopoulos, Philip, 2006. Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter, *American Economic Review* 96(1): 152-175.

Rosefielde, Steven, 2003. The Riddle of Post-War Russian Economic Growth: Statistics Lied and Were Misconstrued, *Europe-Asia Studies* 55(3): 469-481.

Rosen, Seymour, 1980. *Education in the USSR. Current Status of Higher Education*. Washington, DC: U.S. Government Printing Office.

Russian Census, 2010. Moscow: Goskomstat.

Russian State Standard Resolution, 1995.
Available at http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=LAW;n=138471

Russian State Standard Resolution, 2003.
Available at http://zakonprost.ru/content/base/part/744053/

Russian Statistical Yearbook, 2012. Moscow: Goskomstat.

Schultz, Theodore, 1975. The Value of the Ability to Deal with Disequilibria, *Journal of Economic Literature* 13(3): 827-846.

Solomon, Lewis, 1975. The Definition of College Quality and Its Impact on Earnings, *Exploration in Economics Research* 2: 537-587.

Soviet Law, 1943. "O razmerah i poryadke naznacheniya stipendiy v vysshih uchebnyh zavedeniyah i tehnikumah i ob osvobozhdenii studentov ot prizyva v Krasnuyu Armiyu[On size and order of scholarships in higher educational institutions and technical schools and on exemption of students from recruiting in Red Army] September 15, 1943. Available at http://istmat.info/node/20313.

Soviet Law, 1958. "Ob ukreplenii svyazi shkoly s zhizhn'yu i o dal'neyshem razvitii sistemy narodnogo obrazovaniya v SSSR" [On strengthening the relationship between the school and life and on further development of the system of public education in the USSR], December 24, 1958. Available at http://www.libussr.ru/doc_ussr/usr_5337.htm.

Soviet Law, 1964. "O povyshenii zarabotnoy platy rabotnikov prosvescheniya, zdravoohraneniya, zhilischno-kommunal'nogo hozyaystva, torgovli i obschestvennogo pitaniya i drugih otrasley narodnogo hozyaystva, neposredstvenno obsluzhivayuschih naselenie" [On the increase of wages of employees in education, health care, housing and communal services, trade and public catering and other service sectors of the national economy], July 15, 1964.
Available at http://www.libussr.ru/doc_ussr/usr_6110.htm.

Soviet Law, 1967. "O vseobschey voinskoy obyazannosti" [On compulsory military service], October 12, 1967. Available at http://www.libussr.ru/doc_ussr/usr_6676.htm.

Soviet Law, 1974. "O narodnom obrazovanii" [On public education], August 2, 1974. Available at http://www.consultant.ru/online/base/?req=doc;base=ESU;n=2054

Soviet Law, 1989. "O predostavlenii otsrochki ot prizyva na deystvitel'nuyu voennuyu sluzhbu studentam dnevnyh (ochnyh) vysshyh uchebnyh zavedeniy i uchaschimsya dnevnyh (ochnyh) srednih spetsial'nyh uchebnyh zavedeniy, srokah i poryadke prohozhdeniya voinskoy sluzhby litsami, imeyushchimi vysshee obrazovanie" [On the granting of deferment from active duty to students of day-time higher education institutions and students of day-time secondary professional schools, terms and procedure of military service for individuals with higher education], March 31, 1989. Available at http://www.libussr.ru/doc_ussr/usr_15534.htm.

UNESCO, 2009. *Global Education Digest 2009. Comparing Education Statistics Across the World,* Montreal, Canada: UNESCO Institute for Statistics.

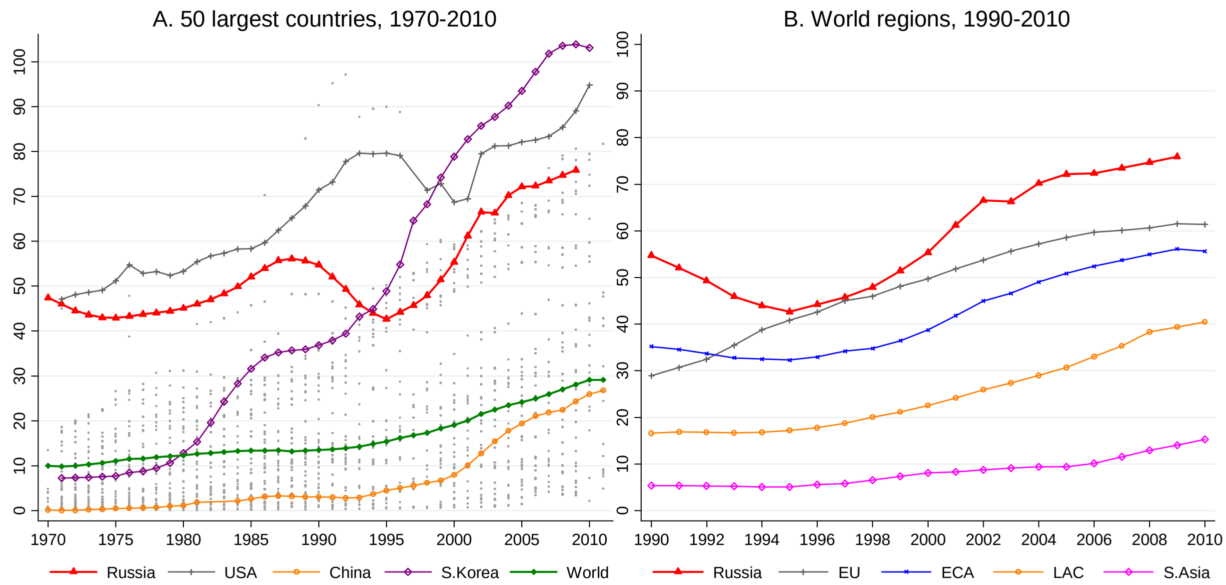
USSR Census, 1989. Moscow: Goskomstat.

Wooldridge, Jeffrey, 2002. *Econometric Analysis of Cross-Section and Panel Data*, Cambridge, MA: MIT Press.

Wooldridge, Jeffrey, 2007. Inverse Probability Weighted Estimation for General Missing Data Problems, *Journal of Econometrics* 141(2): 1281-1301.

Zhang, Liang, 2005. *Does Quality Pay? Benefits of Attending a High-Cost, Prestigious College*. New York: Routledge.

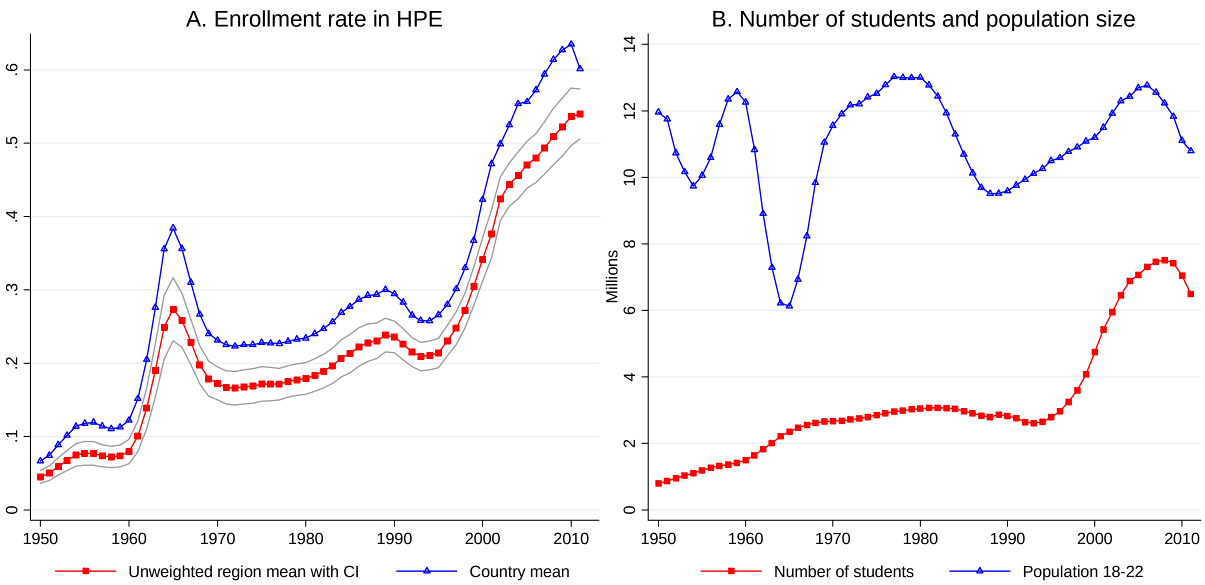
Figure 1: Enrollment Rates in Tertiary Education: International Comparison



Notes: This figure displays enrollment rates in tertiary education (ISCED 5 and 6) in Russia and selected countries and world regions. Panel A also shows the scatterplot of enrollment rates in world's 50 most populous countries as well as the world's average trend (green line) from 1970 to 2010. Panel B shows enrollment trends in selected world regions from 1990 to 2010. EU=European Union; ECA=Eastern Europe and Central Asia; LAC=Latin America and Caribbean. Tertiary education includes both higher education institutions such as universities and post-secondary schools such as community colleges in U.S. and secondary professional schools (*technicums*) in Russia. Enrollment rates are expressed as a percentage of the total population of the five-year age group following on from secondary school leaving.

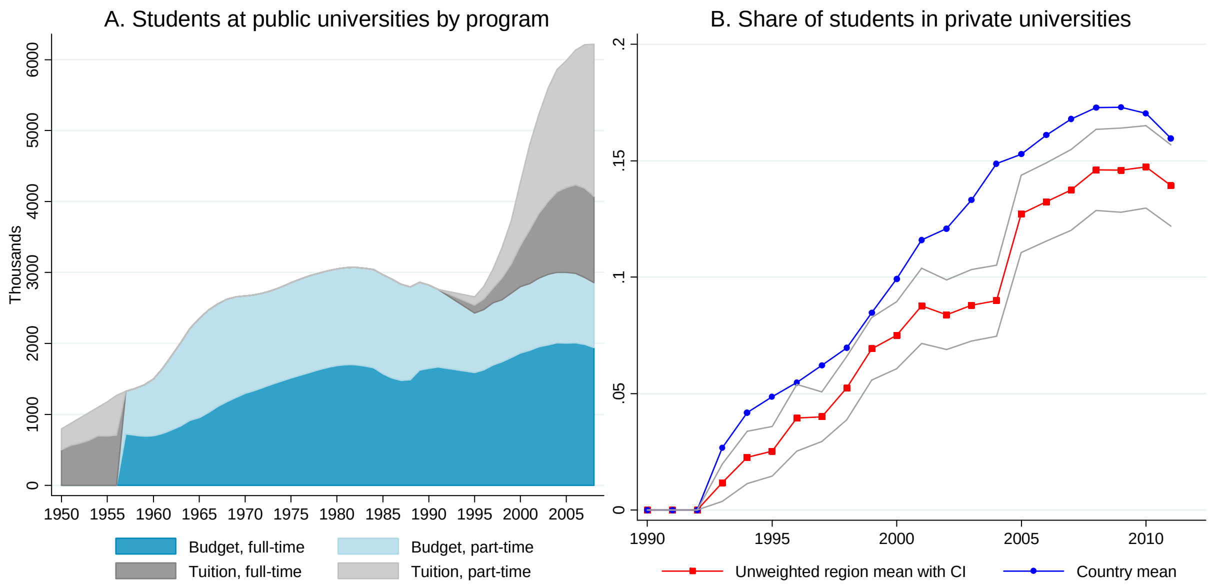
Source: World Bank Education Statistics.

Figure 2: Expansion of Higher Education in Russia



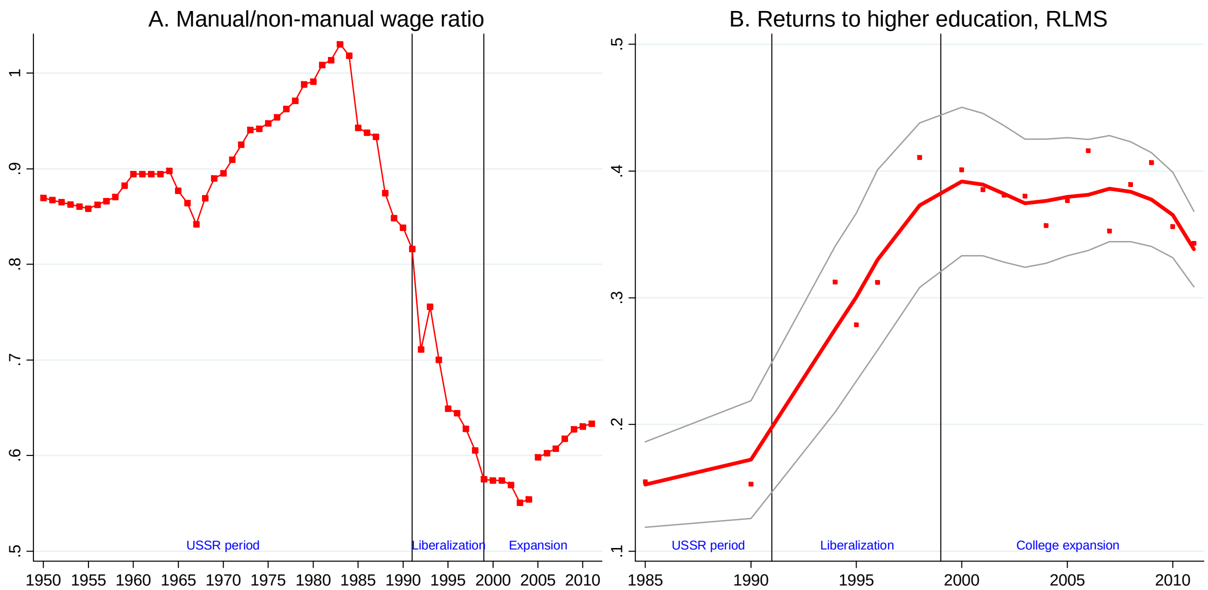
Notes: HPE stands for higher professional education, which is part of tertiary education. Panel A depicts the average region-cohort enrollment rate in higher education. The enrollment rate is calculated as the number of students in HPE institutions divided by the college age population 18-22 in a given year and region. The regional mean is calculated across 79 Russian regions and presented with the 95 percent confidence interval. The regional mean (red line) deviates from the country mean (blue line) due to less populated regions having lower enrollment rates. Panel B reports the total number of students in public and private HPE institutions and the size of college age population (18-22). The spike in enrollment in the 1960s is associated with a sharp drop in the size of college age cohort, while an increase in enrollment in the 2000s is mainly caused by the increasing number of students.

Figure 3: Compositional Shifts



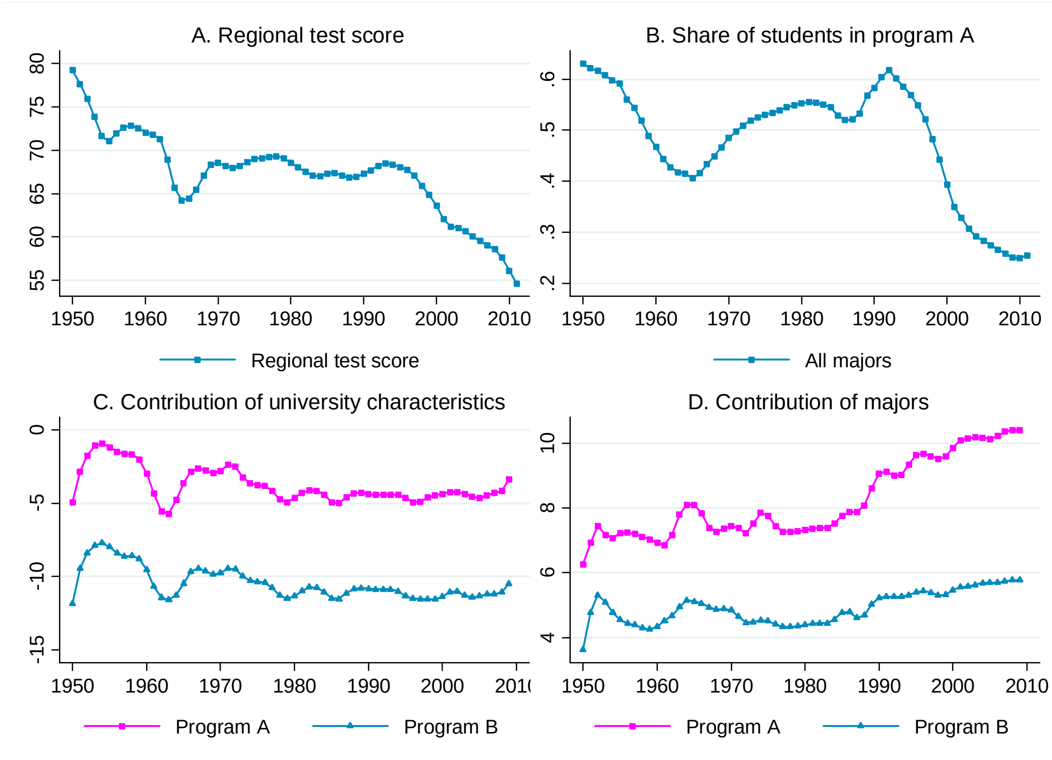
Notes: Panel A depicts the number of students in full-time and part-time programs broken down by the type of financing (budget vs. tuition) at public universities in Russia in 1950-2008. Part-time programs include evening and correspondence programs. The tuition was charged in 1940-1956. Panel B reports the share of students in private universities. CI is a 95 percent confidence interval around regional mean. The difference between national and regional averages is due to regions with higher number of students having a higher share of students in private universities (e.g., Moscow).

Figure 4: Skill Wage Premium



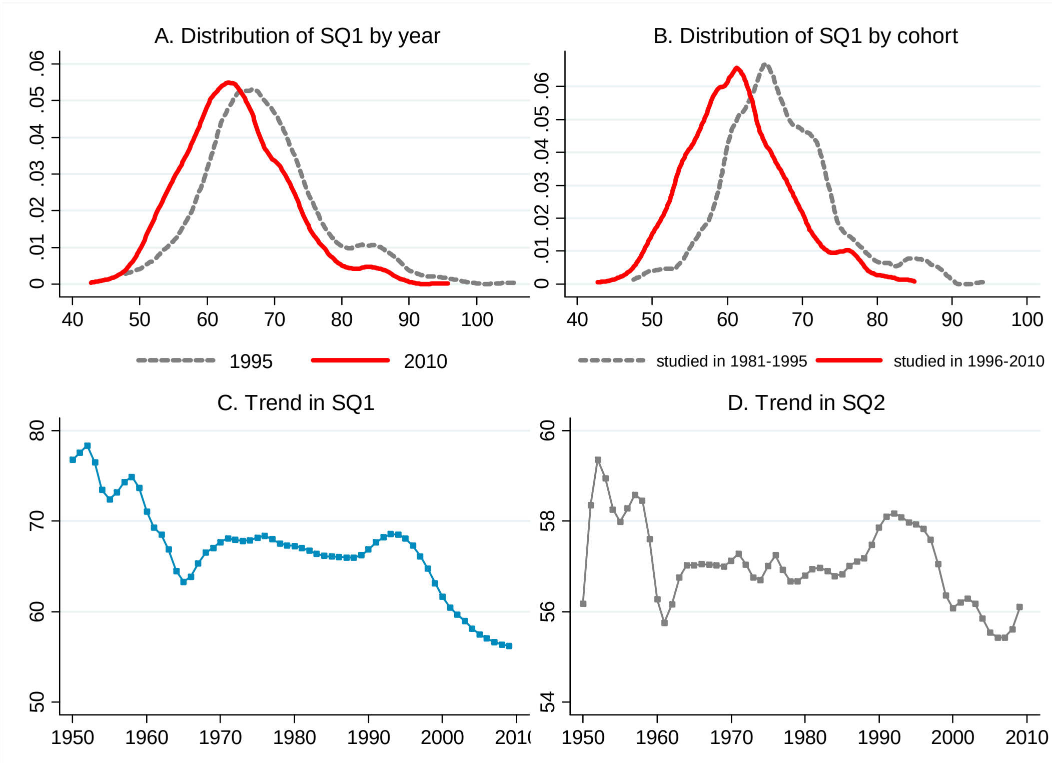
Notes: In Panel A, the wage ratio is the ratio of wages of manual workers to wages of non-manual workers. The wage ratio is calculated for the industrial sector (manufacturing + mining + electricity + selected industrial services) in 1950-2004 and for the manufacturing firms only in 2004-2011. In Panel B, returns to higher education are calculated from the OLS regression of the log of monthly earnings at primary job on college degree, gender, age, age squared, place of birth, a dummy for Russian nationality, and 7 federal districts. Estimation is performed for each year separately using the sample of adults (25 to 55 years old). Reported are the estimated coefficients on college degree and the overall trend fitted using non-parametric smoothing (lowess; bandwidth=0.4). Also shown is the 95% confidence interval computed using robust standard errors. The vertical lines demark 1991 as the breakup of USSR and 1999 as the year when the number of college graduates significantly exceeded the 1980s level.

Figure 5: Elements of the Aggregate Index of Student Quality



Notes: Each panel depicts one of the four elements of the aggregate index of student quality according to equation (4). Program A is a full-time budget program at public HPE institutions. Program B represents part-time and/or tuition programs at public and private HPE institutions. Panel A of Figure 5 reports the average regional test score, which is estimated based on observed region-cohort enrollment rates (see Appendix A2 for the details of calculations). Panel B depicts the share of students in program A. Panel C reports the deviation of university-major test scores from the regional score due to university characteristics by program. Panel D depicts the deviation of university-major test scores from the regional score due to different majors by program. The contributions of university characteristics and majors in Panels C and D are predicted based on estimates in Table 2 according to equation (3). All estimates in Figure 5 are obtained for the RLMS sample of college graduates with non-missing information on major and university location.

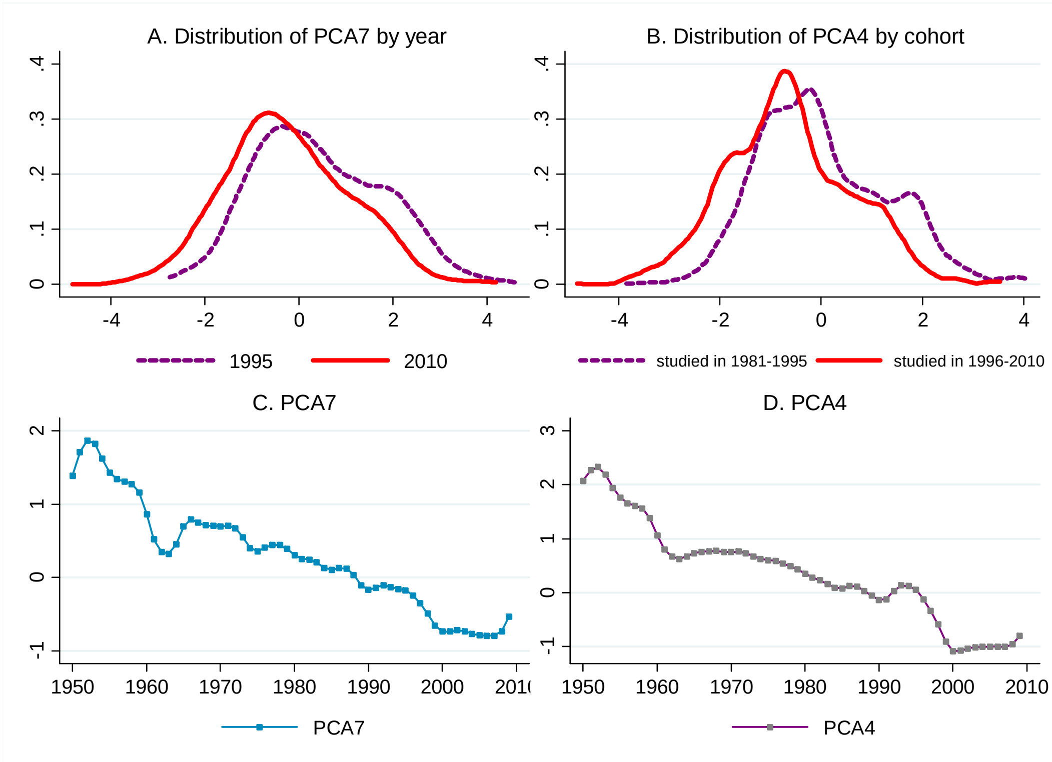
Figure 6: Aggregate Index of Student Quality



SQ1: mean=65.02, sd=8.59, min=40.84, max=105.92;
SQ2: mean=56.78, sd=6.72, min=38.35, max=78.59.

Notes: The aggregate index of student quality, *SQ*1, is constructed according to equation (4) based on the four elements shown in Figure 5. Panel A depicts the kernel densities of *SQ*1 for college graduates age 20-60 surveyed in 1995 and 2010. Panel B plots the kernel densities of *SQ*1 for the cohorts educated in 1981-1995 and 1996-2010 (surveyed in different years). All kernel densities are calculated using the Epanechnikov kernel with the optimal bandwidth. Panel C shows changes in the student quality measure *SQ1* over time. Panel D reports trends in *SQ2*, which is a linear prediction of EGE test scores based on university characteristics, major and program of study using the estimates of column 3 and 4 in Table 2; it does not account for changes in the enrollment rate. All estimates in Figure 6 are obtained for the RLMS sample of college graduates with non-missing information on major and university location.
Summary statistics on measures of quality:

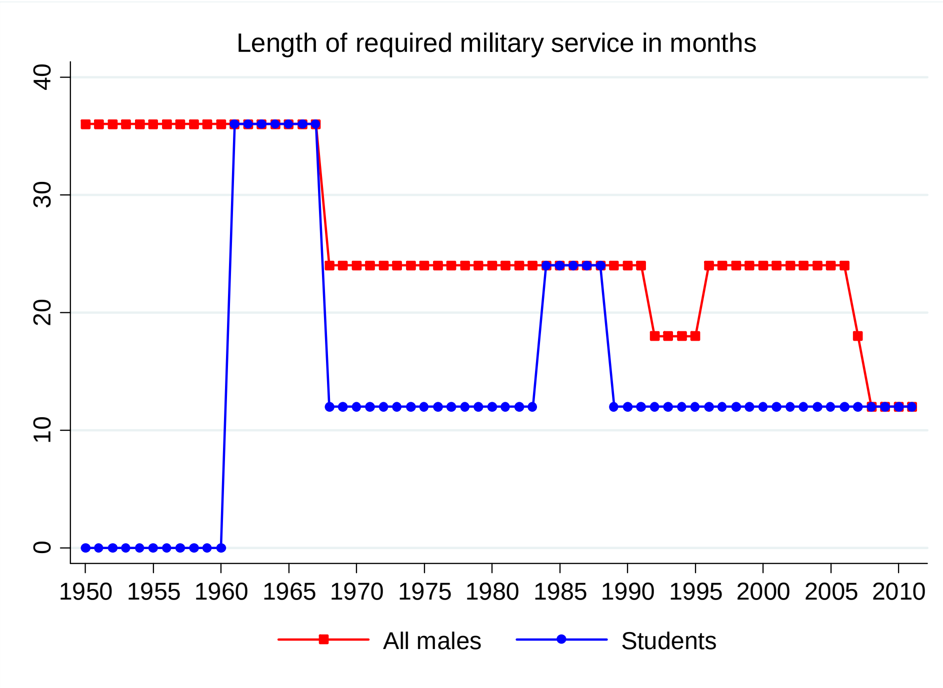
Figure 7: Composite Index of College Quality



PCA4: mean=-7.10e-10, sd=1.21, min=-3.44, max=4.91;
PCA7: mean=-8.47e-11, sd=1.41, min=-4.81, max=5.29.

Notes: PCA7 is the first principal component extracted from SQ1, tuition, the predicted share of foreign students, log of predicted educational space, log of predicted institutional revenue, wages in education, and faculty-student ratio. PCA4 is the first principal component extracted from *SQ1*, tuition, wages in education, and faculty-student ratio. Panel A depicts the kernel densities of *PCA7* for college graduates age 20-60 surveyed in 1995 and 2010. Panel B plots the kernel densities of *PCA7* for the cohorts educated in 1981-1995 and 1996-2010 (surveyed in different years). All kernel densities are calculated using the Epanechnikov kernel with the optimal bandwidth. Panel C shows changes in the first principal component *PCA7* over time. Panel D reports trends in *PCA7*4 All estimates in Figure 7 are obtained for the RLMS sample of college graduates with non-missing information on major and university location.

Figure 8: Exclusion Restriction



Notes: The figure depicts the length of required military service for all males and male students in Russia in 1950-2011. Both variables are used as an exclusion restriction in the selection model for college participation.

Figure 9: Price Heterogeneity

Figure 9 shows flattening the price-quality profiles based on OLS and mixed model.

Figure 10: Decomposition

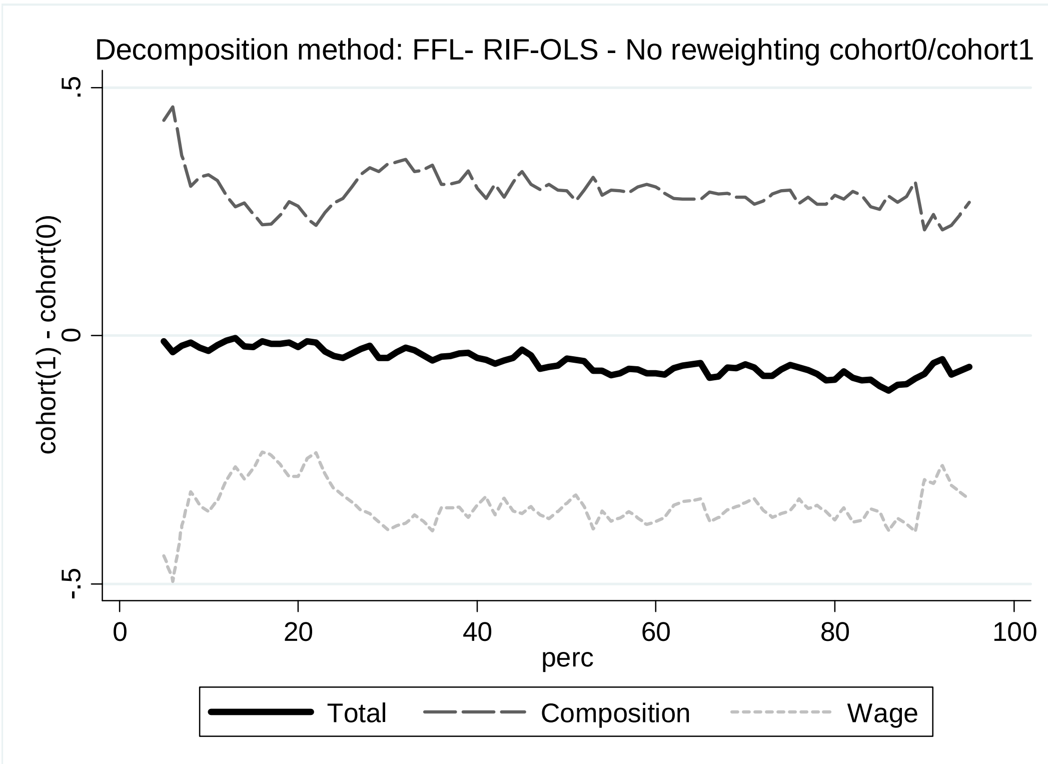


Table 1: Summary Statistics

	1995	2000	2005	2010	1995-2011
	All respondents (N=118,967)				
Female (share)	0.538	0.545	0.549	0.549	0.548
Age	39.24	38.21	38.42	38.87	38.64
	(11.57)	(11.26)	(11.54)	(11.87)	(11.61)
College graduates (share)	0.191	0.185	0.209	0.255	0.219
Years of schooling	11.65	11.90	12.02	12.24	12.04
	(2.54)	(2.21)	(2.19)	(2.21)	(2.24)
Wage (rubles), N=83,913	14.52	9.58	19.05	26.56	20.37
	(17.34)	(9.65)	(16.72)	(22.19)	(19.28)
	College graduates (N=25,991)				
Female (share)	0.558	0.590	0.613	0.627	0.612
Age	40.74	39.97	39.18	38.16	39.04
	(10.42)	(10.25)	(11.08)	(10.93)	(10.86)
Years of schooling	15.15	15.09	15.08	15.05	15.08
	(0.656)	(0.532)	(0.507)	(0.417)	(0.497)
Wage (rubles), N=21,518	19.64	13.15	25.58	35.03	28.22
	(22.44)	(12.67)	(20.41)	(26.99)	(24.75)
	College graduates with university characteristics (N=21,520)				
Female (share)	0.666	0.631	0.611	0.626	0.630
Age	40.86	40.45	39.21	38.13	39.11
	(10.16)	(10.17)	(11.05)	(10.92)	(10.82)
Years of schooling	15.08	15.07	15.08	15.06	15.07
	(0.510)	(.472)	(0.501)	(0.421)	(0.468)
Wage (rubles); N=18,042	16.02	11.86	25.67	34.86	28.60
	(16.86)	(10.11)	(20.45)	(26.39)	(24.67)
Selected majors					
Economics and business	10.93	11.24	17.32	21.26	17.50
Health care	9.74	8.70	5.68	4.82	5.85
Pedagogy	5.77	6.16	5.54	4.67	(5.37)
Machine building	5.96	5.76	5.20	5.01	5.14
Law	2.58	3.08	5.54	6.34	5.33
Private universities (share)	0.002	0.005	0.036	0.074	0.044
Local branches (share)	0.089	0.097	0.117	0.151	0.124
Non-central cities (share)	0.111	0.112	0.115	0.162	0.134
Selected types of university					
Humanities	2.78	2.81	4.79	6.15	4.86
Classical	17.30	20.21	17.52	16.43	17.70
Medical	9.54	8.17	5.48	4.57	5.56
Pedagogy	16.10	18.34	15.33	15.74	16.17
Economic	7.36	6.96	10.20	11.98	10.16
Engineering	30.02	27.44	29.91	28.29	28.71
Law	1.59	1.87	2.19	2.43	2.26

Notes: The sample consists of individuals aged 20-60. All figures in the table represent sample means. Estimation sample includes individuals with non-missing values of explanatory variables. The sample composition is unweighted.

Table 2: Determinants of EGE, 2010-2012

	Mean	Deviation		Average EGE	
	(SD)	Program A	Program B	Program A	Program B
Public university	0.984	...	3.838***	...	3.851***
	(0.124)		(0.423)		(0.423)
Location type (omitted: Moscow city)					
Central city-main campus	0.685	-1.044	-2.162***	-0.789	-1.972***
	(0.464)	(0.731)	(0.666)	(0.719)	(0.646)
Central city-branch	0.036	-1.222	-1.534*	-1.087	-1.458*
	(0.187)	(0.875)	(0.800)	(0.862)	(0.781)
Small city-main campus	0.097	-4.993***	-4.337***	-4.718***	-4.162***
	(0.297)	(0.733)	(0.688)	(0.721)	(0.668)
Small city-branch	0.053	-4.409***	-4.264***	-4.305***	-4.225***
	(0.224)	(0.757)	(0.674)	(0.741)	(0.652)
University has branches	0.651	0.617***	-0.024	0.591***	-0.017
	(0.476)	(0.143)	(0.163)	(0.143)	(0.162)
Founding date (omitted: founded after 2000)					
Before 1917	0.157	5.807***	2.953***	5.757***	2.798***
	(0.364)	(0.480)	(0.499)	(0.483)	(0.496)
1918-1930	0.216	4.322***	2.314***	4.321***	2.182***
	(0.411)	(0.465)	(0.479)	(0.468)	(0.476)
1931-1950	0.225	4.499***	2.423***	4.485***	2.334***
	(0.417)	(0.466)	(0.479)	(0.469)	(0.476)
1951-1960	0.128	4.114***	1.926***	4.087***	1.864***
	(0.335)	(0.463)	(0.478)	(0.466)	(0.475)
1961-1970	0.059	3.252***	1.738***	3.161***	1.632***
	(0.237)	(0.497)	(0.503)	(0.498)	(0.499)
1971-1980	0.056	3.564***	1.711***	3.515***	1.616***
	(0.230)	(0.495)	(0.511)	(0.499)	(0.507)
1981-1990	0.018	0.814	0.282	0.773	0.295
	(0.133)	(0.643)	(0.664)	(0.644)	(0.660)
1991-2000	0.120	1.611***	1.758***	1.650***	1.761***
	(0.325)	(0.460)	(0.457)	(0.463)	(0.454)
University specialty (omitted: agriculture)					
Architecture	0.021	0.530	1.367***	0.533	1.252***
	(0.145)	(0.391)	(0.352)	(0.388)	(0.346)
Humanities	0.050	2.408***	2.214***	2.488***	2.149***
	(0.218)	(0.367)	(0.331)	(0.367)	(0.331)
Classical	0.342	4.293***	3.082***	4.300***	3.003***
	(0.474)	(0.233)	(0.208)	(0.232)	(0.208)
Medicine	0.017	2.223***	3.437***	2.270***	3.270***
	(0.130)	(0.504)	(0.501)	(0.500)	(0.501)
Pedagogy	0.082	0.632**	0.495**	0.688**	0.523**
	(0.275)	(0.281)	(0.264)	(0.280)	(0.264)
Economy	0.085	4.544***	3.631***	4.637***	3.486***
	(0.280)	(0.311)	(0.299)	(0.310)	(0.298)
Engineering	0.308	2.983***	1.393***	3.027***	1.279***

	(0.461)	(0.228)	(0.203)	(0.228)	(0.203)
Governance	0.018	4.249***	3.295***	4.417***	3.186***
	(0.134)	(0.629)	(0.477)	(0.624)	(0.476)
Law	0.003	6.073***	4.088***	6.046***	3.890***
	(0.057)	(0.984)	(0.700)	(1.019)	(0.700)
Military	0.002	1.800**	1.241	1.898**	0.928
	(0.052)	(0.980)	(1.931)	(0.935)	(1.849)
Intercept		-1.304**	-8.992***	50.447***	43.014***
		(0.547)	(0.804)	(0.548)	(0.802)
Field major (57 dummies)		Yes	Yes	Yes	Yes
University region (78 dummies)		Yes	Yes	Yes	Yes
Year effects		Yes	Yes	Yes	Yes
Number of observations	27,171	17,408	9,649	17,411	9,652
R-squared		0.715	0.681	0.616	0.510

Notes: Estimation sample includes universities with non-missing variables. Robust standard errors are in parentheses.

Table 3: Principal Component Analysis

	Factor #	Eigenvalue	Proportion
PCA 4			
	1	1.477	0.369
	2	1.164	0.291
	3	0.869	0.217
	4	0.488	0.122
PCA 7			
	1	2.008	0.286
	2	1.395	0.199
	3	1.143	0.163
	4	0.995	0.142
	5	0.575	0.082
	6	0.476	0.068
	7	0.405	0.057

Notes: (a) PCA4 is constructed based on four different measures of quality. An aggregate index of student quality (SQ1) is computed according to equation (4). A proxy for tuition (SQ2) is predicted based on the estimates presented in Table 2. A ratio of wages in education to average wages in the economy is a country-level average by year. A faculty-student ratio in HPE is calculated as the number of faculty in HPE divided by the total number of students at public universities in a given year. (b) PCA7 is constructed based on seven different measures of quality, four measures of quality used in PCA4 and three additional measures. Share of foreign students is predicted based on the estimates in Table S1. Log of instructional space per student in square meters is predicted based on the estimates in Table S1. Revenue per faculty in thousand rubles is predicted based on the estimates in Table S1.

Table 4: Estimates of the Returns to College Quality: OLS, Heckman, and IPW

	OLS	Heckman Probit	Heckman Main	IPW1	Probit for IPW2	IPW2
			Panel A			
*SQ*1	0.004***		0.005***	0.005***		0.006***
	(0.001)		(0.001)	(0.000)		(0.000)
Female	-0.285***	2.176***	-0.272***	-0.268***	1.777***	-0.264***
	(0.010)	(0.116)	(0.010)	(0.009)	(0.120)	(0.009)
Experience	0.029***	-0.124***	0.027***	0.026***	-0.089***	0.025***
	(0.002)	(0.005)	(0.002)	(0.001)	(0.005)	(0.001)
Experience squared	-0.080***	0.113***	-0.083***	-0.071***	0.043***	-0.071***
	(0.004)	(0.008)	(0.004)	(0.004)	(0.008)	(0.004)
Urban	0.324***	0.486***	0.351***	0.304***	0.400***	0.300***
	(0.014)	(0.037)	(0.014)	(0.011)	(0.036)	(0.011)
Log of entry size	-0.051*	-1.672***	-0.131***	-0.052	-1.493***	-0.087
	(0.026)	(0.056)	(0.029)	(0.029)	(0.054)	(0.029)
Wage ratio		2.975***			2.542***	
		(0.151)			(0.159)	
Cohort competition		-0.298***			-0.316***	
		(0.093)			(0.099)	
Army service - all		-0.012***			-0.012***	
		(0.003)			(0.003)	
Army service- students		0.087***			0.073***	
		(0.005)			(0.005)	
F-test for excl. restr.						
District effects	Yes	Yes	Yes	Yes	Yes	Yes
Year effect	Yes	Yes	Yes	Yes	Yes	Yes
N of observations	18,034	83,577	18,034	18,034	118,403	17,996
R-squared	0.378		0.378	0.400		0.411
			Panel B			
*SQ*1	0.066***		0.067***	0.056***		0.052***
	(0.006)		(0.008)	(0.006)		(0.006)
*SQ*1 squared	-0.045***		-0.045***	-0.037***		-0.033***
	(0.005)		(0.006)	(0.004)		(0.004)

Notes: Exclusion restrictions are (1) ratio of average wages of manual workers to average wages of non-manual workers in the economy at age 17, (2) a dummy variable for "high" and "low" competition for a university slot (= 1 for cohorts born between September 1942 and August 1945; = -1 for cohorts born between September 1946 and August 1948; = 0 otherwise), (3) length of mandatory service in the army for men at age 17, (4) length of mandatory service in the army for college students at age 17. Bootstrapped standard errors are in parentheses.

Table 5: Returns to College Quality Before and After Expansion

	Baseline			Extended		
	All	Before Expansion	During Expansion	All	Before Expansion	During Expansion
*SQ*1	0.052***	0.135***	0.104***	0.080***	0.160***	0.105***
	(0.006)	(0.016)	(0.014)	(0.008)	(0.018)	(0.016)
*SQ*1 squared	-0.033**	-0.095***	-0.077***	-0.054**	-0.110***	-0.078***
	(0.004)	(0.011)	(0.011)	(0.005)	(0.013)	(0.012)
SQ2	0.104***	0.123***	0.151***	0.125***	0.161***	0.176***
	(0.007)	(0.015)	(0.018)	(0.009)	(0.018)	(0.021)
*SQ*2 squared	-0.081***	-0.099***	-0.121***	-0.098***	-0.129***	-0.144***
	(0.006)	(0.013)	(0.016)	(0.007)	(0.015)	(0.018)
PC4	0.102***	0.091***	0.065***	0.111***	0.106***	0.066***
	(0.005)	(0.012)	(0.013)	(0.006)	(0.013)	(0.015)
*PC*4 squared	-1.694**	-4.668***	-3.169**	-1.910**	-5.590***	-3.261**
	(0.291)	(0.873)	(0.678)	(0.329)	(0.955)	(0.757)
PC7	0.064***	0.057***	0.100***	0.076***	0.062***	0.115***
	(0.004)	(0.009)	(0.008)	(0.004)	(0.010)	(0.009)
*PC*7 squared	-0.924*	-2.084**	0.703	-0.775	-2.957***	1.716**
	(0.212)	(0.510)	(0.384)	(0.223)	(0.569)	(0.426)

Notes: Specification of IPW2 is the same as in Table 4. "Before expansion" is estimation for a cohort of individuals who graduated from college in 1986-1998. "During expansion" is estimation for a cohort of individuals who graduated from college in 1999-2011. Extended specification includes parents' education among the covariates. Bootstrapped standard errors are in parentheses.

Table 6: Expanded list of covariates

- Soviet vs. market
- Baseline vs. extended

Table 7: Panel mixed model

	SQ1	SQ2	PCA4	PCA7
Quality	0.005***	0.009***	0.091***	0.069***
	(0.001)	(0.001)	(0.009)	(0.007)
Female	-0.291***	-0.303***	-0.298***	-0.273***
	(0.016)	(0.016)	(0.015)	(0.015)
Experience	0.030***	0.031***	0.027***	0.028***
	(0.002)	(0.002)	(0.002)	(0.002)
Experience	-0.085***	-0.086***	-0.084***	-0.085***
Squared	(0.005)	(0.005)	(0.005)	(0.005)
Urban	0.334***	0.327***	0.310***	0.302***
	(0.024)	(0.024)	(0.024)	(0.024)
Log of entry size	-0.076*	-0.104***	-0.087**	-0.117***
	(0.039)	(0.037)	(0.036)	(0.036)
District effects	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
N of observations	18,034	18,034	18,034	18,034

Report covariances.

Table 8: OLS regression of returns to quality on college enrollment

Enrollment	0.412***
	(0.143)
Cohort	0.052***
	(0.012)
Period	0.035***
	(0.008)
Constant	-0.331***
	(0.038)
N of observations	170
R squared	0.428

Notes: Enrollment is actual enrollment in higher education at age 17; cohort=1 if 17-year old in 1950-1959, =2 if 17-year old in 1960-1969, =3 if 17-year old in 1970-1979, =4 if 17-year old in 1980-1989, =5 if 17-year old in 1990-1999, =6 if 17-year old in 2000-2010; period=1 if observed in the labor market in the 1985-1990, =2 of observed in the labor market in the 1995-1998,=3 if observed in the labor market in the 2000-2010.

Appendix A1: Description of Variables

Variable name	Definition and sources
	A. Individual-Level Variables

General notes:

1. *The source for all individual-level variables is RLMS*
2. *Estimation sample covers 1995-1996, 1998, 2000-2011 time periods; variables are available for all years, unless noted otherwise*
3. *The sample is restricted to individuals aged 20-60 at the time of the survey.*

Female	=1 if female
Age	Year of survey minus year of birth; the mode of birth year is computed in cases of inconsistencies across rounds
College major	58 categories; coded by authors using the 2004 Russian Classification of Fields of Study; see web appendix W2 for details on coding of majors
College degree	=1 if received degree of higher professional education
Full-time program	=1 if studied in a full-time program; 0 if studied in correspondence (distant learning) or part-time evening program; 2009-2011
University name	Identified based on question "write down names of higher educational institutions attended"; available in 2004, 2005, 2010, 2011; extrapolated to other years based on individual panel id for cases of consistent reporting of college major
Adjusted years of schooling	Assigned years of schooling according to the highest level of education attained: the grade level in secondary general school +1.5 years if completed vocational technical school; 13 years for secondary professional education; 15 years for higher professional education; and 18 years for post-graduate degree
Actual years of schooling	Sum of years of schooling at all levels of education, including multiple degrees at the same level
Urban residence	= 1 if resides in urban location at the time of survey
Urban place of birth	=1 if born in urban location
Russian	=1 if ethnicity is Russian
Mother's college	=1 if mother has a college degree
Father's college	=1 if father has a college degree
Wage	Usual earnings (average monthly earnings in the last year)
Hours of work	Usual hours of work (usual hours of work per week x 4)
Tenure	Number of years worked for a current employer

B. University-Level Variables

General notes:

4. *Sources are common to all variables and include Higher School of Economics Monitoring of Quality (HSE EGE database; http://www.hse.ru/ege); Federal Portal "Russian Education" (www.edu.ru); the 2012 Efficiency Assessment Report by the Ministry of Education and Science; University websites; other Internet resources.*

5. *University-level variables are linked to each RLMS respondent based on the name and location of campus provided in 2004-2005 and 2010-2011 surveys and college major provided in 1998-2011 surveys.*
6. *Categories of college majors in RLMS are made consistent with the HSE EGE database to ensure accurate linking.*

Public university	=1 if university is publicly-owned; 0 otherwise (private and NGOs)
Main campus	=1 if main campus; 0 if satellite campus or branch
University with branches	=1 if university has branches in other locations
University location	3 categories: Moscow city, other central city (i.e., regional capital that is also the largest city in a region); non-central city
University region	81 subjects of the Russian Federation or regions (Moscow and Leningrad regions include Moscow city and St. Petersburg, correspondingly)
University type	Classical university or one of the 10 specialized schools: agriculture, architecture and construction, humanities, medicine, pedagogy and sports, economy incl. trade and services, polytechnic, governance incl. former schools of communist party, law, and military
University founding date	Year when the university is officially established as an institution of higher education for main campuses or year when a satellite campus is opened at a given location for branches
Freshmen enrollment, n	Number of enrolled freshmen by disaggregated college major, program of study, and year, 2010-2012
Program of study A	Publicly-financed full-time curricula at public universities
Program of study B	Tuition-paid full-time curricula at both public and private universities and correspondence curricula (the latter is added by the authors based on admission lists)
Average EGE	Average test score per subject from the unified state exam (EGE) for enrolled freshmen by college major, program of study and year, 2010-2012; weighted by n
Tuition	Annual real tuition for tuition-paid programs by college major and year, 2010-2012; deflated using annual national CPI (2010=1); weighted by n
$SQ1$, $SQ2$, and $TQ2$	Predicted test scores/tuition as described in the text
Revenues	Total revenue from all sources per faculty, 2012
Space	Total area of educational and laboratory buildings per student available to university on the right of ownership or operational use, 2012
Foreign students	Share of foreign students in the total number of students, 2012

C. Region-Level Variables

General notes:

1. *Russian regions include two federal cities (Moscow city and St. Petersburg), and 81territories.*
2. *Due to multiple changes in the administrative-territorial structure of Russia, all past regional data are collected based on the most recent classification of regions according to the 2008 amendment to the Constitution of the Russian Federation.*
3. *At the time of the survey, respondents resided in 32 regions, but they graduated from universities located in 73 regions and all 15 former USSR republics.*
4. *Region-level quality measures are linked to RLMS respondents using the region of their university and the year when they turned 20-years old (medium college age). Thus, 1955 is the first year when aggregate quality measures are linked to the RLMS sample for a 60-year old individual surveyed in 1995.*

5. *If respondents graduated from other former republics of the USSR, we linked indicators of the Russian region that is closest in terms of geographic distance and educational attainment based on the 1989 USSR census (e.g., Kaliningrad region is chosen for Baltic republics, Kaluga region is for Belarus, etc.). Results do not change if these observations are dropped.*
6. *Education industry is defined based on the Soviet classification of industries OKONH in 1950-2001 and the International Standard Industry Classification ISIC in 2002-2010; education industry includes culture prior to 1976.*
7. *Letter "R" denotes time periods when data are available by region; letter "C" indicates if country-level cohort-specific data are used.*

Population age 18-22	Number of people age 18-22, R1950-2010 Sources: Census 1959, 1970, 1979, 1989, 2002; Demoscope Weekly (www.demoscope.ru/); Goskomstat Central Statistical Database (GCSD); values in 1950-1958 are imputed using the age distribution by region from the 1959 Census and age-specific mortality coefficients for Russia available at http://www.mortality.org/
Cohort size	Number of people age 17, R1950-2010 Sources: see above
Number of students, *S*	Total number of students at higher educational establishments, R1950-2010 Sources: Russian yearbooks (annual issues from 1956 to 2011); GCSD *Notes: missing values for 1951-55 are imputed using the linear interpolation between the two available data points.*
Enrollment rate	Number of students divided by population age 18-22, R1950-2010
Construction in HPE	Construction of education buildings and labs in HPE, 3-year moving average, in square meters per students, C1950-1989, R1990-2010 Sources: Russian yearbooks (annual issues from 1990 to 2011); GCSD *Notes: missing values for 1950-69 are imputed using the lagged volume of construction in secondary education (R-squared for prediction=0.93).*
Faculty-student ratio	Number of instructors and professors at public universities divided by the number of students at public universities, C1950-2010, R1990, R1995, R2000-2003, R2009 Sources: Education in Russia (2003, 2008); GCSD *Notes: prior to 1970, the faculty-student ratio was legally set at 0.6.*
Relative wage in education	Average accrued monthly wage in education industry divided by the national average accrued monthly wage in all sectors; C1950-1989, R1990-2010; Sources: GCSD; Russian yearbooks (annual issues from 1960 to 2011) *Notes: Linear interpolation is used to impute missing values in 1951-54, 56-57, and 61-62.*
IV	

D. Cohort-Level Variables

Wage ratio	Ratio of average wages of manual workers to average wages of non-manual workers in the economy at age 17 Sources: Russian yearbooks (annual issues from 1960 to 2011)
Compulsory schooling laws	= 1 for cohorts born between September 1942 and August 1945; = -1 for cohorts born between September 1946 and August 1948; = 0 otherwise Sources: Soviet Law 1958, Soviet Law 1964
Military service - all	Length of mandatory service in the army for male non-students at age 17

Military service - students	Length of mandatory service in the army for male college students at age 17 Sources: Federal Law 1996, 2006; Gatsko (2008); Soviet Law 1943, 1967, 1989
Approved slots	Log of the number of accepted freshmen at public universities at age 17

Appendix A2: The Relationship between the Average Test Score and Enrollment Rate

We assume that test scores of high school graduates, S, are a continuous random variable with the probability density $f(s)$, and that college enrollment is merit-based and depends upon applicants passing a threshold value, s^*. The population mean of test scores is a weighted average of test scores:

$$E(S) = E(S|S < s^*) \cdot (1 - e) + E(S|S > s^*) \cdot e, \tag{A1}$$

where e is the enrollment rate defined as the ratio of the number of admitted students to the size of entering cohort, and $E(S|S > s^*)$ corresponds to the average test score of freshmen used in the text, $\bar{\theta}_{rc}$.

First, we consider a simple case of the uniform distribution of test scores on a scale from 0 to 100, with $f(s) = \frac{1}{100}$ for $0 < s < 100$ and $\mu = 50$. By substituting $E(S|S < s^*) = \frac{s^*}{2}$ and $E(S|S > s^*) = \frac{s^*+100}{2}$ into equation (A1), we obtain the following formula for threshold s^*:

$$E(S) = \frac{s^*}{2} \cdot (1 - e) + \frac{s^*+100}{2} \cdot e = \frac{s^*}{2} + 50e \quad \rightarrow \quad s^* = 2\mu - 100e. \tag{A2}$$

Therefore, the average test score of students is a negatively sloped linear function of enrollment rate:

$$E(S|S > s^*) = \frac{s^*+100}{2} = \mu - 50e + 50 = 100 - 50e. \tag{A3}$$

In reality, Russian test scores have a truncated normal distribution $S \sim \mathcal{N}(50, 15)$ on a scale from 0 to 100. We can write the expected value of standardized test scores for admitted students, $Z \sim \mathcal{N}(0, 1)$ conditional on $a < Z < b$, as follows:

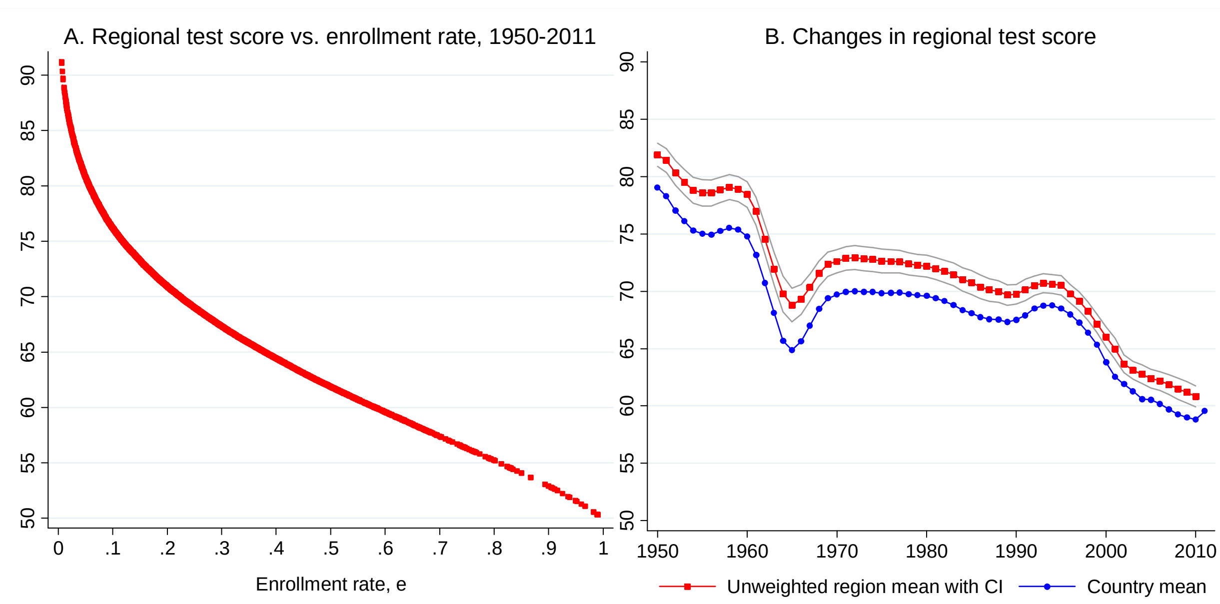
$$E(Z|Z > z^*, Z \in [a, b]) = \frac{\int_{z*}^{b} z\phi(z)dz}{P(Z>z^*)P(a<Z<b)} = \frac{\int_{z*}^{b} z\frac{1}{\sqrt{2\pi}}exp^{-\frac{z^2}{2}}dz}{[\Phi(b)-\Phi(z*)][\Phi(b)-\Phi(a)]}$$

$$= \frac{exp^{-\frac{z*^2}{2}} - exp^{-\frac{b^2}{2}}}{\sqrt{2\pi}[\Phi(b)-\Phi(z*)][\Phi(b)-\Phi(a)]}, \tag{A4}$$

where z^* is a z-score for the minimum test score of a student admitted last, $\phi(z)$ is the pdf of standard normal distribution, $\Phi(\cdot)$ is its cdf, $a \approx -3.33, b \approx 3.33$.

For each of 4530 observations on positive enrollment rates in region r and cohort c, we (i) calculate z_{rc}^* using the inverse of $\Phi(1 - e_{rc})$, (ii) estimate the conditional expectation of standardized scores as in (A4), and (iii) convert conditional standardized means back to raw scores to obtain $\bar{\theta}_{rc}$. We plot the regional-cohort mean test score of admitted students against the enrollment rate in Panel A of Figure A2. The correlation between $\bar{\theta}_{rc}$ and e_{rc} is -0.95. Panel B of Figure A2 shows cross-cohort fluctuations in the imputed regional average test score of admitted students.

Figure A1: Regional Test Scores Imputed Based on Enrollment Rates



Notes: N=4530. Region-cohort enrollments rates are calculated as the number of students at HPE institutions divided by the population of age 18-22 in a given year. Regional test scores are conditional on passing the minimum entry threshold s^*, which is estimated according to equation (A4) based on observed enrollment rates. Panel A shows a scatterplot between the two variables for the entire sample of 4530 regions-cohorts.

Appendix A3: Supplementary Tables

Table S1: Determinants of University Efficiency Indicators

	Tuition	Foreign	Revenue	Space
Public university	0.251***			
	(0.025)			
Location type (omitted: Moscow city)				
Central city-main campus	-0.407***	-0.027***	0.156	-0.482**
	(0.034)	(0.008)	(0.265)	(0.193)
Central city-branch	-0.377***	-0.035***	-0.492	-0.729***
	(0.045)	(0.008)	(0.320)	(0.232)
Small city-main campus	-0.451***	-0.029***	0.129	-0.225
	(0.034)	(0.008)	(0.246)	(0.202)
Small city-branch	-0.405***	-0.035***	-0.815***	-0.707***
	(0.035)	(0.008)	(0.307)	(0.210)
University has branches	0.016**	-0.003	0.047	-0.313***
	(0.008)	(0.004)	(0.077)	(0.081)
Founding date (omitted: founded after 2000)				
Before 1917	0.200***	0.011*	0.561***	1.051***
	(0.030)	(0.006)	(0.198)	(0.193)
1918-1930	0.152***	0.002	0.414**	0.889***
	(0.030)	(0.005)	(0.195)	(0.169)
1931-1950	0.158***	0.006	0.538***	0.874***
	(0.030)	(0.004)	(0.187)	(0.170)
1951-1960	0.159***	0.001	0.483**	0.955***
	(0.030)	(0.003)	(0.211)	(0.158)
1961-1970	0.119***	0.003	0.367*	0.958***
	(0.031)	(0.004)	(0.217)	(0.172)
1971-1980	0.110***	-0.001	0.180	1.237***
	(0.032)	(0.003)	(0.298)	(0.171)
1981-1990	-0.024	-0.008**	0.546**	0.985***
	(0.036)	(0.003)	(0.240)	(0.229)
1991-2000	0.076***	-0.001	0.364**	0.167
	(0.029)	(0.002)	(0.156)	(0.121)
University specialty (omitted: agriculture)				
Architecture	0.254***	0.008	-0.102	-0.103
	(0.019)	(0.005)	(0.380)	(0.246)
Humanities	0.333***	0.010***	0.317**	-0.576***
	(0.017)	(0.003)	(0.155)	(0.131)
Classical	0.307***	0.008***	0.162	-0.220*
	(0.011)	(0.002)	(0.159)	(0.130)
Medicine	0.013	0.050***	0.009	-0.338***
	(0.029)	(0.010)	(0.163)	(0.129)
Pedagogy	0.117***	0.005*	-0.209	-0.679***
	(0.013)	(0.003)	(0.192)	(0.139)
Economy	0.421***	0.008***	0.351**	-0.483***
	(0.017)	(0.003)	(0.172)	(0.145)
Engineering	0.176***	0.010***	0.243*	0.186

	(0.011)	(0.003)	(0.146)	(0.120)
Governance	0.121***	0.009***	0.412*	-0.565***
	(0.033)	(0.003)	(0.221)	(0.190)
Law	0.337***	0.005	0.530**	-0.127
	(0.060)	(0.005)	(0.267)	(0.228)
Military	0.267***	-0.002	1.153*	1.106
	(0.080)	(0.008)	(0.601)	(0.815)
Field of studies	Yes	No	No	No
University region	Yes	Yes	Yes	Yes
Intercept	10.547***	0.036***	6.491***	1.998***
	(0.045)	(0.008)	(0.223)	(0.217)
N of observations	13,322	1,428	1,428	1,428
R-squared	0.535	0.287	0.199	0.348

Notes: 53 field of studies dummies, 82 university region dummies. Robust standard errors are in parentheses.