

Arabsheibani, G. Reza; Mussurov, Altay

**Working Paper**

## Returns to schooling in Kazakhstan: OLS and instrumental variables approach

IZA Discussion Papers, No. 2462

**Provided in Cooperation with:**

IZA – Institute of Labor Economics

*Suggested Citation:* Arabsheibani, G. Reza; Mussurov, Altay (2006) : Returns to schooling in Kazakhstan: OLS and instrumental variables approach, IZA Discussion Papers, No. 2462, Institute for the Study of Labor (IZA), Bonn,  
<https://nbn-resolving.de/urn:nbn:de:101:1-2008060602>

This Version is available at:

<https://hdl.handle.net/10419/33760>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

IZA DP No. 2462

## Returns to Schooling in Kazakhstan: OLS and Instrumental Variables Approach

G. Reza Arabsheibani  
Altay Mussurov

November 2006

# Returns to Schooling in Kazakhstan: OLS and Instrumental Variables Approach

**G. Reza Arabsheibani**

*University of Wales Swansea, WELMERC  
and IZA Bonn*

**Altay Mussurov**

*KIMEP*

Discussion Paper No. 2462  
November 2006

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### Returns to Schooling in Kazakhstan: OLS and Instrumental Variables Approach<sup>\*</sup>

This paper examines rates of return to schooling in Kazakhstan using OLS and instrumental variable (IV) methodologies. We use spouse's education and smoking as instruments. We find that spouse's education is a valid instrument and that conventional OLS estimates that assume the exogenous nature of schooling, and hence do not control for endogeneity bias, may underestimate the true rates of return. The results indicate that the returns to schooling in Kazakhstan have increased with transition. This may reflect the relative scarcities of highly educated people in Kazakhstan with human capital that employers require and, following the market reforms, reward accordingly.

JEL Classification: C13, I21, J24

Keywords: human capital, instrumental variables, rate of return to education

Corresponding author:

G. Reza Arabsheibani  
Department of Economics  
University of Wales Swansea  
Singleton Park  
Swansea, SA2 8PP  
United Kingdom  
E-mail: [G.Arabshuibani@swansea.ac.uk](mailto:G.Arabshuibani@swansea.ac.uk)

---

<sup>\*</sup> We are grateful to Patrick Gunning, Andrew Henley, John Leonard, Alan Marin, Philip Murphy, and John Sessions for their helpful comments and suggestions on previous drafts of this work. The authors thank an anonymous referee for the constructive comments that led to substantial improvements of the final version. We would like to thank the staff at the Kazakhstan Agency on Statistics (KAS) in Almaty, who gave us access to the data and provided detailed information. In particular, we are grateful to Ludmila Agaltsova, Nina Krivko, Vera Osokina, Galina Semibratova, and Juri Shokomanov. However, in all cases, the usual disclaimer applies.

## 1. Introduction

Due to the ideology of the previous system, labour compensation in the formerly planned economies of the Commonwealth of Independent States (CIS) was based on the idea of the state ownership of the work force. The central planning authority allocated workers to industries and jobs across regions and undertook the responsibility of training the work force, bearing all the costs of training. Wages were set according to a tariff wage grid for each job category and a compensating wage differential was paid to those who worked in regions with adverse weather and working conditions. Given the fact that agents' expectations were formed in a 'stable price' environment, the relationship between wage and effort became weak. More importantly, the universally observed positive relationship between education and earnings weakened.<sup>2</sup> In some cases, as Graeser (1988) observes, the relationship became perverse.

Returns to schooling under planning were low in many transition economies.<sup>3</sup> However, recent empirical evidence show that returns to schooling increase as market reforms take place. Using metadata from 39 studies of 11 transition economies, Fleisher et al. (2005) report that returns to schooling tended to rise almost immediately in the Czech Republic, Estonia, Poland, Romania, Russia, Slovak Republic, and Slovenia following reform. Münich, Svejnar, and Terrell (2005) report that both women and men experienced similar and significant increases in returns to education in the Czech Republic, Estonia, Hungary, Poland, and Slovenia in the early transition period.

---

<sup>2</sup> The central planner allocated labour to the priority sectors of the economy through the highly arbitrary wage differentiation mechanisms. According to Katz (2001), this policy of wage compression often explained the low returns to schooling observed in the Soviet Union. Wage premiums for highly-educated workforce were eroded by the pay differentials in favour of manual workers with low educational attainment.

<sup>3</sup> See, for example, Gregory and Kohlhase (1988) for the Soviet Union, Andrén, Earle, and Săpătoru (2005) for Romania and Fleisher, Sabirianova, and Wang (2005) for China.

In this paper, we examine returns to schooling in Kazakhstan, a country that has received limited attention in the literature.<sup>4</sup> Existing evidence found in Rama and Scott (1999) show favourable rates of return to schooling, at least in the early reform period.<sup>5</sup> We use recently collected data, the 2001 Kazakhstan Household Budget Survey (KHBS), to examine returns to human capital in the late reform period. In this context we address three key questions. Firstly, do the returns to schooling in Kazakhstan differ from other transition economies? Secondly, do the OLS estimates where the endogeneity bias is ignored differ significantly from the estimates where this bias is taken into consideration? Thirdly, do the returns to schooling differ for men and women?

These questions are of interest for several reasons. Firstly, the pace and extent of reforms in Kazakhstan had a dramatic impact on the labour markets and, in turn, may have influenced rewards to human capital. Secondly, it is well established that conventional OLS estimates of the returns to schooling suffer from a bias if schooling is endogenous. However, if an individual obtained his or her education under the previous economic regime in which educational access was largely centrally determined, it seems highly unlikely that education was chosen with regard to the expected returns in transitional setting.<sup>6</sup> This may suggest that endogeneity of education is less of a concern in transition economies. If this is true, then the OLS approach would be more appropriate. In this respect, the impact of endogeneity bias in Kazakhstan is an

---

<sup>4</sup> We are unaware of any study that estimates returns to human capital in Kazakhstan before the transition.

<sup>5</sup> These authors use the 1996 Kazakhstan Living Standards Measurement Survey (LSMS). The sample contains around 2,000 randomly selected households and 2,278 individuals who report a positive income in the form of salary, bonus, profit, pension, allowance and occasional earnings. The OLS estimated rates of return are in the region of 7-9%.

<sup>6</sup> Even if we assume that individuals adjust their education to market information with a lag, it is likely that only the most recent graduates would be altering educational choices based on expectations of returns in a market system.

interesting case. Thirdly, it is possible that the transition process, in association with an underlying Islamic culture, has allowed the resurgence of more patriarchal attitude to the acquisition of human capital by ethnic Kazakh women. If true, this is likely to have generated a reduction in the returns to education for ethnic Kazakh women and, therefore, the incentives to acquire education in the first place. The answers to these questions have important policy implications.

Our empirical findings show that returns to schooling in Kazakhstan are higher than the estimates found in Rama and Scott (1999). The returns are also higher for women than men. Our results suggest that OLS estimates of the returns to schooling that assume the exogenous nature of schooling and, therefore, do not control for endogeneity bias, may underestimate the true rates of return. Furthermore, we demonstrate the importance of controlling for endogeneity bias and sample selection bias simultaneously.

The structure of this paper is as follows. In Section 2, we examine changes in the Kazakh labour market. Section 3 discusses education in Kazakhstan under planning and reform. In the next section, we introduce the data set and variables used for the analysis. Section 5 presents our methodology and Section 6 provides the empirical results. Section 7 concludes the paper.

## **2. Labour market**

Labour market conditions in Kazakhstan worsened after independence, when central planners withdrew from managing the state-owned enterprises (SOEs). Price liberalisation, initiated in January 1992, together with inflation and currency reform, led to a significant fall in real wages. GDP per capita, measured at purchasing power parity,

declined from \$4,240 in 1991 to \$3,420 in 1996, with real wages declining by 65% over the same period (UNICEF, 2005).

In the immediate post-reform period, cuts in real wages were accompanied by the aggregate fall in labour demand. High unemployment rates followed the restructuring of the SOEs. In July 1993, around 200 enterprises introduced part-time work or stopped production, affecting around 132,000 workers, or 1.8% of the labour force (Sziracki, 1995). The unemployment rate rose from 7.5% in 1994 to 13.5% in 1999 (KAS 2004, 2005). The decline in recorded employment was heavily concentrated in the industrial sector. Industrial employment declined by 41% over the period 1989-97 (UNECE, 2005).<sup>7</sup>

The fall in employment rates was sharper for women than for men, and this trend continued in the late 1990s. Data on officially registered unemployed in 1997 suggests an average of 171 unemployed women per 86 unemployed men (UNICEF, 2005). According to the 2001 KHBS, as shown in Table 1, the unemployment rate for men (women) was 13.6% (15.8%). Table 1 also shows that participation rates for both sexes were high in 2001.

Self-employment grew substantially in Kazakhstan following labour market adjustment. The growth in self-employment reflects the collapse of formal sector jobs and wages more than the emergence of new private opportunities. In addition, most self-employment was not full-time, but reflected work done in parallel to a formal, low-paying job. Many who are on forced or unpaid leave would temporarily

---

<sup>7</sup> According to a UNECE project, the fall in industrial employment in Kazakhstan was larger than in several Central and Eastern European (CEE) countries (Czech Republic, Poland, Romania, Serbia and Montenegro, Slovakia, and Slovenia) and in the CIS economies (Belarus, Russia, Turkmenistan, and Uzbekistan).



engage in private sector activities to compensate for the loss of income from the main job. Agriculture, hunting and related service activities are the most common forms of self-employment in Kazakhstan. The share of self-employed workers engaged in these three sectors (as percentage of total self-employed workers) was 62.2% in 2001, 67.4% in 2003, but then declined to 66% in 2005 (KAS 2004, KAS, 2006).<sup>8,9</sup>

Table 1  
Labour market indicators (percent) in 2001

| Gender:                                       | Men                                    | Women                                  |
|---|--|--|
| Labour force participation rates <sup>a</sup> | 76.2 <sup>b</sup> (83.2 <sup>c</sup> ) | 63.3 <sup>b</sup> (78.8 <sup>c</sup> ) |
| Unemployment rates <sup>a</sup>               | 13.6 <sup>b</sup> (9.0 <sup>c</sup> )  | 15.8 <sup>b</sup> (12.6 <sup>c</sup> ) |
| Self-employment rates <sup>a</sup>            | 60.4 <sup>b</sup> (37.4 <sup>c</sup> ) | 48.6 <sup>b</sup> (42.6 <sup>c</sup> ) |

*Notes:* <sup>a</sup> Definitions: Working age population (for men aged 16-63 years and for women aged 16-58 years); Labour force = employed (working age) plus unemployed (working age); Labour force participation rate = labour force as a percentage of working age population; Unemployment rate = unemployed as a percentage of the labour force; Self-employment rate = self-employed (working age) as a percentage of employed. <sup>b</sup> Our own estimations based on the 2001 KHBS. <sup>c</sup> *Source:* KAS (2004), Table 6.1 (p. 289). Labour force participation rates, unemployment rates and self-employment rates for working age population.

Despite job losses the employment rate in Kazakhstan remained high by international standards. According to our data, the employment rate was 58% in 2001.<sup>10</sup> We also found that the male employment rate is significantly higher than the female rate in all localities with the exception of Almaty and Astana where the employment rates are similar (Table 2). Similar to the evidence reported by Rutkowski (2006, Table 1) for Georgia, Ukraine, Latvia and Lithuania, it seems that labour force adjustment in

<sup>8</sup> The changing patterns of employment in the post-transition period in Kazakhstan are similar to other CIS economies. See Rutkowski (2006) for a detailed analysis of the differences in labour market adjustments between the CIS economies and CEE countries.

<sup>9</sup> The share of self-employment in Kazakhstan is lower than in the low-income CIS economies (Azerbaijan and Kyrgyzstan), but much higher than in Russia and Belarus (Rutkowski, 2006).

<sup>10</sup> At 58% the employment rate for Kazakhstan is similar to Ukraine (60%) and Georgia (56.8%) in 2001, somewhat lower than in Russia (75.8%) in the same year, higher than in Southern and Eastern Europe (Bosnia and Herzegovina, Croatia, Macedonia, and Serbia and Montenegro) with the average of 45.7% in 2001, but similar as in CEE where the average was 59.3% also in 2001 (see Rutkowski 2006, Table 1).

Kazakhstan displayed a combination of high labour force participation rate and high unemployment rate in the late 1990s. However, labour market conditions tightened between 2001 and 2004, with real wage increasing by 14% (IMF, 2005). The unemployment rate declined from 13.5% in 1999 to 8.2% in 2005 (KAS, 2005; KAS, 2006). Recent evidence also shows that participation continues to remain high at 82.7% for men and 76.3% for women in 2005 (KAS, 2006). In the last five years Kazakhstan has witnessed a long awaited economic recovery. However, high youth unemployment, a large pool of long-term unemployed, regional differences in the unemployment and employment rates, and wage inequalities continue to dominate labour market transition in Kazakhstan (Rutkowski, 2006; KAS, 2006).

Table 2  
Employment rates (percent) by municipality type in 2001<sup>a</sup>

| Gender:       | Men               | Women             |
|---------------|-------------------|-------------------|
| Almaty/Astana | 69.3 <sup>b</sup> | 69.6 <sup>b</sup> |
| Large city    | 65.2 <sup>b</sup> | 46.8 <sup>b</sup> |
| Medium city   | 65.7 <sup>b</sup> | 42.4 <sup>b</sup> |
| Small city    | 58.0 <sup>b</sup> | 37.0 <sup>b</sup> |
| Village       | 60.8 <sup>b</sup> | 29.9 <sup>b</sup> |

*Notes:* <sup>a</sup> Definition: Employment rate = employed as a percentage of working age population. “Small City” refers to a population point with a number of households between 3,000 and 10,000. “Average City” refers to a population point with more than 10,000 households but less than 30,000. “Large City” refers to a population point with greater than 30,000 households. “Astana/Almaty” refers to the two major cities in Kazakhstan: the capital city, Astana and the financial capital, Almaty. <sup>b</sup> Our own estimations based on the 2001 KHBS.

### 3. Educational system

In the 1930s, following accession to the Soviet Union, Kazakhstan enjoyed a remarkably high literacy rate of 83.6%, on a par with advanced Western nations (Shokomanov, 2001). This was a direct result of education reform carried out by the Communist government. Schools were established to provide all forms of education for

children, with priorities given to children of working class origin. A network of vocational schools was established to train unskilled workers. In 1958, the Soviet Education Law abolished university tuition fees. However, during the Communist era the educational system in Kazakhstan was highly responsive to the needs of the totalitarian regime. Priority was given to educational specialisations with possible future military applications. Employment opportunities after graduation were constrained by the assignment system. University graduates were required to take up posts for a specified number of years (usually 3 years) after graduation (Oxenstierna, 1990). Placement of graduates from vocational-technical schools (VTSs) and, in some cases, school-leavers was also administratively regulated.

At the onset of transition Kazakhstan enjoyed high levels of human capital.<sup>11</sup> However, the provision of the basic forms of education deteriorated in the initial stages of transition. Expenditure on education, measured as the share of GDP, fell from 6.5% in 1991 to 3.2% in 2001 (KAS, 2002). The decline in enrolment rates has been especially sharp in the early reform period (Table 3). Enrolment rates in the nursery level of education for children between one and six years of age declined from 43.9% in 1992 to 10.2% in 1998.<sup>12</sup> Secondary education enrolment rate fell from 95% in 1992 to 78.7% in 1996. Recent evidence found in Shokomanov (2001, p. 183) suggest that 7.6% of children aged 15, 5.3% of children aged 16 and 6.9% of children aged 17 did not complete basic secondary education during the academic year 1999-2000. In the late

---

<sup>11</sup> According to the 1989 Soviet Census, around 11.5% of the Kazakh population reported university education (KAS, 2002).

<sup>12</sup> On the supply side, the number of nurseries fell from 8,881 in 1991 to 1,103 in 2001 (KAS, 2002). Given the continuous fall in output, many of the SOEs had to abandon the provision of affordable (nominal fees were charged) nursery education for the children of employees. The number of state nurseries also declined. On the demand side, fertility rate (births per woman aged 15-49) fell from 2.84 in 1989 to 1.84 in 1999 (UNICEF, 2005). According to UNDP (2004), the fall in demand is also related to falling family incomes and rising childcare fees. In addition, the shift into the informal economy and self-employment may raise the share of women who are available to provide childcare at home.

1990s over 30% of VTSS were closed, and the number of vocational students fell from 225,600 in 1991 to 87,327 in 2001 (KAS 2002, pp. 78-79).

Table 3  
Enrolment rates (percent), 1992, 1994, 1996, 1998 and 2000

|           | 1992 | 1994 | 1996 | 1998 | 2000 |
|-----------|------|------|------|------|------|
| Nursery   | 43.9 | 35.1 | 20.7 | 10.2 | 7.7  |
| Primary   | 98.7 | 99.5 | 99.5 | 99.8 | 99.6 |
| Secondary | 95   | 88.9 | 78.7 | 82.6 | 93.4 |

Source: Shokomanov (2001, Figure 6.2).

Following the elimination of the state monopoly on educational provision, VTSS, colleges and universities introduced tuition fees.<sup>13</sup> University scholarships and student loans are now awarded by the state on a competitive basis. However, only 9.5% of university students were funded by student grants in 2003, 6.2% were funded by student loans, 2% of university students were funded under the state educational order, and the vast majority of university students (82.3%) were privately funded (UNDP, 2004).

It has long been established that the labour force in the Soviet Union was highly educated (Oxenstierna, 1990; Katz, 2001). This was also evident in late-transition.<sup>14</sup> Table 4 shows that the percentage of men with a university degree increased from 15.5% in 2001 to 19.2% in 2005. Similarly, the share of men with a vocational qualification rose from 9.3% in 2001 to 14.6% in 2005. During the same period the

<sup>13</sup> The State Order System (SOS) is now responsible for the allocation of scholarships and grants to encourage vocational training. VTSS and colleges that qualify under the SOS undertake the training, as requested by the state, and receive funding accordingly.

<sup>14</sup> We must note that following economic transition the quality of education deteriorated, particularly in rural and remote areas, and in those schools with Kazakh as the language of instruction (UN 2004, p. 33). The UN (2004, p. 33) reports that the content of education in secondary schools is not optimal for providing such skills as critical thinking, decision-making and team-work, and is unable to meet the needs of the labour market.

percentage of women with a university degree rose by 4.3 percentage points from 20.2% to 24.5%. There is much less of a shift into vocational qualification compared with men as only 9% of women reported such qualification in 2005. One possible explanation for this result is that young women are reluctant to obtain VTS training, and prefer to enter a college that provides training for non-manual occupations. Finally, for both genders, there is a sharp fall in the percentage with a high school qualification from 38.3% (33.7%) in 2001 to 32.4% (28.2%) in 2005 for men (women).

Table 4  
Employment shares (percent) by education, 2001-2005

|                       | 2001 | 2002 | 2003 | 2004 | 2005 |
|-----------------------|------|------|------|------|------|
| <i>Panel A. Men</i>   |      |      |      |      |      |
| University degree     | 15.5 | 15.9 | 17.2 | 18.2 | 19.2 |
| Incomplete degree     | 2.5  | 2.5  | 3.0  | 3.0  | 2.9  |
| College               | 24.7 | 26.0 | 26.1 | 25.8 | 25.1 |
| Vocational (VTS)      | 9.3  | 15.6 | 16.7 | 15.0 | 14.6 |
| High school           | 38.3 | 33.4 | 31.4 | 32.7 | 32.4 |
| Incomplete secondary  | 6.9  | 5.1  | 4.7  | 4.5  | 5.0  |
| Primary               | 2.8  | 1.5  | 0.9  | 0.8  | 0.8  |
| <i>Panel B. Women</i> |      |      |      |      |      |
| University degree     | 20.2 | 20.5 | 21.4 | 23.0 | 24.5 |
| Incomplete degree     | 2.6  | 2.5  | 3.2  | 3.3  | 3.6  |
| College               | 30.5 | 32.0 | 31.2 | 30.4 | 29.6 |
| Vocational (VTS)      | 6.2  | 9.7  | 11.0 | 10.3 | 9.0  |
| High school           | 33.7 | 30.4 | 28.7 | 29.0 | 28.2 |
| Incomplete secondary  | 5.0  | 4.0  | 3.9  | 3.5  | 4.0  |
| Primary               | 1.8  | 0.9  | 0.6  | 0.5  | 1.1  |

Source: KAS (2004, 2005, 2006).

## 4. Data

The data used are from the 2001 KHBS. The survey for this data was conducted by the KAS for 12 months and completed in January 2002. The survey covers around 12,000 households comprising 45,736 individuals. Interviews were conducted in all of

the country's 14 regions. The sample survey was randomly selected from a register of dwellings based on the territorial division of the housing register.<sup>15</sup> This insured that each household had an equal opportunity of being selected. Data were collected, among other things, on a range of basic demographic, health and education indicators, and on labour market behaviour and outcomes.

The labour market questionnaire, answered by all individuals in the household over the age of 15, provides information on household income and employment. This questionnaire was conducted on a quarterly basis, but income questions were asked on a monthly basis. In accordance with the ILO, the respondent is considered to be employed if he/she worked for at least one hour in the past seven days and received some form of monetary payment or payment in kind. The education questionnaire covers all members of the household aged six and above. The education level must be certified by the respondent who is obliged to produce a valid education diploma.

Our sample does not provide information on the number of years of schooling. The respondents are asked about their highest level of education. Therefore, we construct  $S$ , years of schooling, in the following way: if no qualification or nursery education is indicated  $S=1$ , if primary  $S=3$ , if general secondary  $S=8$ , if high school  $S=10$ , if VTS  $S=10$ , if college  $S=12$ , if degree  $S=15$  and if postgraduate  $S=20$ .

The wage is defined from income for December 2001. This month is used because it is the most recent in the survey and has the highest number of respondents. The dependent variable used in the earnings equations is the log of monthly cash earnings received from the main job. We exclude earnings from secondary jobs or from

---

<sup>15</sup> The compilation of the housing register was part of the 1999 National Population Survey.

agricultural production and non-monetary benefits. Main job wages are net of payroll and income taxes, and exclude pensions and welfare payments. Our data suggest that non-wage compensation, a pervasive theme of economic transition in many of the CIS economies, represents only a small share of the household income.

The present survey does not ask questions about hours worked. It is important to note that the use of monthly wage variable may be problematic since it may be distorted due to possible differences in the average hours of work of male and female workers.<sup>16</sup> However, evidence presented in KAS (2001) suggests that these differences are small. This survey provides information on wages and hours of work for as many of the 28 occupations in 18 industries as possible, for the whole country. The survey found that men and women in Kazakhstan supply very similar hours.<sup>17</sup>

The length of training is not observed in the present data and therefore we are unable to control for experience or tenure (years in the present job) with a direct measure. Thus, we use age which is entered in linear (Age) and quadratic forms (Age squared). Additional variables include a set of dummies representing type of settlement (Village) and locality (Astana).

In columns 1 and 2 of Appendix Table A.1, we present summary statistics for the overall sample. To focus on working age adults, we exclude students, children who are less than 16 years of age, disabled and pensioners. In addition, there are some respondents who did not report their wage, education, or household composition

---

<sup>16</sup> We must mention that the use of monthly earnings to estimate the earnings function in transition economies was adopted by various other researchers (see, for example, Pailhe, 2000 and Newell and Reilly, 2001 for the Czech Republic, Hungary and Slovakia; Hunt, 2002 for East Germany; Adamchik and Bedi, 2003 for Poland; Brainerd, 1998 and Arabsheibani and Lau, 1999 for Russia).

<sup>17</sup> Similar findings are also reported for the Soviet Union by Katz (2001). She reports that the “usual hours of work per week are 42.67 hours for men and 40.83 hours for women” (Katz 2001, Table 5.1).

information. These observations are also excluded from the analysis. Our empirical analysis of earnings is restricted to the sample of individuals whose wage was positive at the time of the survey. The sample retained includes 12,549 individuals between the ages of 16 (the school-leaving age) and 63 (58) in the case of men (women), who also report a wage from the main job (63 is the state retirement age for men and 58 is for women). Our sample consists of 6,690 working men and 5,859 working women. There are considerable differences in the characteristics of men and women with respect to education as revealed in the summary statistics of the variables. For women, the average length of time spent in education is 0.8 years higher than for men.

## 5. Empirical framework

The conventional method of estimating the rate of returns to education is Mincer's (1974) semi-logarithmic earnings function. The simple Mincerian earnings function, adopted by various researchers, can be presented as:

$$\ln Y_i = \beta_0 + \beta_1 S_i + \beta_2 X_i + \beta_3 X_i^2 + \varepsilon_i \quad (1)$$

where  $\ln Y_i$  is the natural logarithm of the observed wage for individual  $i$ ,  $\beta_0$  is a constant term,  $S_i$  is the number of years of schooling of individual  $i$  and  $\beta_1$  is the average rate of return to one additional year of schooling;  $X_i$  is post-schooling experience of the individual worker which is entered in linear and quadratic forms,  $\varepsilon_i$  is assumed to be normally distributed with mean zero and variance  $\sigma_i^2$ . In our case we use age and its square instead of experience.

To control for endogeneity bias, we adopt an IV approach and use spouse's education as an instrument. We follow the methodology of Trostel, Walker, and Wooley



(2002) which explores the independence of wife's education from husband's earnings and its interaction with husband's education. They rely on the assortative nature of marriage as married couples share common interests and behavioural traits, and they usually share a common level of schooling (Pencavel, 1998). On average their IV estimates using spouse's education to instrument for schooling are over 20% higher than the corresponding OLS estimates, suggesting that conventional OLS estimates may be biased downwards. Following Trostel et al. (2002), we also rely on the assortative nature of marriage.<sup>18</sup>

We also use information about smoking habits at age 18 to instrument schooling. The rationale behind our choice of instrument stems from the correlation between health habits and education that can be partially explained by the unobserved differences in discount rates across individuals (Fuchs, 1982). Therefore, some habits, such as smoking, can serve as a good proxy for discount rates and can be used to instrument schooling (see, for example, Evans and Montgomery, 1994; Chevalier and Walker, 1999; Fersterer and Winter-Ebmer, 2003).

Normally, to address the problem of endogeneity of schooling the following two-equation model describing log of earnings ( $\ln Y_i$ ) and years of schooling ( $S_i$ ) is applied:

$$\ln Y_i = \beta S_i + \delta Z_i + \varepsilon_i \quad (2)$$

$$S_i = \alpha Z_i + \eta_i \quad (3)$$

---

<sup>18</sup> We found from the work of Sviatova, Kuandykov, and Chulkina (1988) evidence in support of this argument for Kazakhstan. Using data on marriage certificates issued in 1954 and 1984, the authors study marriage structure of Almaty's population in relation to birth places and nationality of married couples. They found positive assortative mating in all basic national groups of the population.

where  $Z_{1i}$  is a vector of exogenous variables that determine individual earnings ( $\ln Y_i$ ),  $Z_i$  is a vector of exogenous variables that influences the schooling decision,  $E(Z_i', \eta_i) = 0$ ,  $Z_{1i}$  is a subset of  $Z_i$ .

The return to schooling is measured by  $\beta$  and since  $S_i$  is defined as the residual from a regression of  $\eta_i$  on  $Z_i$ , it is uncorrelated with  $Z_{1i}$  by construction. Defining  $\alpha Z_i = \alpha_1 Z_{1i} + \alpha_2 Z_{2i}$ , we assume that  $\alpha_2 \neq 0$ . This assumption results from the endogeneity of  $S_i$  and is identical to the rank condition needed for identifying Eq. (2) in the absence of sample selection (Wooldridge 2002, p. 568). Eq. (2) is then estimated by substituting the fitted values from the first stage regression of  $S_i$  on  $Z_i$  as defined in Eq. (3).

In the case of females, there is an additional problem. It can be argued that women who work may not be a random sample of all women. If working women self-select then results from regressions according to the uncorrected earnings Eq. (2) might be biased since some of the factors that increase the likelihood of women working may also be factors that make their earnings high or low. To correct for the sample selection bias when a possibility exists that  $S_i$  may not be truly exogenous, we follow the suggestion made by Wooldridge (2002). We are aware of only three other studies, that of Garcia, Hernandez, and Lopez-Nicolas (2001), Arrazola and De Hevia (2003) and Das, Newey, and Vella (2003) that address the two issues simultaneously.

The selection Eq. (4) determines female labour force participation decision. We can define  $I_i$  as:

$$I_i = \psi P_i + u_i \tag{4}$$

where  $I_i$  is the endogenous selection process which determines the individual decision of a woman to participate,  $P_i$  is a set of explanatory variables determining the labour force participation of women so that  $(\varepsilon_i, u_i)$  is independent of  $P_i$ ;  $u_i \sim \text{Normal}(0,1)$ ;  $E(\varepsilon_i | u_i) = \gamma_i u_i$ .

Since we do not observe  $I_i$  the decision of a woman to participate or not is indicated by:

$$D = 1 \quad \text{if} \quad I > 0 \quad (\text{participate}) \quad (5)$$

$$D = 0 \quad \text{if} \quad I \leq 0 \quad (\text{do not participate}) \quad (6)$$

To derive an estimating equation, we write:

$$\ln Y_i = \beta S_i + \delta Z_{li} + g(\psi, P_i) + e_i \quad (7)$$

where  $g(\psi, P_i) \equiv E(\varepsilon_i | \psi, P_i)$  and  $e_i \equiv \varepsilon_i - E(\varepsilon_i | \psi, P_i)$ .

By definition:  $E(\varepsilon_i | \psi, P_i) = 0$ . If we knew  $g(\psi, P_i)$ , then according to Wooldridge (2002, p. 568), we could just estimate Eq. (4) by 2SLS on the selected sample ( $I_i = 1$ ) using instruments  $[Z_i, g(P_i, 1)]$ . Notice that we do know  $g(P_i, 1)$  up to some estimable parameters:  $E(u_i | P_i, I_i = 1) = \gamma_i \lambda(\psi, P_i)$ . Therefore, we can obtain  $\hat{\psi}_i$  by a probit of  $I_i$  on  $P_i$  using the entire sample and then obtain the estimated inverse Mills ratios,  $\hat{\lambda}_i = \lambda(\hat{\psi}, P_i)$ . Eq. (2) can now be estimated by 2SLS, using instruments  $(Z_i, \hat{\lambda}_i)$ . Identification requires that  $Z_2$  appear in the linear projection of  $S_i$  onto  $Z_1, Z_2$ , and  $\lambda_i = \lambda(\psi, P_i)$  in the selected subpopulation. The hypothesis of sample selection bias, allowing  $S_i$  to be endogenous or not,  $H_0 : \gamma_i = 0$ , is tested using the usual

2SLS  $t$  statistics for  $\hat{\gamma}_i$ . In essence, this estimation procedure is a mixture of the Heckman (1979) two-step procedure and standard 2SLS.

Following the suggestion made by Wooldridge (2002), we exclude education from the selection Eq. (4) to avoid any possible distortion in the analysis produced by endogeneity of education and include all the exogenous variables as instruments and as regressors in the selection equation. Conditions of our model require to have at least one variable to be included in  $P_i$  which does not also appear in  $Z_{1i}$ . Identification is achieved by including an additional variable: number of children less than six years old (Dependent children).<sup>19,20</sup>

As Bound, Jaeger, and Baker (1995) have shown, the instrument validity exists only if there is a strong correlation between the potential instrument and schooling. The authors suggest that if the potential instrument is weakly correlated with schooling, it is likely to result in a large inconsistency of IV estimates. To check for the instrument validity, as suggested by Bound et al. (1995), we perform  $F$ -tests on the excluded variables and report partial  $R^2$  from the first stage regression. We also report the Wu-Hausman test for endogeneity of schooling, which allows us to reject or accept the null hypothesis that the OLS estimates are consistent.

For our IV estimates, we are restricted to the sample of married respondents. This sample consists of 5,198 men and 3,340 women. Columns 3 and 4 of Appendix Table A.1 present the means and standard deviations of the variables that we use in the

---

<sup>19</sup> We presume that fertility decisions are exogenous to decisions about labour force participation and expected returns from that participation. This is an assumption that is implicitly made in the literature.

<sup>20</sup> In the past the non-linearity of the selection term was thought to identify the model. However, Heckman (1990) suggests what is needed is a continuous variable in the selection equation that is excluded from the earnings equation. The identifying variable used in our model addresses this point.

IV estimates. On average married men are older and have fewer years of schooling than married women. For women, the average length of time spent in education is 0.7 years higher than for men. The average age of the married men and women is 41.8 and 40.5 years respectively. We found that a much higher proportion of men smoked at the age of 18 when compared to women.

## **6. Estimates of the returns to schooling**

In panel A of Table 5, we present the OLS estimates of returns to schooling for the overall sample of male and female respondents. Our estimates are statistically significant, higher than the international average of 7-8% and certainly higher than in other transition economies. In addition, similar to the evidence from other transition economies reported by Munich et al. (2005), we find that the rate of return to schooling is higher for women, 11.5%, compared to 8% for men. Moreover, our results are similar to those found in middle-income developing countries reported by Psacharopoulos and Patrinos (2004). Given that schooling information is provided in levels rather than years, the obvious extension to Eq. (1) is to re-estimate the earnings function by replacing  $S$  with a set of dummy variables representing levels of education. We present these results in panel B of Table 5. The returns by level of education are hierarchical. Women with a university degree earn 124% more when compared to the reference group, and men with a university degree earn 91% more.<sup>21</sup> The returns to college and vocational qualifications are similar for both genders. These results differ with evidence found in other transition economies by Munich et al. (2005), but similar to the results

---

<sup>21</sup> Following Halvorsen and Palmquist (1980), we compute these estimates using the following formula:  $(e^{\beta}-1)\times 100$ .

reported by Joliffe and Campos (2005) for Hungary in 1995 and 1998. Those with a college degree earn around 52% more than the reference category. Those with a vocational education earn around 27% more. High school graduates earn 0.24 (0.18) log points or 27% (20%) more than the reference category in the case of men (women).

Table 5  
Estimated returns to years of schooling and qualifications for overall sample

| Gender:                        | Men      | Men      | Women    | Women    |
|--------------------------------|----------|----------|----------|----------|
| Technique:                     | OLS      | OLS      | OLS      | OLS      |
| Variable                       |          |          |          |          |
| <i>Panel A</i>                 |          |          |          |          |
| Schooling                      | 0.080*** |          | 0.115*** |          |
|                                | (0.004)  |          | (0.004)  |          |
| <i>Panel B. Qualifications</i> |          |          |          |          |
| High school                    |          | 0.245*** |          | 0.182*** |
|                                |          | (0.042)  |          | (0.054)  |
| Vocational                     |          | 0.243*** |          | 0.230*** |
|                                |          | (0.046)  |          | (0.059)  |
| College                        |          | 0.427*** |          | 0.412*** |
|                                |          | (0.043)  |          | (0.052)  |
| University                     |          | 0.646*** |          | 0.806*** |
|                                |          | (0.044)  |          | (0.053)  |
| Postgraduate                   |          | 0.976*** |          | 1.170*** |
|                                |          | (0.140)  |          | (0.178)  |
| <i>N</i>                       |          | 6,690    |          | 5,859    |

Notes: (i) Robust standard errors are in parentheses. (ii) The reference category in panel B is 'less than high school education'. (iii) Full estimates in Appendix Table A.2.

\*\*\* Significance at the 1 % level.

In Table 6, we present returns to schooling for our sample of married men. The first column gives the OLS estimates of Eq. (1) with an estimated rate of return equal to 7.8% that is slightly lower than for the overall sample of men. In column 2, we use spouse's education as an instrument. The reliability of the IV estimates depends on whether the instrument is valid. The partial  $R^2$  and the first-stage  $F$ -statistic confirm validity of the instrument. Further, the result of the endogeneity test rejects the null that

the OLS estimates are consistent. The IV rate of return is higher than OLS by 3.4 percentage points. Column 3 uses wife's education together with smoking as instruments. The 2SLS rate of return is similar to the results found in column 2. The partial  $R^2$  and  $F$ -statistic confirm that the instruments are not weak. The  $J$ -test of over-identifying restrictions confirms that both instruments are exogenous.<sup>22</sup>

Table 6  
Estimated returns to years of schooling for married men

| Technique:<br>Instrument(s): | OLS                 | IV<br>Wife's schooling | 2SLS<br>Wife's schooling and<br>smoking |
|------------------------------|---------------------|------------------------|---|
| Variable                     |                     |                        |   |
| Schooling                    | 0.078***<br>(0.005) | 0.112***<br>(0.011)    | 0.113***<br>(0.011)                     |
| Partial $R$ -squared         |                     | 0.219                  | 0.229                                   |
| First-stage $F$ -statistic   |                     | 927.35                 | 519.05                                  |
| $P$ -value                   |                     | (0.000)                | (0.000)                                 |
| Exogeneity $F$ -test         |                     | 10.765                 | 12.264                                  |
| $P$ -value                   |                     | (0.001)                | (0.000)                                 |
| $J$ -test                    |                     |                        | 0.259                                   |
| $P$ -value                   |                     |                        | (0.610)                                 |
| $N$                          |                     | 5,198                  |   |

Notes: (i) Robust standard errors are in parentheses. (ii) Full estimates for column 2 are shown in Appendix Table A.3. (iii) Full results for column 3 are available on request.

\*\*\* Significance at the 1 % level.

In Table 7, we present our estimates for the sample of married women. In column 1, we find that the OLS rate of return is 11.9%, a value that is very similar to the OLS estimate of 11.5% for the overall sample of women. In column 2, we report the IV estimate of Eq. (2) that only considers the endogeneity bias.<sup>23</sup> The IV rate of return

<sup>22</sup> When using smoking as the only instrument for the overall (married) sample of men the IV rate of return was 5.6 (2.4) percentage points higher than the OLS estimate. However, for the overall and married samples partial  $R^2$  measures of instrument relevance were under 2%. Full results are available on request.

<sup>23</sup> In the case of females, we reject smoking as an instrument both for the overall and married samples. One possible explanation for this result is that we have only a small fraction of females who smoked at age 18. Full results are available on request.

of 16.7% is higher by 4.8 percentage points when compared to the OLS estimate. In column 3, we report the result from the 2SLS estimate of Eq. (7) that considers selectivity and endogeneity biases simultaneously. Firstly, however, we turn to the results of the probit Eq. (4).<sup>24</sup> As column 2 of Appendix Table A.4 shows, being of Kazakh ethnicity reduces the probability of participation. This may be related to customs and, probably, religious beliefs, which came about after the independence. Rural residency when compared to urban residency also reduces the probability of participation in the labour market. This again may be related to customs and/or alternatively to the diminished employment opportunities. In addition, Kazakh women may be discouraged from working due to wage disadvantage. As shown in Appendix Table A.4, the coefficient for the Kazakh ethnicity is negative and significant, suggesting that Kazakh women have lower wages.<sup>25</sup> Having children in the household reduces the probability of participation. Participation also rises with spouse's education.

Next, we analyse the 2SLS rate of return to schooling. As column 3 of Table 7 shows, the 2SLS rate of return is higher (lower) than OLS (IV) by 1.8 (3) percentage points. The partial  $R^2$  and the first-stage  $F$ -statistic confirm validity of the instruments. The  $J$ -statistic does not reject the null hypothesis that our instruments are exogenous. Further, the result of the endogeneity test rejects the null that the OLS estimate is consistent. The significant selectivity term supports the presence of sample selection bias in the OLS and IV estimates of the rate of return. The negative effect of  $\lambda$

---

<sup>24</sup> We follow the common approach found in the literature in our choice of variables for the selection equation. The participation variable is a dummy that takes value 1 if the respondent reports a wage and 0 otherwise. Married women who are on leave (annual or maternity) are considered as working. The base category includes the inactive married women (working age) as well as the unemployed.

<sup>25</sup> We must note that this earning disadvantage may be assigned to the selectivity problem if the non-Kazakh ethnicities who decided to remain in Kazakhstan may indeed be those who do have better paid jobs while poorly paid non-Kazakh ethnicities decided to emigrate. For a discussion of Kazakhstan's population movement in detail, see Becker, Musabek, Seitenova, and Urzhumova (2005).



suggests that non-participating women, on average, have a comparative advantage to work when compared to participating women. This may be associated with the rise of Islamic sentiments that keeps more women at home even though they, on average, have characteristics that give them an earnings advantage if they were to work.<sup>26</sup>

Table 7  
Estimated returns to years of schooling for married women

| Technique:                      | OLS                 | IV                  | 2SLS                 |
|---------------------------------|---------------------|---------------------|----------------------|
| Instrument:                     |                     | Husband's schooling | Husband's schooling  |
| Selectivity control:            |                     | No                  | Yes                  |
| Variable                        |                     |                     |                      |
| Schooling                       | 0.119***<br>(0.006) | 0.167***<br>(0.013) | 0.137***<br>(0.016)  |
| $\lambda$                       |                     |                     | -0.407***<br>(0.133) |
| Partial <i>R</i> -squared       |                     | 0.222               | 0.170                |
| First stage <i>F</i> -statistic |                     | 950.81              | 291.87               |
| <i>P</i> -value                 |                     | (0.000)             | (0.000)              |
| Exogeneity <i>F</i> - test      |                     | 15.128              | 3.057                |
| <i>P</i> -value                 |                     | 0.000               | (0.080)              |
| <i>J</i> -test                  |                     |                     | 1.658                |
| <i>P</i> -value                 |                     |                     | (0.197)              |
| <i>N</i>                        |                     | 3,340               |                      |

Notes: (i) Robust standard errors are in parentheses. (ii) Full estimates for column 2 are available on request. (iii) Full estimates for column 3 are shown in Appendix Table A4.

\*\*\* Significance at the 1 % level.

Our empirical findings show that the OLS estimates of the returns to a year of schooling are higher than the estimates reported by Rama and Scott (1999). In addition, the OLS based results are in line with recent evidence from transition economies. For example, Gorodnichenko and Sabirianova (2005) find that the rate of return to schooling in Russia was 9.7% in 2002. The results are also comparable to the recent evidence found in Joliffe and Campos (2005) for Hungary in 1998 and Pastore and

<sup>26</sup> See also Ermisch and Wright (1994) for theoretical explanations of negative sample selection effects.

Verashchagina (2006) for Belarus in 2001. We also find that the returns to levels of education are higher than the 1994 and 1996 rates of return reported by Newell and Reilly (1997) using the 1994 Kazakhstan Labour Force Survey and 1996 Kazakhstan LSMS.

With regard to the IV estimates, we find that they are considerably larger than the corresponding OLS estimates. This is consistent with empirical evidence for China (Heckman and Li, 2004), Czech and Slovak Republics (Filer, Jurajda, and Plánovský, 1999), and Ukraine (Gorodnichenko and Sabirianova, 2005). For married women, we find that the IV methodology produces higher rate of return to a year of schooling when compared to the 2SLS that simultaneously accounts for endogeneity and sample selection biases.<sup>27</sup>

## 7. Conclusion

The results from the paper suggest that OLS estimates of the returns to education that assume the exogenous nature of schooling and, therefore, do not control for endogeneity bias, may underestimate the true rates of return. The direction of the OLS bias is in accordance with the majority of empirical studies. For men, the IV estimates are higher than OLS by around 3 percentage points. In the case of women, we found that the rate of return rose to 13.7% (2SLS) after we corrected for selectivity and endogeneity biases.

---

<sup>27</sup> Following the referee's suggestion, we split the sample based on age to address the point raised in Footnote 6. Aged 40 (approximate mean of the working sample) was chosen as the cut off point for both genders to test the hypothesis. The estimated OLS rates of return were higher for workers aged >40 for both genders. However, a  $\chi^2$  test showed that these coefficients are not statistically different from each other. We also found that the IV (2SLS) rate of return was higher (lower) for men (women) aged  $\leq 40$ . However, according to a  $\chi^2$  test, these coefficients are also not statistically different from each other. Full estimates are available upon request.

We found that the gender differences in returns to vocational and college qualifications are small. However, women tend to have lower returns to high school education. This suggests that men have more opportunities to earn a higher salary with a lower level of education compared to women. Further, as shown in Appendix Table A.1, nearly 27% of working women have a university degree compared to about 17% of men. There may be at least two explanations for this. Firstly, higher returns to university education provide an incentive for their greater investment compared to men. Secondly, women may be using education as an instrument in reducing the gender wage gap. While the raw gender wage gap is not extreme, the existence of the gender difference in returns to high school diploma points to discrimination against women.<sup>28</sup> Another interesting result is that ethnic Kazakh women have an earnings disadvantage when compared to European and other Asiatic groups (see columns 3 and 4 of Appendix Table A.2). In addition, they are less likely to participate in the labour market. These findings suggest that future research should concentrate on the pay differences within women.

Finally, we would like to conclude with some general comment about the returns to education in a transitional setting. The evidence from other transition economies show that returns exhibit a rising tendency as market reforms continue (Andrén et al, 2005; Pastore and Verashchagina, 2006). This is certainly true for Kazakhstan when compared to the evidence found in Newell and Reilly (1997) and Rama and Scott (1999). This, in our view, reflects the relative scarcities of more highly educated people in Kazakhstan with human capital that employers require and reward accordingly. In

---

<sup>28</sup> Using the 2001 KHBS, Mussurov (2006) finds that a substantial part of the wage gap remains unexplained and that the log point difference is high (low) against women with higher (lower) salary levels. The author reports a 0.292 log point difference between the 10<sup>th</sup> and 90<sup>th</sup> percentiles. Using the 1996 data, Newell and Reilly (2001) find similar results. The authors report that the difference is of the order of 0.327 log points.

addition, the results indicate that market reforms in Kazakhstan led to productivity augmenting factors being rewarded accordingly. Efforts should continue to concentrate on monitoring the effect of education on earnings as well as gender differences in the returns to education and pay gaps as they remain an important research area. This will assist policy makers in improving educational and labour market policies.

## References

Adamchik, V.A. and Bedi, A.S. (2003). 'Gender pay differentials during the transition in Poland', *Economics of Transition*, 11, pp. 697-726.

Andrén, D., Earle, J.S. and Săpătoru, D. (2005). 'The wage effects of schooling under socialism and in transition: Evidence from Romania, 1950–2000', *Journal of Comparative Economics*, 33(2), pp. 300-323.

Arabsheibani, G. R. and Lau, L. (1999). 'Mind the gap: An analysis of gender wage differentials in Russia', *Labour*, 13, pp. 761-774.

Arrazola, M. and De Hevia, J. (2003). 'Returns to education in Spain: New evidence of the differences between men and women', Working Paper, Universidad Carlos III de Madrid, Department of Statistics and Econometrics.

Becker, C.M., Musabek, E.N., Seitenova, A.S. and Urzhumova, D.S. (2005). 'The immigration response to economic shock: lessons from Kazakhstan', *Journal of Comparative Economics*, 33, pp.107-132.

Bound, J., Jaeger, D.A. and Baker, R.M. (1995). 'Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak', *Journal of the American Statistical Association*, 90, pp. 443-50.

Brainerd, E. (1998). 'Winners and losers in Russia's economic transition', *American Economic Review*, 88(5), pp. 1094-1116.

Chevalier, A. and Walker, I. (1999). 'Further results on the returns to education in the UK', mimeo, Department of Economics, University of Warwick.

Das, M., Newey, W.K. and Vella, F. (2003). 'Nonparametric estimation of sample selection models', *Review of Economic Studies*, 70, pp. 33-58.

Ermisch, J.F. and Wright, R.E. (1994), 'Interpretation of negative sample selection effects in wage offer equations', *Applied Economic Letters*, 1, pp. 187-189.

Evans, W. N. and Montgomery, E. (1994). 'Education and health: Where there's smoke there's an instrument', Working Paper No. 4949, Cambridge, MA: NBER.

Fersterer, J. and Winter-Ebmer, R. (2003). 'Smoking, discount rates, and returns to education', *Economics of Education Review*, 22, pp. 561-566.

Filer, R.K., Jurajda, Š. and Plánovský, J. (1999). 'Education and wages in the Czech and Slovak republics during transition', *Labour Economics*, 6, pp. 581–593.

Fleisher, B.M., Sabirianova, K. and X. Wang. (2005). 'Returns to skill and the speed of reforms: Evidence from Central and Eastern Europe, China, and Russia', *Journal of Comparative Economics*, 33, pp. 351-370.

- Fuchs, V. (1982). 'Time preference and health: An exploratory study' in Fuchs, V. (ed.), *Economic Aspects of Health*, Chicago: University of Chicago Press.
- Garcia, J., Hernandez, P.J. and Lopez-Nicolas, A. (2001). 'How wide is the gap: An investigation of gender wage differentials using quantile regressions', *Empirical Economics*, 26(1), pp. 149-169.
- Gorodnichenko, Y. and Sabirianova, K. (2005). 'Returns to schooling in Russia and Ukraine: A semi-parametric approach to cross-country analysis', *Journal of Comparative Economics*, 33(2), pp. 324-350.
- Graeser, P. (1988). 'Human capital in a centrally planned economy', *Kyklos*, 41 (1), pp. 75-98.
- Gregory P.R. and Kohlhase, J.E. (1988). 'The earnings of the Soviet workers: Evidence from the Soviet Interview Project', *Review of Economics and Statistics*, 70, pp. 23-35.
- Halvorsen, R. and Palmquist, R. (1980). 'The interpretation of dummy variables in semilogarithmic equations', *American Economic Review*, 70(3), pp. 474-475.
- Heckman, J. J. (1979). 'Sample selection bias as a specification error', *Econometrica*, 46(6), pp. 1251-1271.
- Heckman, J. J. (1990). 'Varieties of selection bias', *American Economic Review*, 80(2), pp. 313-318.
- Heckman, J.J. and Li, X. (2004). 'Selection bias, comparative advantage and heterogeneous returns to education: Evidence from China in 2000', *Pacific Economic Review*, 9(3), pp. 155-171.
- Hunt, J. (2002). 'The transition in East Germany: When is a ten-point fall in the gender wage gap bad news?', *Journal of Labor Economics*, 20(1), pp. 148-169.
- International Monetary Fund. (2005). *Public Information Notice (PIN) No. 0587*, Washington, D.C., <http://www.imf.org/external/np/sec/pn/2005/pn0587.htm>.
- Joliffe, D. and Campos, N.F. (2005). 'Does market liberalization reduce gender discrimination? Econometric evidence from Hungary, 1986-1998', *Labour Economics*, 12, pp. 1-22.
- Katz, K. (2001). *Gender, Work and Wages in the Soviet Union*, Basingstoke: Palgrave.
- Kazakhstan Agency on Statistics. (2001). *Basic Labour Market Indicators*, Kazstatinform, Almaty.
- Kazakhstan Agency on Statistics. (2002). *Education in Republic of Kazakhstan*, Kazstatinform, Almaty.
- Kazakhstan Agency on Statistics. (2004). *Economic Activity of Kazakhstan's Population (1991-2003)*, Kazstatinform, Almaty.

- Kazakhstan Agency on Statistics. (2005). *Economic Activity of Kazakhstan's Population (1995-2004)*, Kazstatinform, Almaty.
- Kazakhstan Agency on Statistics. (2006). *Basic Labour Market Indicators*, Kazstatinform, Almaty.
- Mincer, J. (1974). *Schooling, Experience and Earnings*, New York: Columbia University Press.
- Mussurov, A. (2006). 'The gender pay gap in Kazakhstan', Working Paper, KIMEP, Department of Economics.
- Münich, D., Svejnar, J. and Terrell, K. (2005). 'Is women's human capital valued more by markets than by planners?', *Journal of Comparative Economics*, 33, pp. 278-299.
- Newell, A. and Reilly, B. (1997). 'Rates of return to educational qualifications in the transition economies', Discussion Paper No. 03/97, University of Sussex, Department of Economics.
- Newell, A. and Reilly, B. (2001). 'The gender pay gap in the transition from communism: Some empirical evidence', *Economic Systems*, 25, pp. 287-304.
- Oxenstierna, S. (1990). *From Labour Shortage to Unemployment? The Soviet Labour Market in the 1980s*, Almqvist & Wiksell International Press, Stockholm.
- Pailhe, A. (2000). 'Gender discrimination in Central Europe during the systemic transition', *Economics of Transition*, 8, pp. 505-535.
- Pastore, F. and Verashchagina, A. (2006). 'Private returns to human capital: A case study of Belarus', *Economics of Education Review*, 25, pp. 91-107.
- Pencavel, J. (1998). 'Assortative mating by schooling and the work behavior of wives and husbands', *American Economic Review*, 88 (2), pp. 326-329.
- Psacharopoulos, G. and Patrinos, H.A. (2004). 'Returns to investment in education: A further update', *Education Economics*, 12(2), pp. 1325-1343.
- Rama, M. and Scott, K. (1999). 'Labor earnings in one-company towns: Theory and evidence from Kazakhstan', *The World Bank Economic Review*, 13 (1), pp. 185-209.
- Rutkowski, J. (2006). 'Labor market developments during economic transition', World Bank Policy Research Working Paper 3894, <http://econ.worldbank.org>.
- Shokomanov, U. (2001). *Tendencii Chelovecheskogo Razvitiia v Kazahstane*, Kazakhstan Agency on Statistics, Almaty, Kazakhstan.
- Sviatova, G.S., Kuandykov, E.U. and Chulkina, M.P. (1988). 'Demographic genetic characteristics of the population of a large multiethnic city (The example of Alma-Ata). An analysis of assortative mating and migrations', *Genetika*, 24(7), pp. 1269-1275.

Sziraczki, G. (1995). 'Emerging labor market policy in Kazakhstan', Labor Market Papers No. 3, ILO, Geneva.

Trostel, P., Walker, I. and Woolley, P. (2002). 'Estimates of the economic return to schooling for 28 countries', *Labour Economics*, 9(1), pp. 1-16.

UN (United Nations). (2004). *Kazakhstan: Achievements, Issues and Prospects*, Almaty, Kazakhstan.

UNDP (United Nations Development Programme). (2004). *Kazakhstan National Human Development Report 2004*, Almaty, Kazakhstan.

UNECE (United Nations Economic Commission for Europe). (2005). 'Evolution of the Industrial Sector in Transition Economies: A Statistical Overview', <http://www.unece.org/ie/wp8/documents/tradewp82005add2.pdf>.

UNICEF (United Nations Children's Fund). (2005). TransMONEE Database, <http://www.unicef-icdc.org/resources/>.

Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*, Cambridge, Mass: MIT Press.



Appendix Table A.1  
Descriptive statistics

| Sample:<br>Gender:<br>Variable | All<br>Men          | All<br>Women        | Married<br>Men      | Married<br>Women    |
|--------------------------------|---------------------|---------------------|---------------------|---------------------|
| Log Wage                       | 8.99<br>(0.90)      | 8.78<br>(1.03)      | 9.06<br>(0.91)      | 8.74<br>(0.84)      |
| Age                            | 39.36<br>(10.42)    | 39.44<br>(9.46)     | 41.87<br>(9.09)     | 40.57<br>(8.21)     |
| Age <sup>2</sup>               | 1658.08<br>(837.76) | 1645.57<br>(733.93) | 1836.25<br>(781.55) | 1714.10<br>(659.79) |
| Schooling                      | 11.15<br>(2.15)     | 11.96<br>(2.15)     | 11.24<br>(2.17)     | 11.99<br>(2.10)     |
| Nursery/None                   | 0.0005<br>(0.0002)  | -<br>-              |                     |                     |
| Primary                        | 0.007<br>(0.001)    | 0.003<br>(0.0006)   |                     |                     |
| Incomplete secondary           | 0.063<br>(0.002)    | 0.039<br>(0.002)    |                     |                     |
| Complete secondary             | 0.341<br>(0.005)    | 0.232<br>(0.005)    |                     |                     |
| VTS                            | 0.172<br>(0.004)    | 0.094<br>(0.003)    |                     |                     |
| College                        | 0.243<br>(0.005)    | 0.363<br>(0.005)    |                     |                     |
| University                     | 0.169<br>(0.004)    | 0.266<br>(0.005)    |                     |                     |
| Postgraduate                   | 0.008<br>(0.0003)   | 0.0008<br>(0.0003)  |                     |                     |
| Smoke18                        | 0.378<br>(0.005)    | 0.030<br>(0.002)    | 0.378<br>(0.005)    | 0.028<br>(0.002)    |
| Kazakh                         | 0.553<br>(0.006)    | 0.481<br>(0.006)    | 0.562<br>(0.006)    | 0.504<br>(0.006)    |
| Village                        | 0.419<br>(0.006)    | 0.298<br>(0.005)    | 0.420<br>(0.006)    | 0.345<br>(0.005)    |
| Astana                         | 0.024<br>(0.001)    | 0.027<br>(0.001)    | 0.023<br>(0.001)    | 0.027<br>(0.001)    |
| Dependent children             | 0.421<br>(0.006)    | 0.202<br>(0.005)    | 0.450<br>(0.006)    | 0.239<br>(0.005)    |
| <i>N</i>                       | 6,690               | 5,859               | 5,198               | 3,340               |

*Note:* Standard deviations are in parentheses.

Appendix Table A.2

Returns to years of schooling and qualifications for overall sample

| Technique:       | OLS                                 | OLS                                 | OLS                               | OLS                               |
|------------------|-------------------------------------|-------------------------------------|-----------------------------------|-----------------------------------|
| Gender:          | Men                                 | Men                                 | Women                             | Women                             |
| Variable         |                                     |                                     |                                   |                                   |
| Schooling        | 0.080 <sup>***</sup><br>(0.004)     |                                     | 0.115 <sup>***</sup><br>(0.004)   |                                   |
| High school      |                                     | 0.245 <sup>***</sup><br>(0.042)     |                                   | 0.182 <sup>***</sup><br>(0.054)   |
| Vocational       |                                     | 0.243 <sup>***</sup><br>(0.046)     |                                   | 0.230 <sup>***</sup><br>(0.059)   |
| College          |                                     | 0.427 <sup>***</sup><br>(0.043)     |                                   | 0.412 <sup>***</sup><br>(0.052)   |
| University       |                                     | 0.646 <sup>***</sup><br>(0.044)     |                                   | 0.806 <sup>***</sup><br>(0.053)   |
| Postgraduate     |                                     | 0.976 <sup>***</sup><br>(0.140)     |                                   | 1.170 <sup>***</sup><br>(0.178)   |
| Age              | 0.072 <sup>***</sup><br>(0.006)     | 0.072 <sup>***</sup><br>(0.006)     | 0.019 <sup>**</sup><br>(0.007)    | 0.020 <sup>**</sup><br>(0.008)    |
| Age <sup>2</sup> | -0.0008 <sup>***</sup><br>(0.00008) | -0.0008 <sup>***</sup><br>(0.00008) | -0.0002 <sup>**</sup><br>(0.0001) | -0.0002 <sup>**</sup><br>(0.0001) |
| Kazakh           | -0.021<br>(0.021)                   | -0.023<br>(0.021)                   | -0.070 <sup>***</sup><br>(0.020)  | -0.072 <sup>***</sup><br>(0.020)  |
| Village          | -0.536 <sup>***</sup><br>(0.022)    | -0.535 <sup>***</sup><br>(0.022)    | -0.346 <sup>***</sup><br>(0.022)  | -0.345 <sup>***</sup><br>(0.022)  |
| Astana           | 0.457 <sup>***</sup><br>(0.054)     | 0.457 <sup>***</sup><br>(0.054)     | 0.490 <sup>***</sup><br>(0.051)   | 0.491 <sup>***</sup><br>(0.051)   |
| Constant         | 6.937 <sup>***</sup><br>(0.130)     | 7.511 <sup>***</sup><br>(0.126)     | 7.083 <sup>***</sup><br>(0.149)   | 8.021 <sup>***</sup><br>(0.154)   |
| R <sup>2</sup>   | 0.179                               | 0.181                               | 0.169                             | 0.173                             |
| N                |                                     | 6,690                               |                                   | 5,859                             |

Note: Robust standard errors are in parentheses.

\*\* Significance at the 5% level.

\*\*\* Idem., 1%

Appendix Table A.3  
Returns to years of schooling for married men

| Technique:                 | OLS                                | OLS                              | IV                                 |
|----------------------------|------------------------------------|----------------------------------|------------------------------------|
| Equation:                  | Earnings                           | Schooling                        | Earnings                           |
| Variable                   |                                    |                                  |                                    |
| Schooling                  | 0.078 <sup>***</sup><br>(0.005)    |                                  | 0.112 <sup>***</sup><br>(0.011)    |
| Age                        | 0.042 <sup>***</sup><br>(0.010)    | 0.034<br>(0.025)                 | 0.038 <sup>***</sup><br>(0.010)    |
| Age <sup>2</sup>           | -0.0005 <sup>***</sup><br>(0.0001) | -0.0003<br>(0.0003)              | -0.0005 <sup>***</sup><br>(0.0001) |
| Kazakh                     | -0.011<br>(0.024)                  | 0.315 <sup>***</sup><br>(0.053)  | -0.024<br>(0.024)                  |
| Village                    | -0.561 <sup>***</sup><br>(0.025)   | -0.463 <sup>***</sup><br>(0.054) | -0.530 <sup>***</sup><br>(0.027)   |
| Astana                     | 0.547 <sup>***</sup><br>(0.058)    | 0.476 <sup>***</sup><br>(0.181)  | 0.522 <sup>***</sup><br>(0.058)    |
| Instrument                 |                                    | 0.482 <sup>***</sup><br>(0.015)  |                                    |
| Constant                   | 7.662 <sup>***</sup><br>(0.220)    | 4.886 <sup>***</sup><br>(0.536)  | 7.367 <sup>***</sup><br>(0.242)    |
| $R^2$                      | 0.175                              | 0.260                            | 0.169                              |
| Partial $R^2$              |                                    | 0.219                            |                                    |
| First Stage $F$ -statistic |                                    | 927.35                           |                                    |
| $P$ -value                 |                                    | 0.000                            |                                    |
| Exogeneity $F$ - test      |                                    |                                  | 10.765                             |
| $P$ -value                 |                                    |                                  | (0.001)                            |
| $N$                        |                                    | 5,198                            |                                    |

Note: Robust standard errors are in parentheses.

\*\*\* Significance at the 1% level.

Appendix Table A.4

Returns to years of schooling for married women

| Technique:                 | OLS                  | Probit                | OLS                 | 2SLS                 |
|----------------------------|----------------------|-----------------------|---------------------|----------------------|
| Equation:                  | Earnings             | Participation         | Schooling           | Earnings             |
| Variable                   |                      |                       |                     |                      |
| Schooling                  | 0.119***<br>(0.006)  |                       |                     | 0.137***<br>(0.016)  |
| Age                        | 0.025*<br>(0.014)    | 0.164***<br>(0.018)   | 0.185<br>(0.121)    | -0.021<br>(0.020)    |
| Age <sup>2</sup>           | -0.0002<br>(0.0001)  | -0.001***<br>(0.0002) | -0.002<br>(0.001)   | 0.0002<br>(0.0002)   |
| Kazakh                     | -0.079**<br>(0.028)  | -0.279***<br>(0.041)  | -0.022<br>(0.118)   | -0.029<br>(0.037)    |
| Village                    | -0.282***<br>(0.029) | -0.256***<br>(0.040)  | -0.330*<br>(0.185)  | -0.207***<br>(0.035) |
| Astana                     | 0.540***<br>(0.065)  | 0.088<br>(0.141)      | 0.173<br>(0.187)    | 0.501***<br>(0.065)  |
| Dependent children         |                      | -0.496***<br>(0.029)  | -0.233<br>(0.368)   |                      |
| Instrument                 |                      | 0.070***<br>(0.009)   | 0.494***<br>(0.056) |                      |
| $\lambda$                  |                      |                       | 1.271<br>(0.995)    | -0.407***<br>(0.133) |
| Constant                   | 6.830<br>(0.290)     | -3.151***<br>(0.360)  | 2.342<br>(3.335)    | 7.744***<br>(0.570)  |
| $R^2$                      | 0.157                |                       | 0.239               | 0.149                |
| Pseudo $R^2$               |                      | 0.130                 |                     |                      |
| $\chi^2(9)$                |                      | 862.37                |                     |                      |
| $P$ -value                 |                      | 0.000                 |                     |                      |
| Log likelihood             |                      | -2790.37              |                     |                      |
| Partial $R^2$              |                      |                       | 0.170               |                      |
| First Stage $F$ -statistic |                      |                       | 291.87              |                      |
| $P$ -value                 |                      |                       | (0.000)             |                      |
| Exogeneity $F$ -test       |                      |                       |                     | 3.057                |
| $P$ -value                 |                      |                       |                     | (0.080)              |
| $N$                        | 3,340                | 5,030                 | 3,340               | 3,340                |

Note: Robust standard errors are in parentheses.

\* Significance at the 10% level.

\*\* Idem., 5%.

\*\*\* Idem., 1%.