HUMAN CAPITAL ALLOCATION IN THE TRANSITIONAL ECONOMIES OF UKRAINE AND RUSSIA

by

Zhylyevskyy Oleksandr

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Arts in Economics

Economics Education and Research Consortium

The National University of "Kyiv-Mohyla Academy"

2002

Approved by ____
Chairperson of Supervisory Committee

____

____

____

Program Authorized
to Offer Degree ____

Date ____

National University of "Kyiv-Mohyla Academy"

Abstract

HUMAN CAPITAL ALLOCATION IN THE TRANSITIONAL ECONOMIES OF UKRAINE AND RUSSIA

by Zhylyevskyy Oleksandr

Chairperson of the Supervisory Committee: Professor Serhiy Korablin
Institute of Economic Forecasting at the National
Academy of Sciences of Ukraine

On the micro level, human capital theory provides many valuable insights into individual behavior regarding education, occupation, mobility, and health. On the macro level, it facilitates understanding of growth patterns across countries. Given the key theoretical and practical role of human capital concept, in this study I apply the basic framework developed by Becker ([1964] 1993a), Mincer (1958 and 1974) and others to Ukraine and Russia. Modifying the methodology of Mulligan and Sala-i-Martin (1995a) and Bils and Klenow (2000a), I construct a feasible index of human capital stock per capita. On the basis of the State Committee of Statistics of Ukraine Household Survey and the Russia Longitudinal Monitoring Survey data sets, I calculate a series of values of the index to determine and compare the allocation of human capital by sector of employment, economic status, and gender in 2000. I find that in both countries education and science sectors have the highest human capital stock per capita and trade sector has the lowest one. Entrepreneurs and self-employed are relatively less skilled than the employees, a pattern similar to that of developing economies. The unemployed are the least skilled and women, in most cases, have slightly higher human capital stock per capita than men.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I wish to express my deep gratitude to my advisor, Prof. Ghaffar A. Mughal, for his invaluable advice and careful guidance throughout the work on this thesis.

I thank Prof. Jerzy D. Konieczny, Prof. Jonathan Willner, and Prof. Charles N. Steele for their helpful suggestions and comments on the early drafts of the paper. I am grateful to Prof. Roy J. Gardner and Prof. Stefan Lutz for making thorough reviews of the final version of the thesis.

Much of the work would not have been possible without assistance of Dr. Frances L. Dancy and other administrative officers of Russia Longitudinal Monitoring Survey (Carolina Population Center, The University of North Carolina at Chapel Hill), who kindly agreed to disclose to me RLMS confidential data.

I am also highly indebted to Prof. Anil K. Bera, The University of Illinois, for drawing my attention to several insightful articles on testing the specification of Tobit II models.

GLOSSARY

Firm-specific human capital. Part of human capital that refers to knowledge about the structure and functioning of a particular firm and is acquired mostly through learning-by-doing and experience.

General human capital. Part of human capital that is acquired through 'general' training (e.g., ability to read and write, knowledge of foreign languages, communication skills etc.).

Human capital. Stock of abilities, skills, and knowledge embodied in workers (including 'bad habits', degree of mobility etc.), the acquisition of which is, in principle, under the control of an individual, as opposed to innate qualities.

Learning-by-doing. A specific type of human capital accumulation that "occurs in part not as a results of deliberate efforts, but as a side-effect of conventional economic activity" (Romer 2001, 120).

Chapter I

INTRODUCTION

Education is primarily an investment good. Once highly questioned, this statement has by the end of the twentieth century won almost unequivocal recognition from the laity. At the same time, human capital theory, which focuses on investment into various skills and abilities, including education, of a particular worker and labor force in general, has gained many supporters within academia.

The apparent success of the theory stems from its fruitful utilization both on the micro and macro levels. On the micro level, as noted by Yoram Ben-Porath 35 years ago, its "application ... to decisions on individual improvement, and in particular improvement of earning capacity, has provided a framework for the understanding of many aspects of observed behavior regarding education, health, occupational choice, mobility, etc., as rational investment of present resources for the purpose of enjoying future returns" (Ben-Porath 1967, 352).

On the macro level, human capital theory facilitates understanding why growth and development patterns vary across countries. Higher average human capital stock is unambiguously associated with larger output per worker and income per capita (Romer 2001). Besides this traditional, pure level effect, human capital stock may have a strong positive association with the growth rate itself, since, as proposed by many researchers, "human capital speeds the adoption of technology" (Bils and Klenow 2000a, 1162). Technically, human capital accumulation may accelerate the expansion of the knowledge frontier, a particular economy is faced with.

The above considerations alone make human capital an issue worth exploring *per se*, irrespective of the country a researcher focuses on. The transitional economy's context, nevertheless, may well enhance the theoretical and practical significance of such a study in the case of Ukraine and its neighbor, Russia.

It should be emphasized, in advance, that it is not the scarcity of relevant literature that makes the research an especially challenging intellectual endeavor. Apart from the creation of new and fundamental reconstruction of old institutions coincident in time with many other major or minor changes, transitional economies are undergoing "the great human capital reallocation" (Sabirianova 2000). Sabirianova states that this phenomenon is characterized by (but not limited to) a discontinuous increase in occupational mobility, massive job destruction and creation, dramatic swings in demand for various skills, and reallocation of human capital towards market-oriented occupations.

"Building a new market economy makes the issues of skill transferability, worker career adjustment, and returns to investment in previous occupations especially important." (Sabirianova 2000, 5). Over time, the degree of matching achieved on aggregate between the skills needed to accomplish jobs and those of employees may be the key to success in the transitional process. A prolonged and sizable mismatch implies that scarce human resources are wasted, thus, ruling out efficient performance of the economy in the long run.

The study of complex phenomena of human capital formation and allocation may produce no less valuable insights and policy suggestions than research focusing on institutional change in a transitional economy. These two are most likely tightly interrelated in ways not yet fully explored in the literature. Existing rules and norms of behavior definitely constrain individual decisions about acquisition of skills, degree of occupational mobility, etc. For instance, primary and secondary education may be compulsory for all children (as it is in most developed and many developing countries). Certain types of education may be regarded as inferior and suited only to the less "intelligent" and less diligent youths (e.g., vocational training in FSU), whereas others (in particular, complete higher education) are highly desirable, insofar as they may be raising the social status of the recipient and his or her relatives and friends. Simultaneously, human capital accumulation and allocation may play a non-negligible role in the creation, change, and maintenance of the institutional framework.

Having recognized the interdependence between human capital and the institutional framework as well as the complexity of human capital formation and change in allocation over time, this thesis seeks to focus on the static aspect of human capital allocation. In particular, I try to determine and to compare contemporary allocation of human capital stocks per capita by sector of employment, economic status, and gender of labor force members in Ukraine and Russia. To do this, I accomplish the following three intermediate tasks. First, I reexamine the current theoretical specifications of human capital and develop a feasible index of the human capital stock per capita. Second, using data from Ukrainian and Russian national representative surveys conducted in 2000, I estimate returns to education and experience, which are the two primary components of an individual's human capital. Third, I calculate human capital indices that take into account the composition of Ukrainian and Russian labor forces. The data at hand are sufficient for the three tasks.

Understanding the contemporary allocation of human capital stock is essential for assessing the degree of the skill mismatch and identifying potential obstacles to the economy's future development. Hence, this study allows me to draw policy recommendations in the form of general suggestions regarding the spheres where "corrective effort" may be most needed. I deliberately abstain from specifying any exact measures, as this should be done only after a thorough cost-benefit analysis requiring months (if not years) of additional research.

The rest of the thesis is organized as follows. In Chapter II, I give an overview of the literature on the issue. In Chapter III, I discuss the earnings function and two particular specifications of human capital stock per capita and develop a feasible index of human capital. In Chapter IV, I accomplish an empirical estimation of the allocation of human capital stock and discuss the results. In Chapter V, I conclude.

Chapter II

LITERATURE REVIEW

Though "the idea that acquisition and development of skills embodied in human agents of production could be treated as a form of investment originated in the works of W. Petty, A. Smith, and A. Marshall" (Nesterova and Sabirianova 1998, 6) and such constituents of the stock of human capital as education and health of the labor force have probably never been completely out of the focus of professional economists in the 19th and first half of the 20th century, a more or less rigorous academic study in the sphere of human capital began only in late 1950s. It was pioneered by Ted Schultz, Jacob Mincer, Milton Friedman, Sherwin Rosen, and several others associated with the University of Chicago (Becker 1993a, 15).

One of the classic references is *Human Capital* (1993a) by Becker. The study clarifies and firmly establishes the very concept of human capital as sharing many important features with physical capital. Expenditures on education, training, medical care, etc. are undertaken to raise earnings and improve health and quality of life. Hence, they are similar to business investments in new machinery and equipment and personal financial investments aimed at generating a flow of future income and other useful outputs. However, human capital investment is distinct from its physical or financial counterpart, for a person cannot be separated "from his or her knowledge, skills, health, or values the way it is possible to move financial or physical assets" (Becker 1993a, 16). Becker also provides a solid theoretical foundation for the notion of "returns to education".

One of the initial applications of human capital "ideas" directly to the distribution of labor income was accomplished by Mincer (1958) almost 6 years before the first edition of Becker's classic (1964). Mincer's approach relating earnings to schooling took precedence of the simple regression analysis of Becker and Chiswick (Becker 1993a, 104):

$$\ln E_i = a + \sum_{j=1}^{q_i} \bar{r}_j^{'} S_j + v_i^{'} , \qquad (1)$$

where $\ln E_i$ denotes the natural log of earnings of person *i*, *a* is natural log of earnings of a person without any schooling, $\bar{r}_j^{'}$ is the average rate of return to one year of schooling at jth educational level, $v_i^{'}$ denotes the effect of individual-specific characteristics (implying deviations from the average rate of return), S_j is the number of years studied at jth level, q_i stands for the number of (consecutive) educational levels at which person *i* studied and $\sum_{j=1}^{q_i} S_j = S_i$ is the total number of years of schooling of person *i*.

Later, Mincer (1974) extended Becker – Chiswick earnings function (1) via inclusion of "a crude but very useful measure of on-the-job training and experience: years after finishing school" (Becker 1993b, 393), which is currently conventionally referred to as the quadratic experience term. A careful derivation of the earnings function is provided in Chapter III.

Mincer's approach proved to be particularly suitable for econometric estimation of the rate of return to education and has been subsequently and successfully used in *ad litteram* hundreds of studies. Apart from two EERC MA theses by Leschenko (2001) and Shyshkina (2001), one of the most recent estimations of the returns to human capital in a transitional economy was conducted by Nesterova and Sabirianova (1998) on the basis of micro data from the Russia Longitudinal Monitoring Survey. The researchers employ two variations of the earnings function, the simple one (where the rate of return to education, β_1, is assumed to be constant for all years of schooling):

$$\ln W = \beta_0 + \beta_1 \cdot SCH + \beta_2 \cdot EXP + \beta_3 \cdot EXP^2 + \beta_4 \cdot TEN + \beta_5 \cdot TEN^2 + \varepsilon\,, \quad (2)$$

with *W* denoting real earnings, *SCH* denoting number of years needed on average to obtain the highest degree reported, *EXP* being years of "general" experience (conventionally calculated as age minus *SCH* minus 6), and *TEN* standing for years of work for the current employer (the squares of respective terms are included to capture the concavity of the observed earnings profiles); and the extended one (which allows for different rates of return to, for instance, a year of university-level education and a year of secondary school education):

$$\begin{aligned}\ln W = \beta_0 + \beta_1 \cdot UNIV + \beta_2 \cdot TECH + \beta_3 \cdot PTU + \beta_4 \cdot SEC + \beta_5 \cdot EXP + \\ + \beta_6 \cdot EXP^2 + \beta_7 \cdot TEN + \beta_8 \cdot TEN^2 + \varepsilon\end{aligned}\,, \quad (3)$$

where *UNIV* is a dummy for a university degree, the same for *TECH* (technical school), *PTU* (vocational school), and *SEC* (secondary school). According to Psacharopoulos (1995, 8), the rates of return may subsequently be calculated by dividing the respective coefficients by the average number of years needed to complete the relevant educational level. In this thesis, Nesterova and Sabirianova's findings are used as a benchmark to estimated rates of return for Russia.

The "Mincerian" method allows for eliminating the effect on earnings of innate abilities, which are *not*, by definition, a part of human capital (e.g., via adding into regression indices of inborn health or cognitive ability) and the effect of possible discrimination on the basis of gender or ethnicity (by adding respective dummies). The resulting rates of return should represent the "true valuation" of human capital by the labor market under the assumption that no year of education is over- or undervalued). The "Mincerian" approach also makes the calculation of the wage of totally unskilled possible (even if no zero-schooling worker has been registered in the labor market). This proves useful in obtaining the labor-income-based measure of human capital, which is discussed below.

Despite its unquestionable convenience, the "Mincerian" method has two major drawbacks. First, the approach has been shown to assume "flat age-earnings profiles for different levels of education" (Psacharopoulos 1995, 8) or, equivalently, a *constant mark-up in wage* of a university graduate over, e.g., a secondary school graduate of the same age irrespective of years of experience (as interaction between schooling and experience is seldom easy to implement). Second, by construction, "Mincerian" returns disregard any direct costs incurred by an individual or society in the process of obtaining education since the only cost considered is the opportunity cost of time (the wage an individual would have earned, had he/she worked and not studied).

To overcome both shortcomings define the rate of return to education as (Psacharopoulos 1995, 2): *the rate of discount (r) that equalizes the stream of discounted benefits to the stream of costs at a given point in time.*

For instance, the private rate of return to university education could be obtained from :

$$\sum_{t=6}^{47} \frac{\left(W_u - W_s\right)_t}{(1+r)^t} = \sum_{t=1}^{5} \frac{\left(W_s + C_u\right)_t}{(1+r)^t}, \qquad (4)$$

where W_u denotes average wage (net of taxes) of a university graduate, W_s is average wage (net of taxes) of a secondary school graduate, and C_u is direct cost (fees etc.) of an individual while studying at a university respectively in each relevant period, given that university education lasts for 5 years and after graduation a person on average works for 42 years.

The social rate of return is defined similarly and differs only in that C_u includes all costs (incurred by an individual as well as covered by state budget outlays) and wages are in gross "form".

Psacharopoulos (1995, 4) argues that "a key assumption in a social rate of return calculation is that observed wages are a good proxy to the marginal product of labor", hence, similar to the rates of return from the "Mincer-type" regression, the adequacy of *r* in (4) hinges on the degree of competitiveness of the labor market. This considerably limits the applicability of both methods to the indirect assessment of the stock of human capital via its valuation in the labor markets in transitional economies since perfect competition is rarely observed there. High underemployment in transitional economies and consequent under-compensation of human capital also casts a shadow on the potential regression results. However, as long time-series data are not available and projections for future real wages raise serious doubts, cross-sectional Mincer-type regressions remain the only feasible tool for estimation of the rates of return in such transitional economies as Ukraine and Russia.

On the other hand, the availability of the world rates of return to different levels of education determined by Psacharopoulos (1994 Table 2) as r_{prim} = 20.0 percent, r_{sec} = 13.5 percent, r_{higher} = 10.7 percent gives the possibility to indirectly assess whether labor markets in transitional economies are functioning worse than world average.

The review above of the "returns-to-education" literature has been motivated by the exceptionally important role of the rates of return in human capital specifications employed in the literature. The remaining part of this chapter presents an overview of the traditional methods of evaluation of human capital stock, followed by a brief discussion of several modern approaches (analyzed in detail in Chapter III).

Attempts to assess the stock of human capital have been conducted primarily within the field of growth literature as economists, discouraged by the low predictive power of the standard Solow model, inserted human capital into the aggregate production function. Since both cross-country and cross-time data on stocks of human capital had not been available, researchers tried to create an index or a proxy for human capital by eliciting relevant information from the data-sets at hand.

As summarized by Woessmann (2000), commonly employed proxies included adult literacy rates, school enrollment ratios (gross and net), and average years of schooling calculated by either of three methods: perpetual inventory, projection, and attainment census. As the first two methods are based on assumptions that do not hold during periods of rapid demographic and economic transformation, the more or less satisfactory measure is average years of schooling (of the working-age population) obtained by attainment census method. However, other proxies are still used in empirical studies (e.g., Mankiw, Romer, and Weil (1992) approximated human capital by the secondary school enrollment ratio).

Several ambitious projects aimed at determining the level of educational attainment on a global level have already been accomplished. The best known is Barro and Lee's study (2000) with a data-set covering most countries of the world up to 1990 (with projections for 1995 and 2000).

However, average years of schooling *per se* is a poor proxy (especially, if human capital stock is compared across countries) for two main reasons: decreasing returns to education and differences in the quality of educational systems. Thus, instead of using the already available index, attempts have been undertaken in the literature to develop very rigorous specifications of human capital stock that would overcome the drawbacks of traditional proxies.

A distinctive study of this kind has been conducted by Mulligan and Sala-i-Martin (1995a) who develop the so called Labor-income-based measure of the value of human capital (LIHC). The authors show how to employ a micro-level estimation to obtain an index of the human capital stock per capita, which properly takes into account decreasing returns and may be used for cross-time and, in principle, cross-country comparisons. Detailed discussion of LIHC is given in the next chapter.

Another way of assessing the value of human capital per capita is suggested by Bils and Klenow (2000a). As will be evident from Chapter III, this approach is conceptually different from LIHC in defining the human capital stock *per capita*, although the estimation methodology is the same as the one employed by Mulligan and Sala-i-Martin. Also, as discussed below, Bils and Klenow approach eliminates the deficiencies of LIHC, in particular, its improper treatment of the non-human-capital determinants of wage.

Consequently, in this thesis, I follow Bils and Klenow's approach and modify the suggested index taking into account data constraints so as to assess the human capital stock per capita of relevant labor force groups.

Chapter III

HUMAN CAPITAL SPECIFICATIONS

Modern specifications of human capital stock either directly employ the rates of return to education obtained via empirical estimation of the earnings function, or indirectly rely on this estimation technique to calculate the compensation for the raw labor. Hence, in this chapter I, first, discuss the theoretical background of the earnings function and show the derivation of its simplest version, distinguish between the concepts of conditional and unconditional rates of return to education and experience, explore two particular specifications of human capital and trace their link to the earnings function, and, finally, derive a feasible index of human capital stock per capita taking into account the limitations of data sets at hand.

Earnings Function

This section draws heavily on works by Becker (1993a), Chiswick (1997), and Woessmann (2000).

Let us assume that occupations are homogeneous, information is perfect, adjustment costs are absent, and wages are set one time only. For simplicity of exposition, ignore post-school investments in education, learning-by-doing, general and firm-specific experience, and other components of human capital, besides formal schooling (this assumption is subsequently relaxed). Suppose, in addition, that individuals regard education primarily as an investment good (a way to improve one's life-long earnings profile), in other words, considerations about the utility potentially obtained in the process of studying play a negligible role in personal decisions about acquiring additional years of training.

Define w_0 as the wage of "average" unskilled worker, i.e., compensation for raw labor. Then, for each particular individual let:

W_t be the wage received after obtaining t years of schooling,

C_t be the amount of investments in year t of schooling (both indirect, forgone earnings, and direct, e.g., tuition), measured in same units as wage,

$R_t = (r_t + q_t)$ be individual-specific rate of return on investment in year t of schooling, consisting of average rate of return across all individuals (r_t) and individual deviation from this average (q_t),

$K_t = C_t/W_{t-1} = (k_t + m_t)$ be individual-specific ratio of investment in year t of schooling to forgone earnings, consisting of relevant society's average (k_t) and individual deviation (m_t).

In addition, let X "capture" the effect on wages of everything that is beyond the control of an individual and *does not depend* on education, in such a way that:

$$W_0=Xw_0. \tag{5}$$

Apparently, this specification explicitly rules out homogeneity of individuals and for the sake of tractability disregards the possible (*additional*) effect of "interaction" between X and education on wages. "Advantaged" individuals ($X > 1$) obtain higher earnings than compensation for raw labor, "disadvantaged" ones ($0 < X < 1$) get less.

Assuming that the individual is a welfare maximizer, she will pursue the first year of schooling if costs are less or equal to the present value of the stream of benefits from acquiring education. Given the definition of the rate of return to education, the individual completes one year of schooling if there exists R_1 ($0 < R_1 < 1$) such that:

$$C_1 = \sum_{z=1}^{D} \frac{W_1 - W_0}{(1+R_1)^z} \approx \frac{W_1 - W_0}{R_1}, \qquad (6)$$

where z sequentially indexes years after completion of schooling till retirement (D is total number of years of "economic" life).

Consequently, the wage is:

$$W_1 = W_0 + R_1 C_1 = W_0 + R_1 K_1 W_0 = W_0(1 + R_1 K_1). \qquad (7)$$

For two periods:

$$W_2 = W_1 + R_2 C_2 = W_0(1 + R_1 K_1)(1 + R_2 K_2). \qquad (8)$$

By mathematical induction:

$$W_S = W_0 \prod_{t=1}^{S} (1 + R_t K_t), \qquad (9)$$

where S is the total number of years of education.

Taking natural logs and employing Taylor's expansion ($\ln(1+e) \approx e$ for small e):

$$\ln W_S = \ln W_0 + \sum_{t=1}^{S} \ln(1 + R_t K_t) \approx \ln W_0 + \sum_{t=1}^{S} R_t K_t . \qquad (10)$$

Separating the individual-specific effects and denoting them jointly as V:

$$\begin{aligned} \ln W_S &\approx \ln W_0 + \sum_{t=1}^{S} R_t K_t = \\ &= \ln w_0 + \sum_{t=1}^{S} r_t k_t + [\ln X + \sum_{t=1}^{S} (r_t m_t + k_t q_t + q_t m_t)] = \ln w_0 + \sum_{t=1}^{S} r_t k_t + V \end{aligned} . \qquad (11)$$

The latter can be further simplified so as to become suited for econometric estimation (where V is, as a rule, modeled as the error term). Suppose that returns do not vary with the years of education within a particular educational level and the costs of schooling are exactly equal to forgone earnings (implying $k = 1$). Then, (11) can be rewritten in a way closely resembling (1):

$$\ln W_S = \ln w_0 + \sum_{j=1}^{q} r_j S_j + V \,, \qquad (12)$$

where S_j is the number of years of schooling at educational level j, r_j is respective return, and q stands for total number of (consecutive) educational levels at which a person studied.

Finally, relaxing the assumption that formal schooling is the only component of human capital and adding the quadratic experience term, we arrive at the classic Mincer regression:

$$\ln W_{S,Exp} = \ln w_0 + \sum_{j=1}^{q} r_j S_j + \gamma_1 Exp + \gamma_2 Exp^2 + V \,, \quad (13)$$

where *Exp* denotes total years of labor-market experience, and $\gamma_1 \geq 0, \;\; \gamma_2 \leq 0$ are supposed "to capture" positive, but decreasing returns to experience.

Now, let me point out several problems pertaining to specifications (12) and (13).

Chiswick (1997, 5-8) argues that the assumption $k = 1$ (note the link between (11) and (12)) is only occasionally true. Thus, respective regression coefficients must not be interpreted as the rates of return to education, for they are rather returns to the product, rk. If the amount of investment is less than the opportunity cost, $k < 1$ (e.g., if the student worked while studying or received a stipend), the rate of return is underestimated, vice versa for the case of significant direct costs. Unfortunately, data on k values is seldom available even in developed economies and in most empirical studies k is set equal to 1.

Another considerable problem stems out of the potential interdependence between V and S_j's. This requires a closer look at the structure of the error term.

In principle, V should encompass everything that may influence wage, but is *beyond* deliberate *control* of an individual. In addition, V must include characteristics of an individual that are under his or her deliberate control, but which are acquired without paying particular attention to the fact that they may influence wage. Conceptually, I suggest to view V as consisting of:

- innate abilities;
- "static" characteristics;
- "dynamic" characteristics;
- purely random effects.

Innate abilities refer to such individual features as intelligence, in-born health condition, etc. "Static" characteristics may include gender, race, family background, and, in principle, age. Also, in the case of low labor mobility and regional peculiarities of labor market development or economy's restructuring, these may encompass place of residence of an individual and enterprise-specific characteristics (e.g., type of ownership). In terminology of Michael Spence (1973), "static" characteristics are nothing else, but indices. "Dynamic" features include those characteristics, which are under deliberate control, but still, do not refer to human capital directly. The best example is marital status or the status of a husband, presumably, influencing wage through signaling about "favorable" innate abilities, reliability, etc.

Purely random effects term is self-explanatory. It may refer to "luck" and, by construction of Becker – Chiswick earnings function ((11), (12)), includes individual specific deviations in return to human capital, which, on the aggregate, sum up to zero[1] and are independent of education and experience. This implies the orthogonality of purely random effects to schooling and experience, but does not mean that V *as a whole* is independent of the regressors.

According to the theory, education may be playing the role of a signal (Spence, 1973). If those who have better "innate abilities" and "favorable" "static" and "dynamic" characteristics strive for education (in terms of the model, COV(V, S_j) > 0 for at least some j), then standard OLS assumptions are not satisfied (V is not orthogonal to regressors) and estimated coefficients are biased upward.

Chiswick (1997) identifies one additional possibility for endogeneity pertaining to many empirical studies. "When hourly earnings is the dependent variable, whether observed or constructed, variables for weeks worked per year and hours worked per week are generally not included in the regression equation. Then, to the extent that [the elasticity of earnings with respect to time units worked] does not equal unity and the time worked variable is correlated with an included variable, the coefficient of the included variable is biased." (10).

The earnings function can be readily extended to include other variables that influence wage. Conceptually, the extensions differ along two dimensions:

1) in what detail they specify human capital component of wage;
2) how they attempt to control for "innate abilities" component.

The second dimension implies the inclusion of dummy variables for "static" and "dynamic" characteristics and application of IV estimation to control for "innate abilities bias" (as OLS estimates are thought to be biased upward, i.e., that they *overestimate* the effect of human capital on earnings). As discussed in Verbeek (2000), some commonly employed instruments are month of birth,

[1] This follows directly from the definition of R_t, r_t, and q_t above.

parents' education, residence near college. Unfortunately, IV estimation is not feasible given the limitations of Ukrainian and Russian data sets.

However, as shown by Ashenfelter and Krueger (1994), "innate abilities bias" is not a serious problem. On the contrary, researchers have demonstrated that measurement errors lead to *underestimation* of rates of return. So, the effects may tend to canceled each other and if IV estimation is not feasible, OLS should not be disregarded.

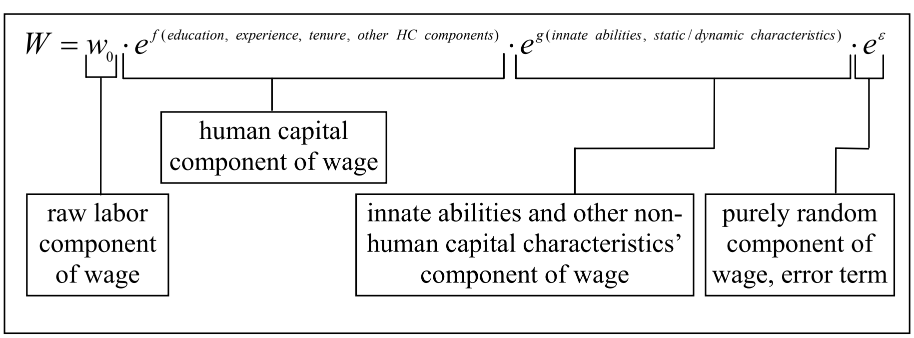
In most general case, the earnings function of a particular individual is:

$$\ln W = f(w_0, O_1, ..., O_L, V_1, ..., V_K), \qquad (14)$$

where O_l denotes the l^{th} characteristic of the individual pertaining to his/her human capital (e.g., number of years of secondary and higher education, experience, index of "acquired health", etc.) and V_k stands for the k^{th} characteristic referring to innate abilities and qualities (for instance, inborn health condition, intelligence, family background, etc.). Exact specification of the functional form and estimation depend upon the data at hand.

Summarizing, in the light of the earnings function wage may be conceptually viewed as shown in Figure 1[2].

Figure 1. Conceptual View of Wage in the Light of the Earnings Function



[2] *e* denotes the base of the natural logarithm.

Conditional vs. Unconditional Returns

Apparently, to determine the rates of return to education and experience both Becker-Chiswick earnings function, (12), and Mincer earnings function, (13), implicitly assume that an individual is employed and earns a positive wage. Otherwise, the natural logarithm of earnings is undefined ("equal to" minus infinity). Hence, many researchers focus primarily on employed and calculate the rates of return on the basis of data for this sub-sample. However, it would be incorrect to claim that the obtained returns hold for the whole labor force, as employed and unemployed may have very different characteristics (including, but not restricted to differences in schooling and acquired experience). Moreover, the rates of return obtained through application of OLS or IV are unable (and not designed) to capture the effect of education and experience on *probability* of being employed.

The real-world phenomenon of unemployment gives rise to the concepts of "conditional" and "unconditional" rates of return, which are implicitly distinguished between in many empirical studies, but not named as such thus far. A simple example may clarify the issue.

Suppose, that better educated and more experienced workers have higher probability of finding a job if currently unemployed or higher probability of staying employed, otherwise. Then, for an individual with certain skills, expected wage is:

$$E(W) = (w_0 \cdot e^{HC} \cdot e^{NHC} \cdot e^{\varepsilon}) \cdot \mathrm{Prob}(employed) + 0 \cdot \mathrm{Prob}(unemployed)\,, \quad (15)$$

where *HC* refers to human capital component of wage, *NHC* stands for "innate abilities", "static", and "dynamic" characteristics' component of wage, and probabilities are conditional on schooling and experience.

Then, the effect of education and experience on wages stems from two effects: first, influence on having positive earnings *per se* and, second, influence (increase) on earnings given that a person is already employed. As OLS and IV can "capture" only the second effect, they improperly estimate (in this case, underestimate) the rates of return for the whole labor force since these estimation techniques are subject to *sample selection bias*.

Fortunately, as discussed by Nesterova and Sabirianova (1999), Verbeek (2000), Stata User's Guide (1999), and other authors, there are approaches that properly deal with sample selection bias. In particular, Nesterova and Sabirianova (1999) suggest that expected wages should be substituted for positive wages for employed and zero earnings for unemployed workers and estimation be done for the whole sample. However, if conditional probabilities are not *a priori* known, this approach is not feasible.

An alternative way of dealing with the problem of sample selection bias is the application of Tobit models. Specification (15) gives rise to a censored model (Tobit I) if the probability distribution of positive wages is conditional on the *same* variables as the probability of observing positive wage (being employed) *per se*. This is rather unrealistic (as some variables, in particular, enterprise ownership, may influence only wages, but not the probability of being unemployed[3]), hence, (15) leads to Heckman's selection model (Tobit II). Then, a proper estimation methodology should take into account the probability of being unemployed, given individual-specific characteristics, the probability distribution of (positive) wages given individual-specific *and* enterprise-specific characteristics, and the probability of being employed, given individual-specific characteristics. Heckman's selection model accomplishes this task (and produces consistent and asymptotically efficient estimates) under certain assumptions regarding the distribution of the error terms. These are clarified in Chapter IV together with a preliminary test for difference in conditional and unconditional returns (i.e., whether education and experience significantly influence the probability of being employed).

[3] Note the static aspect of the analysis.

Summarizing, conditional returns measure the effect of education and experience on earnings provided that an individual is already employed. Unconditional returns additionally take into account the impact of skills on observing a positive wage *per se*. Unconditional rates of returns, thus, measure the "full" impact of education and experience on earnings and apply *both* to the sub-sample of employed and to the one of unemployed.

Labor-Income-Based Measure

This section relies directly on two studies by Mulligan and Sala-i-Martin: (1995a) and (1995b).

For simplicity, let us again assume that the only human capital component is schooling and that individuals have same innate abilities, etc. This implies that the only factor influencing productivity is schooling. Now, let aggregate human capital in an economy be *the quality-adjusted sum of the labor of all its citizens*:

$$H = \int_0^\infty \theta(s) N(s) ds \,, \tag{16}$$

where $N(s)$ is the number of people with s years of schooling, and $\theta(s)$ is the efficiency parameter according to which a worker with s years of schooling contributes to the stock.

Next, define the human capital stock per capita as average human capital stock:

$$h = \int_0^\infty \theta(s) \eta(s) ds \tag{17}$$

where $\eta(s)$ denotes the share of individuals with s years of schooling.

Assuming that the aggregate production function is:

$$Q = F(K, H), \tag{18}$$

where *K* denotes the stock of physical capital and *H* stands for the stock of human capital,

and that a worker's marginal product is equal to her wage, Mulligan and Sala-i-Martin (1995a) show that under normalization, $\theta(0) = 1$:

$$\theta(s) = \frac{w(s)}{w(0)}, \tag{19}$$

since $w(s) = \frac{\partial Q}{\partial N(s)} = \frac{\partial F(K,H)}{\partial H} \cdot \frac{\partial H}{\partial N(s)} = F_H \cdot \theta(s)$ and

$$\theta(s) = \frac{\theta(s)}{\theta(0)} = \frac{F_H \cdot \theta(s)}{F_H \cdot \theta(0)} = \frac{w(s)}{w(0)}.$$

Hence, the human capital stock per capita is:

$$h = \int_0^\infty \theta(s)\eta(s)ds = \int_0^\infty \frac{w(s)\eta(s)ds}{w(0)}, \tag{20}$$

which implies that the human capital stock per capita is the ratio of average wage to the wage of an unskilled worker.

Alternatively, the human capital stock per capita, as follows from specification (17) and normalization, is "the amount of zero-schooling equivalents available in economy" (Mulligan and Sala-i-Martin, 1995a, 11), or the ratio of average productivity to productivity of an unskilled laborer.

Since Mulligan and Sala-i-Martin suggest that *w(0)* be estimated via Mincer-type regression (controlling for gender, race, marital status, and residence in a metropolitan area), in terms of the earnings function:

$$h = \frac{1}{N}\sum_{i=1}^{N}[e^{f(HC_i)} \cdot e^{g(NHC_i)} \cdot e^{\varepsilon_i}] \qquad (21)$$

where *HC* stands for the human capital components of wage (education and experience) and *NHC* denotes innate abilities, etc., ε is the error term, and *f()* and *g()* are some functions.

In my opinion, three particular aspects of the authors' methodology do not seem to be well substantiated. First and apparent, is the wrong treatment of individual characteristics, which are not, by definition, a part of human capital. For simplicity, suppose that the regression includes only one dummy, for gender (say if male then d = 1) and the coefficient before the dummy is significant. Then, irrespective of the labor force composition, the ratio of average wage (which is also an "average" over the sex characteristic) to the exponent of the constant term would not correctly measure the average stock of human capital. The reason for this is that the constant term captures the wage of an unskilled woman. If wages of men and women are ceteris paribus different, then, the higher the proportion of men in the labor force the larger is the departure from the true measure of average human capital (as unskilled men in this specification are implicitly assumed to "have" more or less capital, when they, in fact, do not). The same argument holds for dummies for race, marital status, and the place of residence (under low labor mobility).

Second, instead of using the average wage obtained as the exponent of the fitted value of the regressand (on average values of all explanatory variables), Mulligan and Sala-i-Martin take the per capita labor income for *the whole population*. The latter cannot represent the average wage since it is the ratio of total labor income to total population (including the economically passive).

Third, conceptually, the definition of the human capital stock per capita as average human capital stock is not appealing, since human capital is inseparable from the bearer and cannot be "expropriated" and redistributed across all individuals (unlike physical capital).

Bils and Klenow's Approach

Defining the aggregate human capital similar to Mulligan and Sala-i-Martin (1995a), Bils and Klenow (2000a) postulate that *per capita* human capital stock within cohort *a*, where everyone is, on average, *a* years old, has, on average, *s* years of schooling and, consequently *a-s* years of experience, is (after some simplification of the initial formula):

$$h(a,s) = e^{f(s)+p(a-s)}, \tag{22}$$

where *f()* and *p()* are some functions.

It may be immediately inferred that under such specification, the human capital stock per capita is the human capital stock of an average-skilled worker. In terms of a simpler specification of Mulligan and Sala-i-Martin (abstracting from experience):

$$h = h(\bar{s}), \tag{23}$$

where $\bar{s}$ is the schooling of an average-skilled worker (i.e., average schooling).

In terms of the earnings function:

$$h = e^{f(average\ \ HC)} = \frac{wage\ \ of\ \ an\ \ average - skilled\ \ worker}{w(0) \cdot e^{g(average\ \ NHC)}}, \tag{24}$$

since $e^{\bar{\varepsilon}} = 1$.

Apparently, Bils and Klenow approach avoids the problem of incorrect treatment of non-human-capital components of wage.

The only step to be made from (24) to a feasible index of the human capital stock per capita is the specification of the functional form given the constraints of the data at hand. Combining the approaches of Woessmann (2000) and Bils and Klenow (2000a), let the term for education be linear and the term for experience be quadratic. Then, the human capital stock per capita is:

$$h = e^{\sum_{j=1}^{q} r_j \bar{S}_j + \gamma_1 \overline{Exp} + \gamma_2 \overline{Exp}^2}, \qquad (25)$$

where $\bar{S}_j$ stands for average years of schooling at level *j*, $\overline{Exp}$ denotes average experience, and r_j (j=1,...,q), γ_1, γ_2 are "unconditional" rates of returns (discussed above).

Given data limitations, the rates of return themselves can be obtained by estimating:

$$\ln W = \beta_0 + \sum_{j=1}^{q} \beta_{1j} Ed_j + \beta_2 Exp + \beta_3 Exp^2 + \sum_{l=1}^{L} \beta_{4l} Dem_l + \sum_{p=1}^{P} \beta_{5p} Prop_p + \sum_{k=1}^{K} \beta_{6k} Res_k + \varepsilon, \qquad (26)$$

where Ed_j=1 if *j* is the *highest* educational level completed (0 otherwise), *Dem* stands for individual (demographic) characteristics (age, gender, marital status), *Prop* is a dummy for enterprise/organization ownership, *Res* denotes place of residence dummies.

The regressand, peculiarities of estimation techniques, exact specifications of regression equations for Ukraine and Russia, and other related issues are discussed in Chapter IV.

Since dummy variables are included for the highest educational level completed, the rates of return to education may be obtained from the formula:

$$r_j = \frac{\beta_j - \beta_{j-1}}{T_j - T_{j-1}}, j = 2,...,q$$
$$r_j = \frac{\beta_j}{T_j}, j = 1 \quad , \qquad (27)$$

where T_j is total number of years needed to complete j^{th} educational level, and *j-1* is the prior educational level. Identification of proper prior levels is a non-trivial task and depends upon the institutions of the educational system (this task is tackled in Chapter IV).

It is easy to show that the human capital stock per capita in sub-sample *i* is:

$$h_i = e^{\sum_{j=1}^{q} \beta_{1j}\eta_{ji} + \beta_2 \overline{Exp_i} + \beta_3 \overline{Exp_i}^2}, \qquad (28)$$

where η_{ji} is the proportion of individuals with *highest* educational level *j* in sub-sample *i*, $\overline{Exp}_i$ – average years of experience in sub-sample *i*.

For simplicity, let us consider an example. Suppose, in the sample there are n_c individuals with complete secondary education, n_s with specialized secondary education, n_f with fundamental secondary education, and n_p with primary education. Assume that primary education is a prior level to fundamental secondary, fundamental secondary is a prior level both to specialized secondary and complete secondary (i.e., specialized secondary and complete secondary are non-consecutive) and that normally (on average) primary education lasts T_p years and the estimated coefficient of the dummy for primary education is β_p, fundamental secondary: T_f and β_f, specialized secondary: T_s and β_s, complete secondary: T_c and β_c.

It is necessary to show that:

$$\sum_{j\in(p,f,s,c)} r_j \overline{S}_j = \sum_{j\in(p,f,s,c)} \beta_j \eta_j \; . \qquad (29)$$

By definition of the rates of return, it trivially follows that:

$$\sum_{j\in(p,f,s,c)} r_j \overline{S}_j = r_p \cdot \frac{T_p(n_p+n_f+n_s+n_c)}{(n_p+n_f+n_s+n_c)} + r_f \cdot \frac{(T_f-T_p)(n_f+n_s+n_c)}{(n_p+n_f+n_s+n_c)} +$$
$$+ r_s \cdot \frac{(T_s-T_f)(n_s)}{(n_p+n_f+n_s+n_c)} + r_c \cdot \frac{(T_c-T_f)(n_c)}{(n_p+n_f+n_s+n_c)} =$$
$$= \frac{\beta_p}{T_p} \cdot \frac{T_p(n_p+n_f+n_s+n_c)}{(n_p+n_f+n_s+n_c)} + \frac{\beta_f}{(T_f-T_p)} \cdot \frac{(T_f-T_p)(n_f+n_s+n_c)}{(n_p+n_f+n_s+n_c)} +$$
$$+ \frac{\beta_s}{(T_s-T_f)} \cdot \frac{(T_s-T_f)(n_s)}{(n_p+n_f+n_s+n_c)} + \frac{\beta_c}{(T_c-T_f)} \cdot \frac{(T_c-T_f)(n_c)}{(n_p+n_f+n_s+n_c)} =$$
$$= \frac{\beta_p n_p + \beta_f n_f + \beta_s n_s + \beta_c n_c}{(n_p+n_f+n_s+n_c)} = \sum_{j\in(p,f,s,c)} \beta_j \eta_j .$$

To complete the "proof", I need to specify the assumption regarding the distribution of the error term conditional on group. Suppose either of the two holds:

1) more stringent: Distribution(error|sub-sample=*i*)=Distribution(error). It follows that E(error$_i$)=(E(error|sub-sample=*i*))=E(error)=0.

2) less stringent: Distribution(error|sub-sample) approaches asymptotically Distribution(error), i.e., when both sample and sub-sample grow infinitely large, but the sub-sample size is strictly less than the sample. It follows that plim(error|sub-sample=*i*)=plim(error)=0, where plim stands for probability limit.

These are needed to claim that, on average, *individual-specific deviations* from the average rates of return in a sub-sample as a whole are zero (at least, asymptotically) and (28) holds. Note that for the whole sample (28) is true by definition, since the sum of individual-specific deviations there is preset to zero. By analogy with econometrics estimators, if 1) is true, the index of human capital stock per capita is unbiased, if 2) holds, the index is consistent.

Chapter IV

HUMAN CAPITAL IN UKRAINE AND RUSSIA: EMPIRICAL ANALYSIS

Transitional Context

The following issues related to Ukrainian and Russian context are worth noting. First, in the aftermath of the collapse of the USSR, Ukraine and Russia are expected to exhibit extensive mismatches between human capital of job applicants and employees and job requirements. Citizens with special training are heavily employed in blue-collar jobs. This consistent underemployment of skills, i.e., a "misuse" of human capital, most likely, is due to the temporary general disorganization of the economy with labor markets sending poor and distorted signals. Despite its transitory nature, misallocation of skills may be contributing to acceleration of human capital "depreciation" nation-wide.

Second, a considerable part of skills "inherited" by middle-aged citizens in both countries from the past is obsolete or, at least, is not directly useful under market-economy conditions without substantial additional retraining. Thus, empirically estimated low returns to education may be due not only to skill-job mismatch, but also to skill obsolescence. It is, however, not clear whether "wrong" education should be counted as a part of human capital stock or disregarded altogether. Significant data-related problems inevitably arise, should one try to distinguish between the two education "types" (non-obsolete and obsolete) or a continuum of such "types" (very obsolete, moderately obsolete, etc.).

Third, labor markets in Ukraine and Russia are still in the process of development. Direct state regulation (especially, in such sectors as education, healthcare, and culture) is heavily employed and salary "tariff-nets" are present (e.g., in the form of rigid tying of wages to years of work, and various lumpsum or percentage bonuses for long service, knowledge of foreign languages, for holding professorships and other related "titles"), which makes competitive wage determination (i.e., on the basis on productivity, current and not past accomplishments) an unlikely phenomenon in both countries.

Fourth, labor mobility is rather low within and, definitely, across regions. In Ukraine, significant labor migration is mostly limited to flows of relatively unskilled workers (e.g., in the construction sector) from western regions to the capital. In Russia, the situation is slightly different, since well-paid jobs in oil and gas-extracting industries, located primarily in distant regions, are highly attractive. However, overall "spatial" human capital (not labor *per se*) mobility is rather limited.

Fifth, the populations of the two countries have relatively high educational attainments even in comparison to other transitional economies (e.g., Poland). This is probably due to the fact that secondary education has been compulsory and state-financed for already three generations and higher education has become highly desirable.

The above considerations cast a shadow on the applicability of standard methodologies (which explicitly assume perfectly functioning labor markets and implicitly require high labor mobility) to the cases of Ukraine and Russia. To some extent, the "undesirable" effect of low labor mobility and differences in the speed of labor market development across regions in the two countries may be "absorbed" by inclusion into the regression of dummies for the place of residence. Enterprise-specific impact (which is also present under low mobility) may be partially eliminated by inclusion of dummies for ownership type. Sectoral dummies may be another possibility, but, as discussed below, their presence in the regression considerably complicates the interpretation of the results. The same is true for occupational dummies[4], as pointed out by Chiswick (1997). In any case, returns to education and experience estimated through the earnings function are deemed to turn out relatively low, especially, for compulsory levels of education. The overall explanatory power of the regression (measured by R^2) may also be low, although low R^2 is a peculiar feature of most empirical estimations of the earnings function, even in the setup of a developed economy, e.g., in the study by Ashenfelter and Krueger (1994).

Nevertheless, the transitional context creates some interesting possibilities. For instance, it seems valid to suggest that returns to education are higher for employees in private and foreign-owned firms relative to the ones in state-dominated enterprises (these may not be directly owned). Similarly, newly acquired skills may be commanding a higher premium than older education obtained in communist times (thus, earnings may be inversely related to age, though, the primary cause of this relationship may be different, as discussed below). Vintage aspects may also be playing their role: under-supply of certain professions and oversupply of other ones, a legacy of the past, may be causing disproportionate skill premiums. Certainly, practical application of the above possibilities depends upon the limitations of data available.

Now let me turn to the issue of the relationships among the educational levels in Ukraine and Russia, which is indispensable for calculating returns to schooling when estimation is done on dummies for the highest educational level completed. In essence, this is the problem of finding a proper prior level and determining the total number of years to complete each particular level (refer to (27) above).

The systems of education in Ukraine and Russia are very similar and, apart from a minor detail, the existence of the fundamental higher education, bear a direct

[4] which are not implemented in this thesis because of data limitations.

resemblance to the system of the former USSR.

Roughly, all educational levels may be subdivided into eight groups: primary (which normally takes now a total of 3 years to complete, but 40 years ago lasted for 4 years), fundamental secondary (8 years to complete), complete secondary (10), specialized secondary (11-12), vocational training (10-11), fundamental higher (14), incomplete higher (more than a total of 10 years to complete), and complete higher and above (15-16)[5]. Of these, secondary (and free of charge) education is guaranteed by Ukrainian and Russian Constitutions, thus, in essence, fundamental secondary education is compulsory and most people finish secondary education[6] by either studying another two years at school (complete secondary), or going to vocational training schools, or studying at institutions of specialized secondary education. Sometimes vocational training is undertaken by "holders" of complete secondary education (they have to study for a (additional) year, in contrast to (additional) 2.5 – 3 years for fundamental secondary school graduates.

The case of higher (tertiary) education is slightly less obvious. Statistical offices (and national regulations) distinguish among the three following subtypes: fundamental higher (BA or BSc degree, which take a total of 14 years to complete), incomplete higher (reserved for university drop-outs or those who work while studying), complete higher and above[7] (Specialist, MA, or MSc degree, which take a total of 15 to 16 years to complete). It is worth noting that exact number of years of study for respondents with incomplete higher education cannot be *a priori* determined and depends on data at hand.

---

5 The years in parentheses are valid until a generation of citizens who must now study for 12 years to obtain complete secondary education enters Ukrainian and Russian labor forces. This will happen no sooner than in 8 years in Russia and in 11 years in Ukraine.

6 Still, there are individuals who completed only primary schooling or even failed to do so (in the latter case they are classified as uneducated).

7 "Above" refers to those individuals who undertake further studies (post-graduate education).

As a rule, tertiary education is undertaken by "holders" of complete secondary schooling, while it is extremely uncommon for graduates of vocational schools even to apply to universities. In principle, higher education may be pursued by graduates of specialized secondary schools, though upon graduation from these institutions they are already suited for skilled work.

It may be argued that fundamental secondary education (BA degree) is a proper prior level to complete higher (MA degree). Though this is unambiguously true for U.S., in Ukraine and Russia, it is uncommon for universities to offer the applicants with secondary education distinct BA or BSc programs[8]. Most institutions of tertiary education offer Specialist programs, which falls into the "complete higher" category.

Hence, taking into factual evidence described above, for the purpose of tractability, in this study I suggest the following precedence of educational levels. Fundamental secondary and primary education have no prior level (insofar as these are pursued by the uneducated). Fundamental secondary education itself is the prior level to complete secondary, specialized secondary, and vocational training. Complete secondary, in turn, is the prior level to fundamental higher, incomplete higher, and complete higher and above (as this is the most commonly observed phenomenon).

As far as the total number of years to complete a given education level is concerned (T_j's in (27)), it should be noted that, in principle, I may take the years according to current regulations ("normal" number to obtain a particular level). However, it would be better to use actual averages reported by the respondents themselves. First, it is the only feasible way to determine T_j for university drop-outs (incomplete higher) and T_j's for those educational levels where "normal" number falls into a range (lower bound plus half of the range may be a relatively poor proxy). Second and most important, averages may

[8] The National University of "Kyiv-Mohyla Academy" is an exception, though BA and BSc "holders" are almost automatically re-enrolled into respective MA and MSc programs.

partially eliminate conceptual problems when the suggested precedence of educational levels is violated. For instance, when tertiary education has been pursued by a specialized secondary school graduate. For obvious reasons, her total number of years of study must be higher than "normal", thus, the rates of return would be overestimated if the "normal" years of the suggested precedence are employed[9]. Third, averages may help "to capture" and eliminate the upward bias in case of grade repeats.

Table 7 and Figure 4 in Appendix A visualize the discussion above. In particular, it should be noted that total years to obtain most education levels are close to the "normal" totals.

Data and Samples

For Ukraine, the data used in this paper are obtained from "The Survey on Households' Standards of Living" (in particular, the micro file on members of households) conducted by the State Committee of Statistics of Ukraine (Derzhkomstat) in the year 2000.

The database of the Survey covers 25,133 individuals from 9,318 households and contains information on individuals' age, sex, marital status, educational background (the highest completed level and total years of schooling), employment status (economic status), years of working experience, industry and type of employment, enterprise ownership, region and place of living (city, town, or rural area), type of self employment, earnings/income (separate categories for income from primary job, secondary job, self-employment, entrepreneurial activity, stipends, pensions, unemployment benefits, and other types of income), and a dummy for regular physical training. Though the data set includes a category for the field of education, unfortunately, the information

[9] Another uncommon example is when specialized secondary is pursued by a complete secondary graduate.

is missing for almost ninety percent of the respondents who are labor force members.

For Russia, the data are taken from the Russia Longitudinal Monitoring Survey (RLMS), which is conducted according to the World Bank's methodology of Living Standards Measurement Surveys, in particular, from the open-access micro file on members of households surveyed during Round IX in the year 2000, and several confidential files obtained directly from RLMS administrative office (Carolina Population Center, The University of North Carolina)[10].

RLMS database on adults covers 9,074 individuals and contains very detailed information on their education, job, earnings, and other individual characteristics (almost 290 variables).

Apparently, not all data rows in Ukrainian and Russian samples are relevant for the research undertaken in this thesis. Thus, the original data sets have been subjected to cleaning.

First, I have excluded all individuals who are out of labor force, since human capital, by definition, is a characteristic of labor force members only. It is worth noting, however, that I have not deleted the observations of pensioners, students, and pupils who report positive earnings, since they are, in fact, employed.

Second, I have eliminated the few observations that lack important information on education, experience, age, gender, and marital status as these data are indispensable for proper implementation of estimation methodology.

As a result, the final number of observations left in Ukrainian sample is 8838 and in Russian sample is 4681. The number for Ukraine is above the corresponding one (5430) for the cleaned sample of the last year thesis by

[10] The use of the latter files is subject to the data use agreement.

Shyshkina, who used "The Survey on Households' Standards of Living" of Derzhkomstat for the year 1999. The number for Russia is above the corresponding figure (in the range of 3000) in the by Nesterova and Sabirianova (1999), who used data sets of RLMS Round VI and Round VII.

I have encountered some complications in making Ukrainian and Russian data sets compatible with respect to sector of employment and economic status. Ukrainian data set contains the information on the "branch" of employment for employees and type of activity for entrepreneurs and self-employed. Hence, in this case determination of the sectoral composition of the employed is straightforward. On the contrary, Russian data set contains no standardized variable for sector, but, instead, offers a classification of a respondent's job/occupation according to ISCO-88 (International Standard Classification of Occupations), also known as ILO job code. As RLMS data set contains (detailed) four-digit occupation codes, which, in the overwhelming majority of cases, leave no ambiguity regarding, determination of the sectoral composition is also straightforward, though, extremely tedious and time-consuming. Refer to the table in Appendix C for names of the sectors as well as the details how classification has been accomplished.

Unfortunately, as Ukrainian data set contains no information on occupation (e.g., professional, skilled worker, unskilled worker, etc.) of the respondent, assessment of human capital stock conditional on the professional group is not feasible.

As Ukrainian and Russian data sets are by and large somewhat different in the way they record economic status (and many common ones as student, pupil, pensioner, housewife are irrelevant to this research for obvious reason), I was forced to classify all labor force members into three categories: employees, employers/entrepreneurs/self-employed[11], and unemployed.

Place of residence variables, including settlement type (city, town, village) and region, are readily available in Ukrainian data set and easy to restore for Russian sample with the help of data in confidential RLMS files.

Characteristics of Labor Forces

Although the samples for Ukrainian and Russian labor forces are representative, I am still reluctant to claim that respective figures hold exactly for the whole population. Thus, in this section, I focus on the qualitative aspect of labor force composition. In most cases, I do not offer any reason for observing each particular pattern, for it is not the task of the research.

Educational attainment. As evident from Table 8 and Table 9 in Appendix B, Ukrainian and Russian labor forces exhibit more or less similar composition with respect to schooling. Roughly a quarter of respondents in the samples has higher education (and Russia's share is slightly higher, though the fraction of individuals with complete higher education is lower by 2.5 percentage points). The majority of workers has secondary schooling and the number of those with primary education and without any schooling is small. The main difference between Ukrainian and Russian samples is a significantly higher proportion of respondents with vocational training in Russia (more than a quarter) and, respectively, higher proportion of individuals with complete secondary and specialized secondary in Ukraine (half of the respective sample).

---

[11] Note that RLMS provides no means to unambiguously distinguish among the three, in contrast to Derzhkomstat classification (the best example of this kind is RLMS category of farmer, who may belong to any of the three types depending on particular circumstances). In any case, there may be conceptual difficulties in separating employers and entrepreneurs into distinct subcategories taking into account the economic meaning of the subcategory and not the legal definition.

As far as respective figures for the employed and unemployed are concerned, data unambiguously show that the employed in both samples have higher educational attainment, insofar as the burden of unemployment falls on individuals with secondary schooling, in particular, on "holders" of complete secondary and specialized secondary education in Ukraine (more than 35 and 22 percent respectively) and on graduates of vocational training institutions and (complete) secondary schools (almost 27 and 24 percent respectively) in Russia. The unfavorable trend for individuals with fundamental secondary education is also evident, however, it is relatively more striking in the Ukrainian sample vis-a-vis Russian one[12].

The data also reveal that Ukrainian and Russian women are relatively better educated than men. Women have higher share of respondents with complete higher and fundamental higher education and also with specialized secondary schooling (which is superior to vocational training[13]).

Sectoral composition. As apparent from Table 12 and Table 13 (Appendix B), the two countries have more or less similar structure of employment by sector. The main differences refer to agriculture (as expected, Ukraine's share is almost twice as large reaching more than 16 percent), finance (Russian sample's share is almost 5 times higher), and government (3 times higher in Ukraine). Less conspicuous differences can be observed in construction (5 percent in Ukraine vs. almost 9 percent in Russia), transport and communications (4 percentage points higher in Russia), housing (3.5 percentage points lower in Ukraine), and education (10 and 6 percent in Ukraine and Russia respectively). The largest share of employed is in manufacturing sector (almost 24 percent in Ukraine and more than 20 percent in Russia). Gender structures of sectoral composition are similar. In both countries, men are more likely to be employed in

---

12 Those interested in the employment "breakdown" of Ukrainian and Russian labor forces *within* each educational category are advised to refer to Table 10 and Table 11.

manufacturing, agriculture, construction, transport and communications, and defense sectors. Women "dominate" trade, health care, education, finance, and government bureaucracy[14].

Economic status. Employees in both samples constitute approximately three quarters of the whole labor force. Figures for employers/entrepreneurs/self-employed (8-9 percent) and unemployed (16 percent) are also similar. In Ukraine, men are less likely to be unemployed than women (respective shares are 14.6 and 17.3 percent). Russia's sample case is quite the opposite as the proportions of unemployed men and women are 17.4 and 15.6 percent respectively. In both countries, men are somewhat more likely to engage in various types of self-employment.

Hypotheses, Regression Equations, and Estimation Methodology

As argued in Chapter III, it may be true that individuals with higher stock of human capital are more likely to be selected into the pool of employed. This, in turn, implies that the sub-sample of employed in *not* random, and returns to education and experience obtained through application of OLS to data of the employed produces biased and inconsistent estimates of the rates of return for the whole labor force sample (*sample selection bias*).

This proposition should be tested rigorously; and to do so I suggest using standard logistic model:

Prob(employed|individual characteristics)=G(individual characteristics, parameters):

---

[13] There is a long-established tradition in FSU countries of labeling pupils of vocational training schools as being less intelligent and able than "normal" or specialized school ("tekhnikum") pupils. Unfortunately, this popular statement may well reflect the reality.

[14] Probably, not on the top positions.

$$y = \text{logit}(\alpha_0 + \alpha_1 S + \alpha_2 Exp + \sum_{l=1}^{L} \alpha_{3l} Dem_l + \sum_{k=1}^{K} \alpha_{4k} Res_k), \quad (30)$$

where *y*=1 if an individual is employed, 0 otherwise. *S* is the total number of years of schooling (referred to in regression as *sch_years*); *Exp* (*exp*) is the total number of years of working experience; Dem_l's (individual demographic characteristics) are: *age* (in years), *sex* (1 if male, 0 otherwise), *sxmr* (1 if married man, 0 otherwise)[15]; Res_k's (place of residence dummies) are: *city*, *town* (rural place of residence is a benchmark), and capital city, *Kyiv* (for Ukraine), *Moscow* (for Russia)[16].

Note that as this test is, in essence, preliminary, in specification (30) no dummies for educational level are needed. For obvious reasons *S* is highly correlated with educational dummies and cluttering the illustrative estimation with many dummies is unproductive.

The natural hypothesis is whether education and experience significantly influence the probability of being employed. So, if the hypothesis of no influence of these variables is rejected, then, it may be inferred that conditional and unconditional returns are different, therefore, conditional returns do not apply to the whole labor force[17].

---

[15] I have failed to find that marital status per se has any impact on the probability of being employed. Marital status has also been insignificant in all wage regressions, in contrast to the highly significant positive effect of the gender-marital interaction term, status of a married man. Therefore, hereinafter, in all specifications I retain *sxmr*.

[16] Capital oblasts (*KyivOblast* and *MoscowOblast*) are excluded as I have not found them to significantly affect employment status in logistic specification.

[17] In more technical terms, the issue of interest is (less stringent) joint significance of all variables. This may give a hint (but not a decisive conclusion, as the model is logistic, not probit) whether Heckman's lambda is not constant. Constancy of Heckman's lambda (which is only possible in case when all parameters are zero) implies that estimation of Heckman selection model (either by the full ML or by the two-step procedure) makes no sense.

Let me now turn to the issue of conditional returns to education[18]. As noted in Chapter III, the model to be estimated by OLS is:

$$\ln W = \beta_0 + \sum_{j=1}^{q}\beta_{1j}Ed_j + \beta_2 Exp + \beta_3 Exp^2 + \sum_{l=1}^{L}\beta_{4l}Dem_l + \sum_{p=1}^{P}\beta_{5p}Prop_p + \sum_{k=1}^{K}\beta_{6k}Res_k + \varepsilon, \quad (31)$$

where educational dummies (*Ed*) are: *chigh* (for complete higher education as the highest educational level obtained), *ihigh* (incomplete higher), *fhigh* (fundamental higher), *ptu* (vocational training), *ssec* (specialized secondary), *csec* (complete secondary), *fsec* (fundamental secondary), *prim* (primary), and ownership type dummies (*Prop*) are *foreign* (for foreign or joint-stock entities) and *nongov* (domestic non-governmental enterprises or organizations)[19]. The meaning of other variables has been already explained above with the exception that *Res* additionally includes *KyivOblast* and *MoscowOblast* and specification for Russia does not include *sex*[20]. The error term must satisfy the usual OLS assumptions.

To avoid the problem of heteroscedasticity, the estimation of the variance-covariance matrix of coefficients via White/Huber/sandwich formula (heteroscedasticity-consistent, or robust, errors) is preferable.

The primary hypothesis is whether β_{1j}'s are statistically significantly different from zero. Also, these coefficients are expected to be positive and steadily rising for each consecutive level (for returns to education to be positive). β_2 and β_3 should also be significantly different from zero. Moreover, β_2 is expected to be positive, whereas β_3 to be negative (to "capture" decreasing returns to experience). Other variables are of marginal importance for this

[18] I calculate these only for the sake of comparability with other studies. It has been noted that the validity of OLS model itself is dubious.

[19] State ownership is a benchmark.

[20] Several specifications have been tried, in none of which the coefficient of *sex* (for Russian sample) has been significant. Explanation for this phenomenon is put forward below.

research, although they may be of interest if the issue of wage determination is taken into account.

In the light of specification (31), there are several issues worth noting. First, the content of the regressand should be discussed. In *W* I include (as in the majority of other studies) income from primary place of work (salary/wages/dividends), wages from other jobs, income from self-employment activities, and entrepreneurial income. Other forms of income (stipends, pensions, unemployment benefits) are excluded. For Ukraine this is yearly income. For Russia, *W* is monthly income.

Second aspect is the inclusion of ownership type dummies. To some extent these variables may help to "isolate" the influence of enterprise-specific effects given relatively low labor mobility in Ukraine and Russia. Ownership dummies may also "capture" the possible trade-off between wage and career growth opportunity: in state organizations wage may be lower, but career growth opportunities are higher and retirement benefits schemes are better[21]. Therefore, lower wages in state organization understate the "true" (inclusive of other benefits) valuation of human capital. Thus, this specificity of governmental organizations should be "factored out". In foreign enterprises wages are much higher, but this may be due to "luck" of those who got into these enterprises "in time".

Third, employment status and sectoral dummies are excluded. The reason for this is that the inclusion of a dummy for status or sector significantly raises *difficulty in interpretation of the results*. This issue is pointed out by Chiswick (1997) who argues that the coefficient of schooling is not a proper rate of return when any of these dummies is present (the same argument holds for occupational dummies).

[21] This is at the center of many public discussions, where the allegedly discriminatory clauses of the Law on the State Service regarding pensions are fiercely criticized.

The idea is that schooling may be correlated with status or sector (or occupation) and, therefore, the derivative of the earnings function with respect to schooling (i.e., the influence of education on logarithm of earnings, the *true* return) may *not be equal* to the coefficient of schooling:

$$r_{true} \equiv \frac{\partial \ln W}{\partial S} = \frac{\partial (r_{alleged} S + sector)}{\partial S} = r_{alleged} + \frac{\partial sector}{\partial S}.$$

So, if $\frac{\partial sector}{\partial S} \neq 0$, then, $r_{true} \neq r_{alleged}$, which makes the assessment of the rate of return unfeasible.

As also discussed above, the proper estimation technique of unconditional returns to education and experience is Heckman selection model (Tobit II[22]). The first equation of the model is similar to (31):

$$\ln W = \beta_0 + \sum_{j=1}^{q} \beta_{1j} Ed_j + \beta_2 Exp + \beta_3 Exp^2 + \sum_{l=1}^{L} \beta_{4l} Dem_l + \sum_{p=1}^{P} \beta_{5p} Prop_p + \sum_{k=1}^{K} \beta_{6k} Res_k + \varepsilon_1, \quad (32)$$

where, as previously, *Dem* for Russia does not include *sex*.

The second equation is slightly more difficult to construct. Due to specificity of the conventional form of the estimated loglikelihood function, regressors in the selection equation should constitute a subset of regressors in the first equation. Therefore, I preliminarily suggest the following selection equation:

$$y = \delta_0 + \sum_{j=1}^{q} \delta_{1j} Ed_j + \delta_2 Exp + \delta_3 Exp^2 + \sum_{l=1}^{L} \delta_{4l} Dem_l + \sum_{k=1}^{K} \delta_{5k} Res_k + \varepsilon_2, \quad (33)$$

where *y*=1 if individual is employed, 0 otherwise, *Dem* includes all individual (demographic) characteristics (*age, sex, sxmr*), and *Res* includes all place of residence dummies (*city, town,* capital city, and capital oblast).

---

[22] Tobit III is not feasible, as Ukrainian data set contains no information on hours of work (the structure of Tobit III models is discussed by Verbeek (2000, 212)).

To insure convergence of the maximization algorithm (which is a problem given the complexity of the loglikelihood function and large number of observations) several specifications close to (33) must be tried. In particular, I had to exclude dummies for primary and fundamental secondary education in the selection equation for Ukraine, dummy for capital oblast and constant term in the selection equation for Russia.

To close the model, the standard assumption must be made regarding the distribution of the error terms in (32) and (33):

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \sim NIID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & 1 \end{pmatrix} \right): \qquad (34)$$

the vector of the error terms is normally, independently, and identically distributed with zero mean (vector) and respective variance-covariance matrix.

The theory states that the sample selection bias arises if $\sigma_{12} \neq 0$, equivalently, if $\rho = \frac{\sigma_{12}}{\sqrt{\sigma_1^2}} \neq 0$ (non-zero correlation between the errors). In case of no correlation estimation of (32) through OLS would produce consistent estimates and Heckman selection model is superfluous.

Hence, the hypothesis of zero correlation is the most important one to be tested. Other hypotheses (about the sign and significance of the coefficients) are similar to the ones in OLS regression above, (31).

It is known from the theory that consistency of a maximum likelihood estimator crucially depends on the validity of the "distributional" assumption (here, (34)). In particular, the estimator is definitely inconsistent if the functional form of the distribution is incorrectly specified.

Therefore, a proper specification test must identify whether (34) holds for the estimated model. If estimation of Tobit II is conducted by the two-step procedure, then, as argued by Pagan and Vella (1989), the test should, first, check whether the distribution of the error term in the underlying probit model (i.e., marginal distribution of ε_2 from the selection equation) is normal and, second, make an inference about the conditional distribution of ε_1 given ε_2 (which should also be normal). If the estimation procedure is full maximum likelihood[23], the specification test is much more involved. This highly technical issue is discussed in detail in Appendix F.

Given consistent estimates, returns to education and human capital indices may be calculated according to (27) and (28).

Results and Discussion

Logit (Preliminary test for difference between conditional and unconditional returns). As shown in Table 15 and Table 16 (Appendix D), the data supports the proposition than individuals with higher educational attainment and more experience are more likely to be employed in Ukraine and Russia. The estimated regression for Ukraine is (robust standard errors are in brackets):

$$y = \text{logit}\begin{pmatrix} \begin{array}{llllll} -.67 + .21 \cdot sch\_years & + .12 \cdot exp & - .07 \cdot age & - .37 \cdot sex & + 1.05 \cdot sxmr \\ (.257) \quad (.015) & (.008) & (.008) & (.092) & (.102) \\ + .82 \cdot city + .25 \cdot town + .44 \cdot Kyiv & & & & \\ (.081) \quad (.074) \quad (.215) & & & & \end{array} \end{pmatrix}$$

Marginal effects of *sch_years* and *exp* at the mean level of the regressors (i.e., incremental changes in the probability of being employed due to an additional year of schooling or experience for an "average" individual) are 2 and 1 percentage points respectively.

For Russia:

[23] which is more efficient than the two-step.

$$y=\text{logit}\begin{pmatrix} 1.54+.08\cdot sch\_years+.11\cdot exp-.08\cdot age-.60\cdot sex+.61\cdot sxmr \\ (.314) \quad (.017) \quad (.012) \quad (.011) \quad (.123) \quad (.127) \\ +.86\cdot city+.71\cdot town-.56\cdot Moscow \\ (.112) \quad (.101) \quad (.254) \end{pmatrix}$$

Marginal effects of *sch_years* and *exp* at the mean level of the regressors are 1 and 1 percentage points respectively.

Almost all coefficients are very precisely estimated and are statistically significant at 0.1 percent level. The coefficients of *Kyiv* and *Moscow* are significant at 4 and 3 percent levels respectively. As expected, the coefficients before the years of schooling and experience, as well as relevant marginal effects, are positive. A larger impact of schooling on the probability of being employed in Ukraine vis-a-vis Russia is, probably, due to the dissimilarity in the employment "breakdown" of the top educational category in the samples. For instance, in Ukraine, the unemployment rate in the group of holders of the complete higher schooling is almost 6 percent. In Russia, however, more that 10 percent of such labor force members are jobless (Table 10 and Table 11).

Curiously enough, being male *per se* implies lower probability of being employed. However, for married men this "negative" effect of gender is more than offset in Ukraine and marginally offset in Russia.

This phenomenon requires some additional consideration. Data inspection reveals that single males are younger that married males, in particular, roughly 50 percent of single males in both samples are at most 25 years old, whereas the respective figure for married men is only 9 percent. So, many single males fall into the category of "potential conscripts". A rational employer would certainly not prefer to offer a job to a "potential draftee" and risk facing soon a loss of indirect investments into such an employee (e.g., acquired firm-specific human capital). On the contrary, married males may be viewed as more "responsible and reliable" and are, thus, apt to be employed (positive coefficient of *sxmr*).

Other coefficients have expected sign: older individuals (who may have trouble adapting to new requirements in increasingly competitive labor markets in transitional Ukraine and Russia) face unfavorable effect of age on the employment opportunities, residents of cities and town are more likely to be employed than their rural counterparts. Negative coefficient of *Moscow* dummy is rather difficult to explain. Plausible reasons may include a relatively high turn-over rate of the local labor market and traditionally massive migration flow into Russian capital (as migrants are not expected to find jobs instantaneously).

OLS. The results of the estimation are provided in Table 17 and Table 18[24]. The estimated equation for Ukraine is (robust standard errors are in brackets):

$$\begin{aligned}
\ln W = 6.43 &+ \underset{(.074)}{.76} \cdot chigh + \underset{(.061)}{.50} \cdot ihigh + \underset{(.072)}{.68} \cdot fhigh + \underset{(.067)}{.50} \cdot ptu + \underset{(.061)}{.57} \cdot ssec + \underset{(.060)}{.42} \cdot csec \, + \\
&+ \underset{(.060)}{.35} \cdot fsec + \underset{(.064)}{.36} \cdot prim + \underset{(.092)}{.02} \cdot exp - \underset{(.002)}{.0004} \cdot exp^2 - \underset{(.00005)}{.0005} \cdot age + \underset{(.002)}{.13} \cdot sex + \underset{(.025)}{.19} \cdot sxmr \, + \\
&+ \underset{(.026)}{.41} \cdot foreign + \underset{(.075)}{.05} \cdot nongov + \underset{(.013)}{.41} \cdot city + \underset{(.016)}{.28} \cdot town + \underset{(.017)}{.28} \cdot Kyiv + \underset{(.029)}{.21} \cdot KyivOblast \\
&\qquad (.041)
\end{aligned}$$

For Russia:

$$\begin{aligned}
\ln W = 5.78 &+ \underset{(.368)}{1.63} \cdot chigh + \underset{(.330)}{1.30} \cdot ihigh + \underset{(.337)}{1.52} \cdot fhigh + \underset{(.340)}{1.11} \cdot ptu + \underset{(.332)}{1.24} \cdot ssec + \underset{(.331)}{1.05} \cdot csec \, + \\
&+ \underset{(.333)}{.72} \cdot fsec + \underset{(.337)}{.65} \cdot prim + \underset{(.349)}{.07} \cdot exp - \underset{(.007)}{.0008} \cdot exp^2 - \underset{(.00009)}{.04} \cdot age + \underset{(.005)}{.51} \cdot sxmr \, + \\
&+ \underset{(.031)}{.32} \cdot foreign + \underset{(.126)}{.28} \cdot nongov + \underset{(.032)}{.49} \cdot city + \underset{(.039)}{.09} \cdot town + \underset{(.040)}{.21} \cdot Moscow + \underset{(.075)}{.56} \cdot MoscowOblast \\
&\qquad (.064)
\end{aligned}$$

In Ukrainian sample, apart from the insignificant negative coefficient for *age*, all parameters are very precisely estimated and are significant at 0.1 percent level. In Russian sample, most coefficients are significant at (at most) 3.5 percent significance level and the only marginally significant coefficient is that of primary schooling, *prim* (different from zero at 7 percent level).

---

[24] Note that the hypothesis of constant variance of the error term is decisively rejected. Therefore, robust (White/Huber/sandwich) estimator of coefficients' variance-covariance matrix is employed.

F-statistics decisively reject the joint insignificance of the parameters. As expected, the explanatory power of the regression is rather low (R^2 does not exceed 23 percent).

It is worth noting that the coefficients of interest have the expected sign and relative magnitude (e.g., the effect on earnings of the complete higher education is definitely larger than the effect of the complete secondary schooling[25], the coefficient of exp^2 term is negative reflecting the hypothesized diminishing returns to experience). Other coefficients also have expected sign. In particular, married males, workers of foreign and domestic private firms, and residents of urban settlements obtain significant wage premia that are not related to the acquired schooling and experience.

Tobit II (Heckman selection). Results of maximum likelihood estimation of the Tobit II model are presented in Table 19 and Table 20. The estimated wage equation for Ukraine is (respective errors are in brackets):

$$\begin{aligned}
\ln W = {} & \underset{(.063)}{6.13} + \underset{(.044)}{.87} \cdot chigh + \underset{(.059)}{.63} \cdot ihigh + \underset{(.052)}{.77} \cdot fhigh + \underset{(.044)}{.54} \cdot ptu + \underset{(.043)}{.63} \cdot ssec + \underset{(.044)}{.42} \cdot csec + \\
& + \underset{(.047)}{.32} \cdot fsec + \underset{(.083)}{.36} \cdot prim + \underset{(.003)}{.04} \cdot exp - \underset{(.00005)}{.0006} \cdot exp^2 - \underset{(.002)}{.008} \cdot age + \underset{(.026)}{.07} \cdot sex + \underset{(.027)}{.31} \cdot sxmr + \\
& + \underset{(.092)}{.35} \cdot foreign + \underset{(.013)}{.04} \cdot nongov + \underset{(.016)}{.48} \cdot city + \underset{(.017)}{.30} \cdot town + \underset{(.029)}{.30} \cdot Kyiv + \underset{(.040)}{.21} \cdot KyivOblast
\end{aligned}$$

For Russia:

$$\begin{aligned}
\ln W = {} & \underset{(.367)}{5.73} + \underset{(.328)}{1.67} \cdot chigh + \underset{(.335)}{1.34} \cdot ihigh + \underset{(.337)}{1.56} \cdot fhigh + \underset{(.330)}{1.15} \cdot ptu + \underset{(.329)}{1.28} \cdot ssec + \underset{(.330)}{1.09} \cdot csec + \\
& + \underset{(.335)}{.76} \cdot fsec + \underset{(.346)}{.69} \cdot prim + \underset{(.007)}{.07} \cdot exp - \underset{(.00009)}{.0008} \cdot exp^2 - \underset{(.005)}{.04} \cdot age + \underset{(.031)}{.52} \cdot sxmr + \\
& + \underset{(.126)}{.32} \cdot foreign + \underset{(.032)}{.28} \cdot nongov + \underset{(.039)}{.50} \cdot city + \underset{(.040)}{.10} \cdot town + \underset{(.075)}{.21} \cdot Moscow + \underset{(.064)}{.56} \cdot MoscowOblast
\end{aligned}$$

---

[25] The hypothesis of the equality of these coefficients is decisively rejected for both regressions (p-value is 0.0000).

In Ukrainian sample, the coefficients of interest (on educational levels and two experience terms) are significant at .1 percent level, have predicted sign and reasonable magnitude. The remaining coefficients are all significant at 1 percent level (or better). In Russian sample, the coefficients are also estimated very precisely with the exception of the ones for fundamental secondary schooling (different from zero at 2.5 percent level) and primary schooling (significant at 5 percent level). As previously, married males, workers of foreign and domestic private firms as well as residents of urban settlements obtain significant wage premia that are not related to the acquired schooling and experience.

The Wald test of the independence of the first (wage) and second (selection) equations reveals that it is highly unlikely to observe $\sigma_{12} = 0$ (equivalently, $\rho = 0$). This hypothesis of zero correlation is rejected at .01 percent significance level for Ukraine and 5 percent level for Russia, implying that the application of OLS produces biased and inconsistent estimates of the parameters for the whole sample. In other words, the inequality of conditional and unconditional returns to education and experience is established. Conditional returns apply to the sub-sample of employed only and do not reflect the full effect of human capital on earnings.

As argued above in section on estimation methodology and below in Appendix F specification tests for Tobit II models are especially difficult to conduct and are very seldom attempted in empirical studies. Details on one particular test (which aims at checking the validity of the bivariate normality assumption) are provided in Appendix F. The estimated test statistic is 12.397143 for Ukraine and 9.4725440 for Russia. Under the null hypothesis of bivariate normality the test statistics are distributed as χ^2 with nine degrees of freedom. Therefore, respective p-values are .1918347 (Ukraine) and .394845 (Russia), implying that the null hypothesis cannot be rejected. Although the test seem to support the "correctness" of the model specification, the results should be exercised with extreme caution and high degree of skepticism, since the test statistic employs the outer product gradient version of variance-covariance matrix instead of the matrix of second derivatives of the loglikelihood function[26], which may be more correct when the $(1/N)$ term is substituted for the expectation term[27].

The conditional and unconditional rates of return to education and experience calculated according to (27) taking into account all the peculiarities discussed above (in the section on the transitional context) are presented in tables below. Note that these are the rates of return to *one year of schooling* at each particular level.

Conditional returns. As evident from Table 1, Ukraine and Russia exhibit rather dissimilar profiles of the rates of return to a year of schooling at a particular educational level.

In Ukraine, the rates of return are in the range from 2.3 percent (incomplete higher education) to 9.4 percent (primary) and the rate of return to a year of schooling at the higher level (complete higher) is 6.2 percent. The rates of return do not exhibit a theoretically predicted declining trend. This is, however, not unprecedented, as demonstrated in the studies of the education system of Pakistan by Dr. Shahrukh Rafi Khan.

In Russia, the rates of return are 1.6 – 5 times higher and fall into the range from 6.9 percent (incomplete higher) to 19.1 percent (primary, although this is significant only at 7 percent level) with the rate of return to complete higher education being 10.2 percent. The difference in the patterns is particularly startling for complete secondary schooling, the rate of return to a year of which is 18.5 percent in Russia, but only 3.8 percent in Ukraine. Moreover, apart from the "troughs" for fundamental secondary and incomplete higher education, Russian rates of return exhibit a declining trend. The phenomenon of the "fundamental secondary schooling trough" may be explained by the fact that

[26] famous Cramer-Rao Lower Bound Theorem (Green 2000, 106).

[27] Some discussion is provided by Verbeek (2000, 165).

secondary education is compulsory, thus, as most labor force members have completed this level, fundamental secondary schooling cannot command a large mark-up *per se*. Negative signaling effects on wages for individuals with this type of education and, also, with incomplete higher education (i.e., university drop-outs), vis-a-vis holders of the complete secondary and complete higher education respectively, are, probably, present as well.

Table 1. Conditional Rates of Return to Education and Experience, Ukraine and Russia, percent*

Description	Ukraine	Russia
Complete higher and above	6.2463	10.2116
Incomplete higher	2.2656	6.8956
Fundamental higher	6.6225	11.8557
Vocational training	4.7703	12.9795
Specialized secondary	5.6196	13.1808
Complete secondary	3.7735	18.5430
Fundamental secondary	4.2050	8.4905
Primary	9.3670	19.0593
Experience	1.5397	6.8222
Experience^2	-0.0438	-0.0780

Source: own calculations, based on data from Derzhkomstat and RLMS.

**Note:* The rate of return to primary education in Russia is significantly different from zero only at 7 percent level.

In the light of other studies, obtained returns to education for secondary and tertiary levels for Ukraine are well below the reported world averages of 13.5 and 10.7 percent respectively (Psacharopoulos, 1994)[28]. This may reveal the fact that labor market in Ukraine is functioning worse than world average, and skills are relatively under-compensated (even in comparison with Russia).

On the contrary, Russian rates of return are close to and above the world averages, being even similar to the pattern of OECD countries[29] (8-14 percent overall according to OECD Education policy analysis 1997 as cited in Klazar et al. (2001)).

In general, the obtained rates of return are more or less close to the range reported by Nesterova and Sabirianova (1999) (6-8 percent for Russia, years 1995-96), by Klazar et al. (2001) (7-8 percent for the Czech Republic, 1995), and by other researchers for transitional economies (as cited by Nesterova and Sabirianova (1999)). It is worth noting, however, that the results of the reported studies may not be straightforwardly compared to mine, since due to data limitations the researchers frequently calculate and report the rate of return to a generic year of schooling without distinguishing between the educational levels.

To visualize the experience-earnings profile, Figure 2 plots the marginal effect of a year of experience on wages.

As evident, Russian labor market "favors" experienced workers to a greater extent than its Ukrainian counterpart, as the overall and marginal reward to experience in the relevant range is significantly higher in Russia. Maximal return to experience in Ukraine is obtained by individuals with approximately 18 years of working record (when marginal effect is zero), whereas in Russia this is true for a 44-year record. It is also easy to calculate that in Ukraine the overall (not marginal) impact of experience beyond 35 years on earnings is negative. For Russia the respective figure is 87 years, which appears to be irrelevant as there are no people in the sample with more than 60-years experience record.

---

[28] Comparison of the rate for primary education makes little sense, as there are few people in the sample holding this level.

[29] This similarity in relative terms should not be mistaken for similarity in absolute dimension, for wages in Russia, on average, are far below the ones in advanced economies.

Figure 2. Marginal Effect of Experience on Earnings, Conditional Return, Ukraine and Russia, 2000

Unconditional returns. As predicted, the unconditional rates of return for most educational levels (reported in Table 2) are higher than the conditional rates. Similar to the previous case, Russia unambiguously demonstrates higher returns at every level, the upper bound of the relative difference, nevertheless, is now much lower: 2.3 times (previously, 5 times for fundamental secondary education).

In Ukraine, returns fall into the range from 3.8 percent (fundamental secondary) to 9.2 percent (primary) and the rate of return to a year of compete higher schooling is 8.4 percent (more than a 2.1 percentage point increase from the conditional rate).

In Russia, the range is from almost 7 percent (incomplete higher) to 20.1 percent (primary) and the return to a year of schooling at the highest level is 10.3 percent.

Cumulative returns (estimated coefficients in Tables 17 and 18, Appendix D) are invariably higher in Russia. It is easy to calculate that the earned income of a holder of complete higher education is expected to be 5.33 times higher than the income of the uneducated and 1.78 times higher than the income of a holder of complete secondary education controlling for other wage determinants. Respective ratios for Ukraine are 2.40 and 1.58.

As in the case of conditional returns, the data does not seem to support the hypothesis of decreasing returns to schooling in the case of Ukraine. On the contrary, the diminishing returns pattern is evident for Russian sample.

To visualize the marginal effect of experience on earnings, refer to Figure 3 below. To a large extent it reflects the same pattern as Figure 2 above with one important qualification: significantly higher unconditional returns to experience vs. conditional ones in Ukraine.

Table 2. Unconditional Rates of Return to Education and Experience, Ukraine and Russia, percent

Description	Ukraine	Russia
Complete higher and above	8.3722	10.3239
Incomplete higher	5.7546	6.9541
Fundamental higher	8.9992	12.0413
Vocational training	7.2144	13.0332
Specialized secondary	8.0107	13.2815
Complete secondary	5.4266	18.4987
Fundamental secondary	3.8108	8.9173
Primary	9.2383	20.1236
Experience	3.5785	6.9551
Experience^2	-0.0644	-0.0788

Source: own calculations, based on data from Derzhkomstat and RLMS.

Figure 3. Marginal Effect of Experience on Earnings, Unconditional Return, Ukraine and Russia, 2000

The highest impact of experience on earnings in Ukraine is observed for approximately 28-years working record (44 for Russia). Also, the overall effect of experience on earnings in Ukraine is positive for a longer range: up to 56 years vs. 35 years in the case of conditional returns.

General Comparison of Indices

Indices of the human capital stock per capita calculated according to (28) are presented in Appendix E. Comparing Table 21 and Table 22, Table 23 and 24 I find that indices for Russian sample are higher. This is due to a stronger effect of both education and experience on earnings in Russia. Thus, it may be argued that labor market is functioning better in that country and human capital is valued relatively more there. Possibly, this is a consequence of the fact that Russia is 2 years ahead of Ukraine in reforming the economy (as sometimes popularly argued). Also, to a certain extent, it may be reflecting the positive impact on Russian economy of favorable conditions on the international market for oil, Russia's primary export and source of revenue to federal budget. However, the latter argument is weaker, insofar as Ukrainian economy has also been growing throughout the year 2000.

Sectoral Allocation of Human Capital

For the purpose of tractability and comparability, human capital indices should be normalized. Normalized indices for each of 16 sectors are presented in Table 3 and Table 4 below, where *NT* refers to normalization done with respect to the human capital stock per capita in the country (i.e., via dividing each respective index by the index of the human capital stock per capita for the whole sample), *NM* refers to normalization with respect to the human capital of men, and *NW* – of women.

As shown in Table 3, in Ukraine, the two sectors with highest stock of human capital per capita are science (1.23) and education (1.15). This is expected and intuitively plausible, since high educational attainment is a prerequisite for most positions in these sectors. The "worst-performing" branches, trade (.95), agriculture (.98), and construction (1.00) have stocks below the human capital stock per capita of the whole labor force.

Sectoral human capital allocation conditional on gender is similar to the one for the whole labor force. The most skilled men and women are concentrated in science (1.24 for men and 1.22 for women) and education (1.24 and 1.11). Individuals with relatively low stock are "attracted" to trade (.94 and .96) and agriculture (.95 and .99). The stock of human capital per capita of men in construction (.99) and business services (.99) as well as of women in defense (1.00) sectors is also below the respective gender stocks.

Table 3. Normalized Human Capital Indices, by Sector, Ukraine, 2000

Code	Sector	total	men		women	
		NT	NT	NM	NT	NW
1	Industry (including manufacturing)	1.0243	1.0195	1.0296	1.0314	1.0218
2	Agriculture, hunting, forestry and fishing	0.9819	0.9380	0.9472	0.9966	0.9872
3	Construction	0.9984	0.9771	0.9867	1.0719	1.0618
4	Transport, storage and communication services	1.0031	0.9927	1.0024	1.0267	1.0171
5	Wholesale and retail trade, public catering and related services	0.9524	0.9332	0.9423	0.9641	0.9551
6	Business services (mediation)	1.0034	0.9824	0.9921	1.0338	1.0242
7	IT services (computer related)	1.0796	1.0412	1.0514	1.0792	1.0691
8	Housing, public utilities and related consumer services	1.0118	1.0099	1.0198	1.0138	1.0043
9	Health care, physical training, sports and social work	1.0432	1.1195	1.1304	1.0281	1.0185
10	Education	1.1456	1.2312	1.2432	1.1186	1.1081
11	Culture and arts	1.1096	1.1727	1.1842	1.0831	1.0729
12	Science	1.2313	1.2364	1.2485	1.2273	1.2158
13	Finance, credit, insurance, legal services	1.0961	1.1136	1.1246	1.0888	1.0786
14	Government	1.1141	1.1409	1.1521	1.1048	1.0944
15	Defense and maintenance of public order (army, police etc.)	1.0055	1.0020	1.0118	1.0062	0.9968
16	Other and Non-classified	1.1101	1.1416	1.1528	1.0559	1.0460

Source: own calculations, based on data from Derzhkomstat.

As evident from Table 4, Russia has more or less comparable allocation of human capital. The "best performing" sectors are education (1.40) and science (1.35). However, there are six (in Ukraine, three) sectors (ignoring the non-classified) where workers have lower human capital stock per capita than the one for the whole labor force: trade (.86), defense (.91), transport (.91), surprisingly, IT services (.93), agriculture (.99), and construction (1.00).

Table 4. Normalized Human Capital Indices, by Sector, Russia, 2000

Code	Sector	total	men		women	
		NT	NT	NM	NT	NW
1	Industry (including manufacturing)	1.0796	1.1193	1.1470	1.0161	0.9927
2	Agriculture, hunting, forestry and fishing	0.9949	1.0025	1.0272	0.9866	0.9639
3	Construction	0.9991	0.9937	1.0183	1.0353	1.0115
4	Transport, storage and communication services	0.9124	0.8846	0.9065	1.0505	1.0263
5	Wholesale and retail trade, public catering and related services	0.8563	0.7870	0.8064	0.8770	0.8568
6	Business services (mediation)	1.1352	0.9822	1.0065	1.1686	1.1417
7	IT services (computer related)	0.9293	0.9746	0.9987	0.8897	0.8693
8	Housing, public utilities and related consumer services	1.0912	1.0939	1.1209	1.0881	1.0631
9	Health care, physical training, sports and social work	1.1426	1.3698	1.4037	1.1072	1.0817
10	Education	1.4016	1.6963	1.7382	1.3591	1.3278
11	Culture and arts	1.1235	1.0816	1.1083	1.1454	1.1190
12	Science	1.3531	1.3257	1.3585	1.3673	1.3359
13	Finance, credit, insurance, legal services	1.2287	1.3040	1.3362	1.2209	1.1928
14	Government	1.2327	1.4156	1.4506	1.2147	1.1867
15	Defense and maintenance of public order (army, police etc.)	0.9071	0.8720	0.8935	1.2256	1.1974
16	Other and Non-classified	0.8550	0.5943	0.6090	0.9934	0.9706

Source: own calculations, based on data from RLMS.

Sectoral allocation of human capital conditional on gender appears to be more diversified in Russia than in Ukraine, as men and women have slightly dissimilar patterns in the former. The two sectors where "best" men are concentrated are education (1.74) and government (1.45); health care (1.40) and science (1.36) are top third and fourth sectors respectively. Russian women with highest human capital stock per capita work in science (1.36) and education (1.33). Government (1.19) and health care (1.08) are top fifth and only top eighth respectively. Least-skilled men are found in trade (.81), defense (.89), transport (.91), and IT services (1.00). Women with lowest stock of human capital per capita are concentrated in trade (.86), IT services (.87), agriculture (.96), and industry/manufacturing (.99).

To draw policy implications it may be insightful to refer to Tables 25 and 26 in Appendix E, which present the rankings of sectors with respect to human capital index and its educational component, average earned income, and average age of the employed.

As evident, the situation in Ukraine looks fairly discouraging, since average remuneration in the top two sectors, science and education, is rather low: rankings of science and education by earned income are only 10 and 13 (of 16 sectors, with 16 denoting the lowest average remuneration). It may be hypothesized that such relatively low wages would not and do not attract rational individuals beginning their working careers in Ukraine. This statement is clearly supported by the data, since the rankings of science and education by average age are 1 and 4 (16 denotes a sector with the "youngest" employees). Meanwhile, relatively young labor force members are concentrated in business services (age rank of 16 and income rank of 1).

The situation in Russia does not appear to be much better than in Ukraine, though, less pessimistic in some respects. The top two sectors by human capital per capita index, education and science, are ranked 15^{th} and 5^{th} by average earned income and 5^{th} and 8^{th} by average age of the employed (more favorable state of Russian science sector vis-a-vis the Ukrainian counterpart is noteworthy both in terms of the earned income and age of the employees). Relatively younger labor force members are concentrated in defense (age rank is 16 and income rank is 2) and IT services (15 and 4).

It is hardly questioned that education and science are the two primary sectors where human capital is accumulated ("produced") and new knowledge is discovered. Therefore, the performance of these two sectors (predominantly dependent upon the quality of the employees, i.e., their stock of human capital) has direct consequences for a country's future growth and development. Moreover, as is also widely agreed upon, the two sectors produce important public and merit goods and generate positive externalities for all other branches of the economy. Thus, education and science "deserve" and, in fact, *are* subject to extensive financial support by the state.

In my opinion, the above analysis for Ukraine unambiguously illustrates gloomy prospects for the country, if current budgetary policy remains unchanged. As discussed above, education and science are not attracting younger labor force members, most likely, because of relatively low remuneration. Hence, the two sectors are doomed to shrink and, probably, will lose their top positions by the stock of human capital per capita as their currently highly educated and experienced but, alas, rather old employees start retiring on a large scale in 8-13 years or sooner.

Evidence shows that already now many regions of Ukraine have trouble filling vacancies of teachers in secondary schools. The problem, however, is currently masked by the sharp drop in fertility rates 10 years ago (persisting through all years of Ukraine's independence) and the resulting reduction in the size of the cohort of schooling-age children. As economic situation improves, fertility rates may well rise for all relevant cohorts and Ukraine may face substantial deficit of educators. Hence, the country's stock of human capital may significantly decrease over time for objective reasons.

The situation needs to be radically changed now since, in the future, the necessary reallocation of workers with high stock of human capital to education and science from other sectors (to meet the increased demand for education and productive knowledge) will, most likely, be very costly, if feasible at all. These high costs of meeting the increased demand will stem from the inevitable loss of firm- and sector-specific human capital (a consequence of the reallocation) and limited possibilities of importing educational services from abroad (especially, for elementary and secondary levels).

One possible suggestion for the immediate corrective action by the Ukrainian government is the redirection of currently available, but wasted, resources from heavy subsidization of the inefficient and "dying out" coal-mining industry to financing projects in education and science sectors. Another policy recommendation is to change present wage-setting schemes so as to motivate the young to pursue careers in education and science. The question of whether these suggestions can be implemented requires additional research well beyond the scope of this thesis.

Allocation of Human Capital by Economic Status

As shown by Tables 5 and 6 the allocation of the human capital stock per capita by economic status is similar in both countries.

Employees possess relatively higher stock than employers/entrepreneurs/self-employed (1.04 vs. .93 in Ukraine and 1.06 vs. 1.00 in Russia). The latter group, in turn, is relatively more skilled than the unemployed (.84 in Ukraine and .77 in Russia).

It is rather surprising that, in the two countries, the group of employers, entrepreneurs and self-employed has lower human capital stock than the employees. Probably, this reveals the effect of a subgroup of relatively less skilled people who cannot find a regular job, but are forced to make one's living and be stuck in low-productivity and low paid types of self-employment, e.g., street vending. This resembles the situation in many developing countries. Going into details behind the single value of the index, I find (refer to Table 23 and Table 24 in Appendix E) that lower human capital stock in Ukraine is due to both lower educational component and lower experience component of the index for the employers/entrepreneurs/self-employed. On the contrary, in

Russia the self-employed are slightly more experienced than the employees, but are still relatively low educated.

Table 5. Normalized Human Capital Indices, by Economic Status, Ukraine, 2000

Category	total	men		women	
	NT	NT	NM	NT	NW
employees	1.0367	1.0274	1.0375	1.0454	1.0356
employers/entrepreneurs/self-employed	0.9317	0.9212	0.9302	0.9514	0.9425
unemployed	0.8382	0.8278	0.8359	0.8467	0.8388
whole sample	1.0000	0.9903	1.0000	1.0095	1.0000

Source: own calculations, based on data from Derzhkomstat.

Table 6. Normalized Human Capital Indices, by Economic Status, Russia, 2000

Category	total	men		women	
	NT	NT	NM	NT	NW
employees	1.0565	1.0289	1.0543	1.0819	1.0570
employers/entrepreneurs/self-employed	0.9983	0.9717	0.9957	1.0300	1.0063
unemployed	0.7675	0.7740	0.7931	0.7600	0.7425
whole sample	1.0000	0.9759	1.0000	1.0235	1.0000

Source: own calculations, based on data from RLMS.

As far as gender is concerned, it is worth noting that, in both countries, women possess higher stock of human capital per capita than men. This finding for Ukrainian whole sample holds true for each economic status (refer to *NT* indices). In Russia, the groups of women who are employees and employers/entrepreneurs/self-employed are more skilled than respective groups of men, but unemployed women possess lower human capital stock per capita than unemployed men.

Chapter V

CONCLUSIONS

This research reveals the following facts pertaining to recent state of labor markets and composition of the labor forces in Ukraine and Russia conditional on the sector of employment, economic status, and gender.

First, the pool of the employed is not a random draw from the labor force as individuals with higher educational attainment and longer experience record are less likely to be unemployed. This sample selection bias implies that returns to schooling and experience calculated on the basis of data for the employed are not applicable to the whole labor force.

Second, as predicted, unconditional rates of return for most educational levels are higher than conditional returns (significantly in the case of Ukraine and marginally in the case of Russia). The hypothesis of diminishing returns to education is not supported by Ukrainian data, but is evident in Russian sample. The rates in Ukraine are fairly moderate, smaller than the world average ones, but still falling into the range for transitional economies reported by other researchers. Russian rates of return are, on average, approximately 1.5 times higher and even close to the pattern of advanced economies.

Third, Russia unambiguously demonstrates higher cumulative rates of return to most educational levels (as shown by coefficients of the regressions). The experience-earnings profile is also more "favorable" there. Consequently, human capital indices are lower in Ukraine, revealing relatively smaller compensation for human capital. This may reflect the fact that Russia is ahead of Ukraine in conducting labor market reforms and indicate that the problem of the skill mismatch is more severe in Ukraine.

Fourth, in the two countries, the sectors with highest human capital stock per capita are science and education, as expected. The sector where relatively less skilled individuals are concentrated is wholesale and retail trade. Sectoral allocation of human capital by gender is more diversified in Russia, where the patterns for employed men and women are slightly dissimilar.

Fifth, inspection of income and age rankings of sectors reveals gloomy prospects of Ukraine's future. Education and science are not attracting young labor force members because of, most likely, very low remuneration. Possibly, the two sectors will shrink and lose their top positions during the next decade. This may have strong negative implications for the national economy's growth and development. Probably, immediate corrective actions by the government are needed in these spheres.

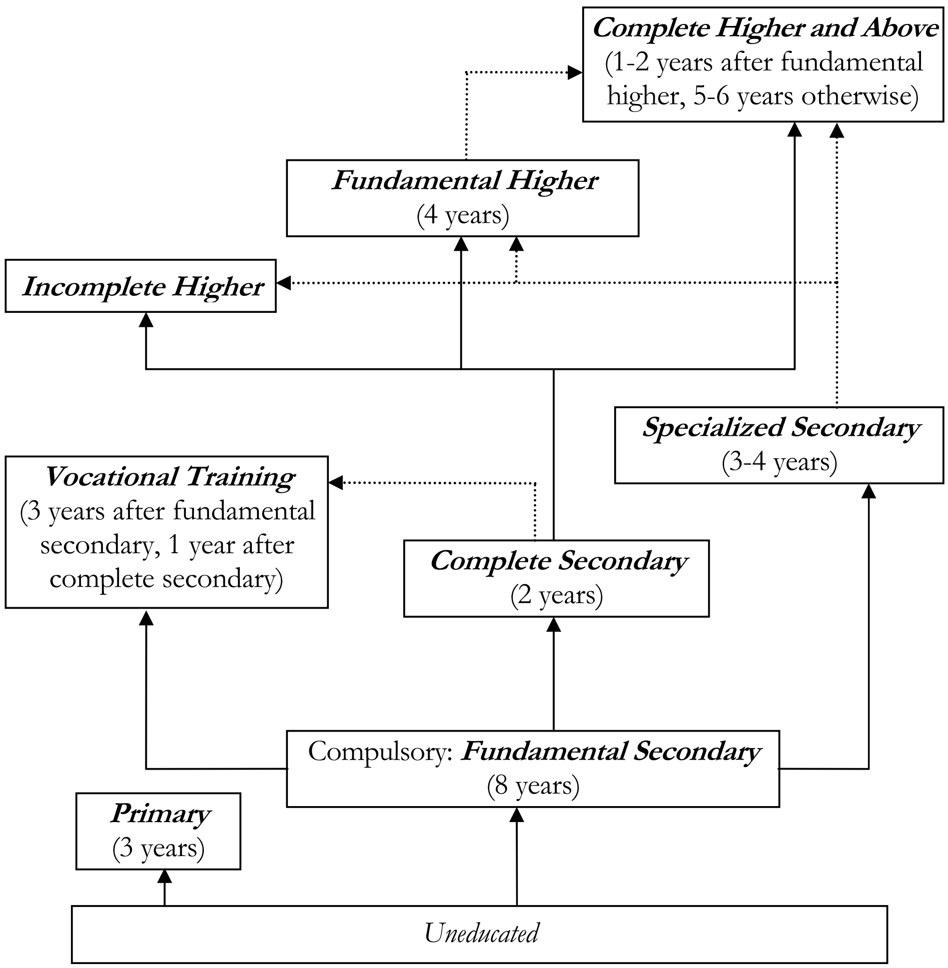
Sixth, of the three socio-economic groups, the unemployed are the least skilled. Unexpectedly, employers, entrepreneurs, and self-employed have lower stock of human capital per capita than employees. This bears some resemblance to the notorious pattern of developing countries where the less-skilled to make for their living have to undertake low-productivity jobs as better options are not available.

Seventh, women have somewhat higher stock of human capital per capita and this phenomenon is observed in all socio-economic groups with the exception of the unemployed women in Russia who are relatively less skilled than the unemployed men in the country.

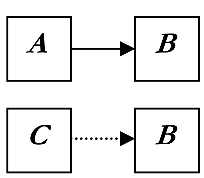
Appendix A

Educational Levels in Ukraine and Russia

Figure 4. Educational Levels Linkages, Ukraine and Russia



Legend:



stands for "A, as a rule, is a prior educational level to B"

stands for "C may be a prior educational level to B" (although, A as a prior level to B is more common)

Number of years to complete a given educational level, *after* a prior level has been completed, is provided in parentheses. Completion of the fundamental secondary education presupposes completion of the primary level.

Table 7. Educational Levels in Ukraine and Russia: A Summary

Level	Description	Total years to obtain, normally	Total years to obtain, Ukrainian dataset average	Total years to obtain, Russian dataset average	Prior Level
1	Complete higher and above	15-16	15.65	15.95	6 or 5 (after 3)*
2	Incomplete higher	>10	13.82	13.91	6 or 5
3	Fundamental higher	14	14.08	14.19	6 or 5
4	Vocational training	10-11	11.47	11.52	7 or 6
5	Specialized secondary	11-12	12.29	12.41	7
6	Complete secondary	10	10.20	10.29	7
7	Fundamental secondary	8	8.36	8.49	9**
8	Primary	3	3.88	3.41	9
9	Uneducated**	–	1.22	1.85	none

Source: own calculations, based on data from Derzhkomstat and RLMS.

Notes: *Refer to Figure 4 above and details in the text.
**Those who failed to complete the primary level.

Appendix B

Composition of Ukrainian and Russian Labor Forces

Table 8. Ukrainian Labor Force Composition by Educational Level, 2000

Level	Description	Share, percent								
		Total			Employed			Unemployed		
		Genders Together	Men	Women	Genders Together	Men	Women	Genders Together	Men	Women
1	Complete higher and above	19.81	18.99	20.60	22.23	21.27	23.20	7.14	5.82	8.21
2	Incomplete higher	1.78	2.14	1.44	1.86	2.22	1.50	1.34	1.58	1.15
3	Fundamental higher	3.54	2.81	4.24	3.74	3.13	4.35	2.48	0.95	3.72
4	Vocational training	18.36	22.16	14.70	18.16	21.93	14.40	19.45	23.5	16.15
5	Specialized secondary	25.75	21.68	29.68	26.45	22.42	30.47	22.06	17.35	25.90
6	Complete secondary	24.17	24.98	23.39	22.06	23.15	20.98	35.22	35.65	34.87
7	Fundamental secondary	5.75	6.30	5.22	4.80	5.21	4.38	10.75	12.62	9.23
8	Primary	0.74	0.85	0.62	0.59	0.59	0.59	1.49	2.37	0.77
9	Uneducated	0.10	0.09	0.11	0.11	0.08	0.13	0.07	0.16	0.00
Total		100.00	100.00	100.00	100.00	100.00	100.00	100.00	100	100.00

Source: own calculations, based on data from Derzhkomstat.

Table 9. Russian Labor Force Composition by Educational Level, 2000

Level	Description	Share, percent								
		Total			Employed			Unemployed		
		Genders Together	Men	Women	Genders Together	Men	Women	Genders Together	Men	Women
1	Complete higher and above	17.20	15.39	18.93	18.45	16.30	20.46	10.89	11.07	10.70
2	Incomplete higher	5.75	5.16	6.31	5.60	5.50	5.70	6.49	3.53	9.63
3	Fundamental higher	2.80	2.53	3.05	3.07	2.91	3.22	1.43	0.76	2.14
4	Vocational training	26.10	31.16	21.27	25.98	31.63	20.70	26.72	28.97	24.34
5	Specialized secondary	20.08	12.72	27.12	20.82	13.11	28.03	16.34	10.83	22.19
6	Complete secondary	17.62	20.28	15.09	16.42	18.61	14.36	23.74	28.21	18.98
7	Fundamental secondary	7.82	9.31	6.39	7.16	8.67	5.75	11.15	12.34	9.89
8	Primary	2.35	3.10	1.63	2.30	3.01	1.63	2.59	3.53	1.60
9	Uneducated	0.28	0.35	0.21	0.20	0.26	0.15	0.65	0.76	0.53
Total		100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Source: own calculations, based on data from RLMS.

Table 10. Employment Status by Educational Category, Ukraine, 2000

Level	Description	Share, percent								
		Genders Together			Men			Women		
		Employees	Employers	Unemployed	Employees	Employers	Unemployed	Employees	Employers	Unemployed
1	Complete higher and above	86.64	7.60	5.77	84.34	11.17	4.49	88.67	4.42	6.90
2	Incomplete higher	76.43	11.46	12.10	75.00	14.13	10.87	78.46	7.69	13.85
3	Fundamental higher	82.43	6.39	11.18	84.43	10.66	4.92	81.15	3.66	15.18
4	Vocational training	74.25	8.81	16.94	74.61	9.89	15.50	73.72	7.25	19.03
5	Specialized secondary	78.47	7.82	13.71	78.30	10.00	11.70	78.59	6.29	15.12
6	Complete secondary	67.42	9.27	23.31	66.02	13.11	20.87	68.85	5.32	25.83
7	Fundamental secondary	63.58	6.50	29.92	61.90	8.79	29.30	65.53	3.83	30.64
8	Primary	67.69	0.00	32.31	59.46	0.00	40.54	78.57	0.00	21.43
9	Uneducated	66.67	22.22	11.11	50.00	25.00	25.00	80.00	20.00	0.00

Source: own calculations, based on data from Derzhkomstat.

Table 11. Employment Status by Educational Category, Russia, 2000

Level	Description	Share, percent								
		Genders Together			Men			Women		
		Employees	Employers	Unemployed	Employees	Employers	Unemployed	Employees	Employers	Unemployed
1	Complete higher and above	80.62	8.94	10.43	75.00	12.50	12.50	84.99	6.18	8.83
2	Incomplete higher	68.40	13.01	18.59	72.03	16.10	11.86	65.56	10.60	23.84
3	Fundamental higher	87.02	4.58	8.40	86.21	8.62	5.17	87.67	1.37	10.96
4	Vocational training	74.71	8.43	16.86	73.91	9.96	16.13	75.83	6.29	17.88
5	Specialized secondary	78.09	8.51	13.40	76.63	8.59	14.78	78.74	8.47	12.79
6	Complete secondary	68.12	9.70	22.18	65.95	9.91	24.14	70.91	9.42	19.67
7	Fundamental secondary	63.11	13.39	23.50	62.91	14.08	23.00	63.40	12.42	24.18
8	Primary	60.91	20.91	18.18	69.01	11.27	19.72	46.15	38.46	15.38
9	Uneducated	46.15	15.38	38.46	62.50	0.00	37.50	20.00	40.00	40.00

Source: own calculations, based on data from RLMS.

Table 12. Employment in Ukraine by Sector, 2000

Code	Description	Share, percent		
		Genders Together	Men	Women
1	Industry (including manufacturing)	23.95	29.20	18.73
2	Agriculture, hunting, forestry and fishing	16.24	19.51	12.98
3	Construction	5.02	7.83	2.23
4	Transport, storage and communication services	8.31	11.72	4.92
5	Wholesale and retail trade, public catering and related services	9.94	7.67	12.20
6	Business services (mediation)	0.67	0.86	0.48
7	IT services (computer related)	0.46	0.41	0.51
8	Housing, public utilities and related consumer services	6.07	6.35	5.80
9	Health care, physical training, sports and social work	9.48	3.21	15.72
10	Education	10.02	4.35	15.66
11	Culture and arts	1.10	0.65	1.56
12	Science	0.65	0.51	0.78
13	Finance, credit, insurance, legal services	1.01	0.59	1.42
14	Government	3.41	1.76	5.05
15	Defense and maintenance of public order (army, police etc.)	2.90	4.43	1.37
16	Other and Non-classified	0.77	0.95	0.59
Total		100.00	100.00	100.00

Source: own calculations, based on data from Derzhkomstat.

Table 13. Employment in Russia by Sector, 2000

Code	Description	Share, percent		
		Genders Together	Men	Women
1	Industry (including manufacturing)	20.26	25.70	15.16
2	Agriculture, hunting, forestry and fishing	8.93	9.36	8.52
3	Construction	8.75	15.76	2.18
4	Transport, storage and communication services	12.38	21.05	4.26
5	Wholesale and retail trade, public catering and related services	11.43	5.55	16.94
6	Business services (mediation)	1.79	0.74	2.77
7	IT services (computer related)	0.66	0.63	0.69
8	Housing, public utilities and related consumer services	9.56	8.78	10.30
9	Health care, physical training, sports and social work	8.52	2.59	14.07
10	Education	6.37	1.75	10.70
11	Culture and arts	1.28	0.90	1.63
12	Science	1.76	1.32	2.18
13	Finance, credit, insurance, legal services	4.81	0.95	8.42
14	Government	0.87	0.16	1.54
15	Defense and maintenance of public order (army, police etc.)	2.55	4.71	0.54
16	Other and Non-classified	0.08	0.05	0.10
Total		100.00	100.00	100.00

Source: own calculations, based on data from RLMS.

Figure 5. Ukrainian Labor Force Composition
by Economic Status, 2000

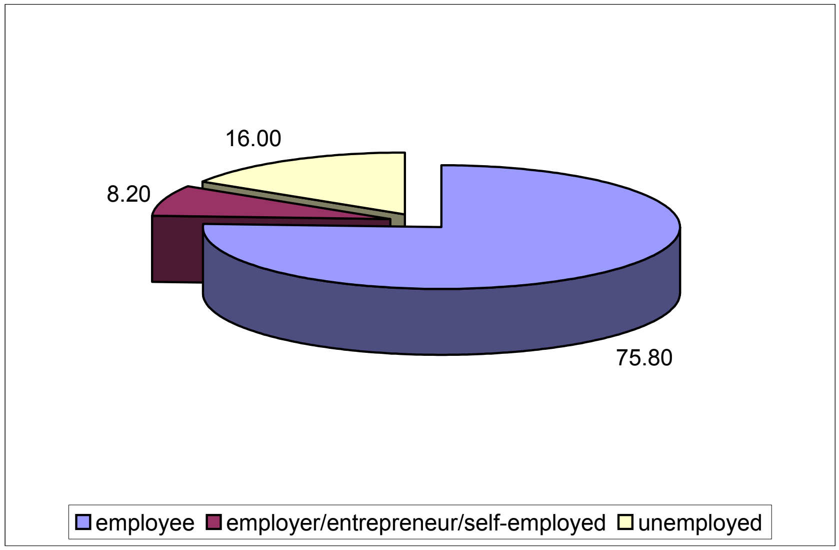


Figure 6. Ukrainian Labor Force Composition
by Economic Status, Men, 2000

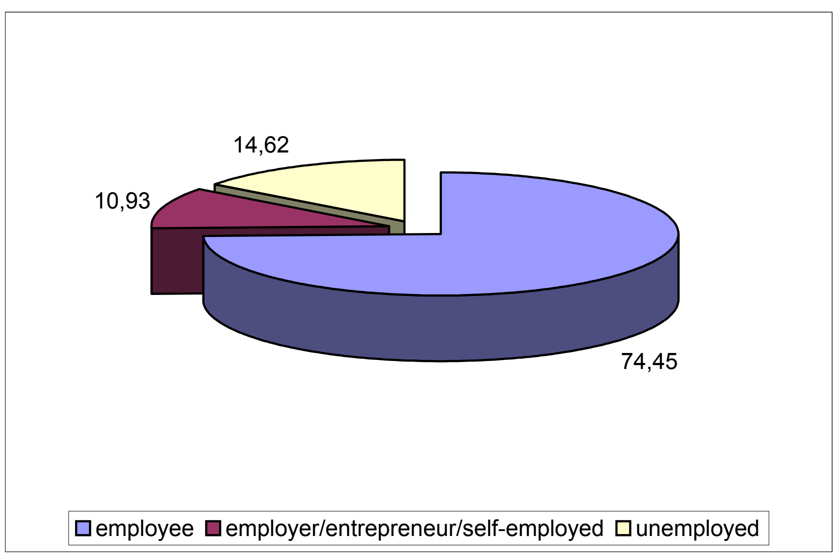


Figure 7. Ukrainian Labor Force Composition
by Economic Status, Women, 2000

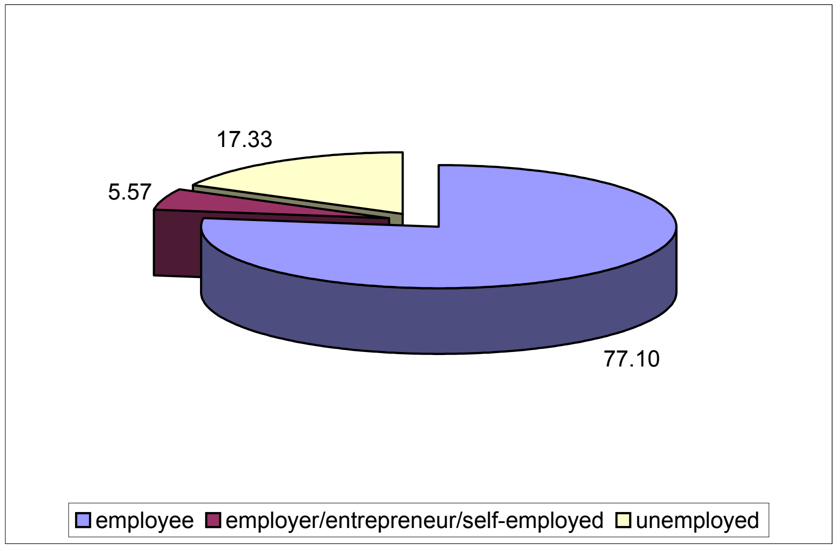


Figure 8. Russian Labor Force Composition
by Economic Status, 2000

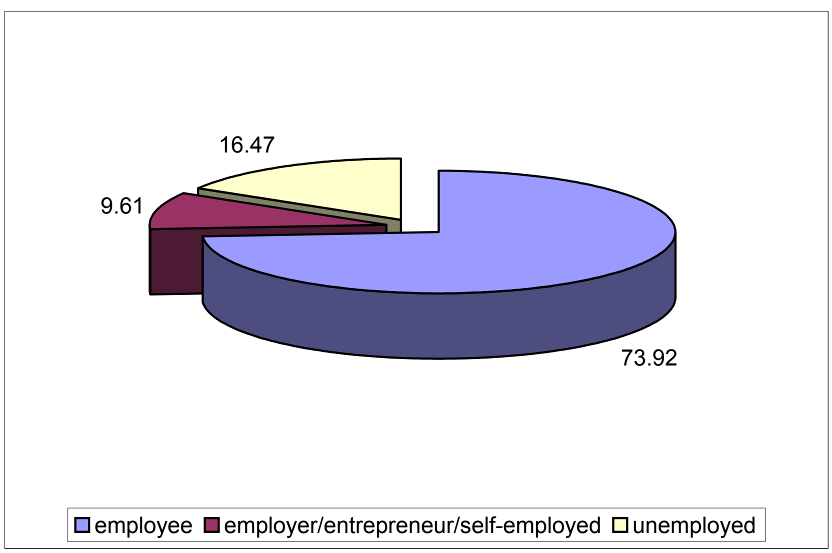


Figure 9. Russian Labor Force Composition

by Economic Status, Men, 2000

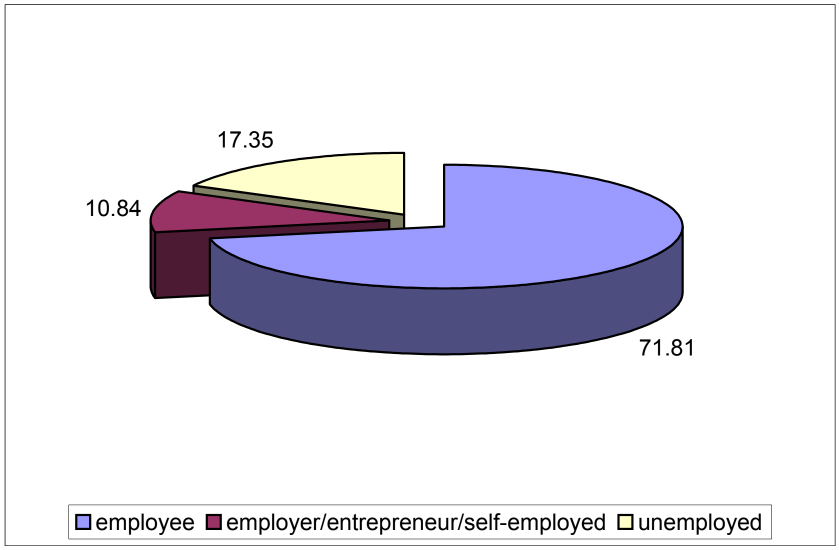
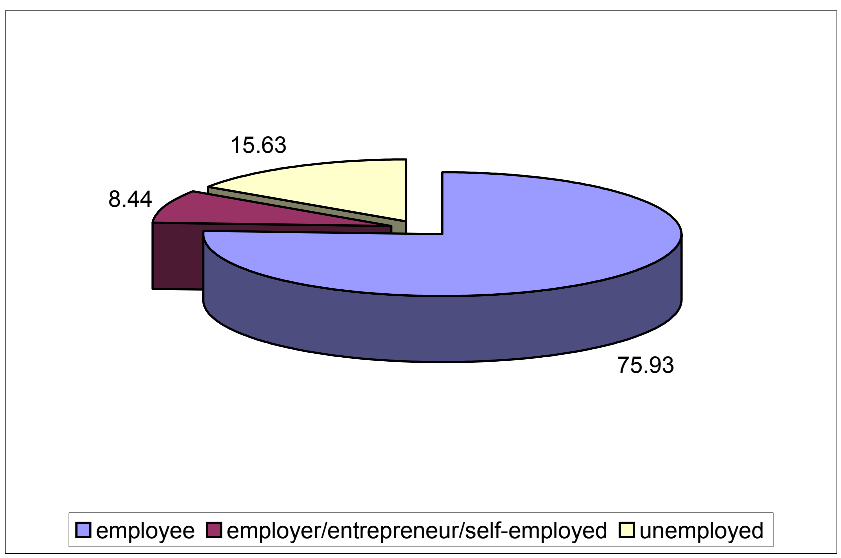


Figure 10. Russian Labor Force Composition

by Economic Status, Women, 2000



Appendix C

Sector Coding and ILO Job Coding

Table 14. Sector Coding and ILO Job Coding

Sector	Derzhkomstat Codes	Thesis Code	ISCO Codes (ILO Job Code in RLMS Database)
Industry (including manufacturing)	1	1	1222, 1312, 1210, 1232, 2143, 2145, 2146, 2147, 2412, 3113, 3115, 3116, 3116, 3118, 3123, 3139, 4132, 6142, 7000, 7111, 7112, 7213, 7211, 7214, 7215, 7216, 7221, 7222, 7223, 7224, 7231, 7232, 7233, 7242, 7243, 7311, 7321, 7322, 7323, 7331, 7324, 7332, 7346, 7416, 7421, 7422, 7423, 7424, 7431, 7432, 7433, 7434, 7435, 7436, 7441, 7442, 8000, 8111, 8112, 8113, 8121, 8122, 8123, 8124, 8131, 8139, 8141, 8142, 8143, 8151, 8152, 8153, 8154, 8155, 8159, 8161, 8162, 8163, 8171, 8172, 8211, 8212, 8223, 8224, 8229, 8231, 8232, 8240, 8251, 8252, 8253, 8261, 8262, 8263, 8264, 8265, 8266, 8269, 8276, 8279, 8281, 8282, 8284, 8285, 8286, 8290, 9311, 9321, 9322, 3117, 3129, 3117, 7439, 9329, 1239, 1319, 3431, 3434, 3439, 4111, 4115, 4190
Agriculture, hunting, forestry and fishing	2, 3	2	1221, 1299, 1311, 1520, 1530, 2213, 2223, 3212, 3213, 3227, 6111, 6112, 6113, 6114, 6121, 6122, 6123, 6124, 6129, 6130, 6141, 6151, 6152, 6153, 6154, 6210, 8271, 8272, 8273, 8275, 8277, 8278, 9211, 9212, 9213, 9332, 2219, 3219
Construction	4	3	1223, 1313, 7100, 7113, 7121, 7122, 7123, 7124, 7129, 7131, 7132, 7133, 7134, 7135, 7136, 7137, 7139, 7141, 7142, 7143, 7212, 8333, 9312, 9313, 7119
Transport, storage and communication services	5, 6	4	1226, 1316, 2144, 3114, 3132, 3141, 3142, 3143, 3144, 3145, 4133, 4142, 4221, 4222, 4223, 5111, 5112, 5113, 7244, 7245, 8311, 8312, 8320, 8321, 8322, 8323, 8324, 8331, 8332, 8334, 8340, 8340, 9151, 9331, 9333, 8329, 8339
Wholesale and retail trade, public catering and related services	7, 8, 10	5	1224, 1225, 1233, 1235, 1314, 3415, 3416, 4211, 4212, 5122, 5123, 5220, 5230, 7411, 7412, 7413, 7414, 7415, 8274, 9111, 9112, 9113, 9153
Business services (mediation)	9	6	1227, 1234, 1317, 1550, 2419, 3413, 3414, 3417, 3421, 3422, 3423, 3429, 4131, 4213
IT services (computer related)	11	7	1236, 2131, 2132, 2139, 3121, 3122, 4113, 4114, 8283
Housing, public utilities and related consumer services	12, 13, 14	8	1228, 1315, 1318, 1510, 2141, 2142, 2149, 3112, 3151, 3471, 5100, 5121, 5141, 5142, 5143, 5149, 7241, 7344, 7437, 9120, 9131, 9132, 9133, 9141, 9142, 9152, 9161, 9162, 9139
Health care, physical training, sports and social work	16, 17, 18	9	2212, 2221, 2222, 2224, 2229, 2230, 2445, 2446, 3133, 3152, 3221, 3222, 3223, 3224, 3225, 3226, 3228, 3229, 3231, 3232, 3241, 3242, 3460, 3475, 5131, 5132, 5133, 5139, 8221
Education	19	10	2300, 2310, 2320, 2331, 2332, 2340, 2351, 2352, 2359, 3310, 3320, 3330, 3340
Culture and arts	20, 21	11	2431, 2432, 2400, 2444, 2451, 2452, 2453, 2454, 2455, 3131, 3472, 3473, 3474, 3479, 4141, 5151, 5152, 7312, 7341, 7342, 7343
Science	22	12	1237, 2000, 2111, 2112, 2113, 2114, 2121, 2122, 2148, 2211, 2441, 2442, 2443, 3111, 3119, 3211
Finance, credit, insurance, legal services	23	13	1231, 2411, 2421, 2429, 3411, 3412, 3419, 3432, 3433, 4121, 4122, 4214, 4215
Government	24	14	1110, 1120, 1130, 2422, 3441, 3442, 3443, 3444, 3449
Defense and maintenance of public order	25	15	3450, 5161, 5162, 5163, 5169, 8222, 0110
Other and Non-classified	15, 26, 27	16	1229, 7313, 7345, 1239, 1319, 1540, 1590, 3431, 3434, 3439, 4000, 4111, 4112, 4115, 4143, 4144, 4190, 5210, 1141, 1142, 1143, 2460, 3480

Source: RLMS Administrative Office.

Appendix D

Regression Results

Table 15. Logistic Model: Effects of Schooling and Experience on the Probability of Being Employed, Ukraine, 2000

```
Logit estimates                                   Number of obs   =       8838
                                                  Wald chi2(8)    =     809.92
                                                  Prob > chi2     =     0.0000
Log likelihood = -3294.2243                       Pseudo R2       =     0.1522

------------------------------------------------------------------------------
             |               Robust
    employed |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   sch_years |   .2053386   .0154912    13.26   0.000     .1749764    .2357009
         exp |   .1186227   .0082336    14.41   0.000     .1024852    .1347602
         age |  -.0663196   .0077156    -8.60   0.000    -.0814419   -.0511973
         sex |  -.3655006   .0923125    -3.96   0.000    -.5464299   -.1845714
        sxmr |   1.057131   .1017272    10.39   0.000     .8577489    1.256512
        city |   .8196145    .080831    10.14   0.000     .6611887    .9780404
        town |   .2466469   .0743307     3.32   0.001     .1009615    .3923324
        Kyiv |   .4434132   .2146725     2.07   0.039     .0226629    .8641636
       _cons |  -.6667167   .2570144    -2.59   0.009    -1.170456   -.1629779
------------------------------------------------------------------------------


Marginal effects after logit
      y  = Pr(employed) (predict)
         =  .88406914
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
sch_years|   .0210453      .00153   13.77   0.000   .018049  .024042   12.0932
    exp  |   .0121577      .00082   14.74   0.000   .010541  .013775   18.4767
    age  |  -.0067972      .00080   -8.52   0.000  -.008361 -.005234   39.8387
    sex*|  -.0376392      .00972   -3.87   0.000  -.056698 -.018581   .490609
   sxmr*|   .1007418      .00934   10.79   0.000   .082435  .119049   .387757
   city*|   .0801822      .00758   10.58   0.000   .065329  .095035   .408690
   town*|   .0244093      .00714    3.42   0.001   .010422  .038397   .310930
   Kyiv*|   .0390174      .01594    2.45   0.014   .007781  .070254   .052161
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Table 16. Logistic Model: Effects of Schooling and Experience on the Probability of Being Employed, Russia, 2000

Logit estimates

Number of obs	= 4681
Wald chi2(8)	= 299.87
Prob > chi2	= 0.0000
Pseudo R2	= 0.0868

Log likelihood = -1912.4927

employed	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
sch_years	**.0791754**	.0165053	4.80	**0.000**	.0468257	.1115252
exp	**.1102207**	.0115278	9.56	**0.000**	.0876266	.1328148
age	-.0802607	.0110813	-7.24	0.000	-.1019795	-.0585418
sex	-.6030933	.1233937	-4.89	0.000	-.8449405	-.3612462
sxmr	.6098441	.1269788	4.80	0.000	.3609702	.858718
city	.863766	.1120526	7.71	0.000	.6441469	1.083385
town	.7073293	.1005263	7.04	0.000	.5103014	.9043572
Moscow	-.56124	.2544928	-2.21	0.027	-1.060037	-.0624433
_cons	1.543919	.3143428	4.91	0.000	.9278185	2.160019

Marginal effects after logit
y = Pr(employed) (predict)
= .85924594

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.	]	X
sch_years	**.0095757**	.00199	4.81	**0.000**	.005676	.013476	11.9990
exp	**.0133303**	.00137	9.72	**0.000**	.010642	.016018	18.1055
age	-.0097069	.00134	-7.26	0.000	-.012329	-.007085	38.2826
sex*	-.0735909	.01529	-4.81	0.000	-.10356	-.043621	.488784
sxmr*	.0703554	.01401	5.02	0.000	.042895	.097815	.382824
city*	.0922604	.01056	8.74	0.000	.07156	.112961	.288186
town*	.0782787	.01014	7.72	0.000	.058413	.098145	.315531
Moscow*	-.081349	.04319	-1.88	0.060	-.166005	.003307	.033326

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 17. Mincer-type Regression, OLS with Robust Errors, Ukraine, 2000

Cook-Weisberg test for heteroskedasticity, Ho: Constant variance
chi2(1) = 91.80, Prob > chi2 = 0.0000

Regression with robust standard errors

Number of obs	=	7424
F(19, 7404)	=	123.54
Prob > F	**=**	**0.0000**
R-squared	**=**	**0.2283**
A. R-squared	**=**	**0.2263**
Root MSE	=	.53734

lnW	Coef.	Rob. Std. Err.	t	P>|t|	[95% Conf.	Interval]
chigh	**.7613906**	.0606272	12.56	**0.000**	.642544	.8802373
ihigh	**.5029823**	.0722431	6.96	**0.000**	.3613653	.6445994
fhigh	**.6779202**	.0668948	10.13	**0.000**	.5467873	.8090531
ptu	**.4998939**	.060797	8.22	**0.000**	.3807145	.6190732
ssec	**.5723863**	.0602816	9.50	**0.000**	.4542172	.6905555
csec	**.4209677**	.0604048	6.97	**0.000**	.3025572	.5393783
fsec	**.3515361**	.0642863	5.47	**0.000**	.2255166	.4775556
prim	**.3634408**	.0916303	3.97	**0.000**	.1838194	.5430623
exp	**.0153969**	.0024582	6.26	**0.000**	.0105781	.0202158
exp2	**-.0004384**	.000046	-9.53	**0.000**	-.0005286	-.0003482
age	-.0005093	.0016346	-0.31	0.755	-.0037136	.0026949
sex	.1331590	.0249953	5.33	0.000	.0841611	.1821569
sxmr	.1936638	.0259579	7.46	0.000	.142779	.2445486
foreign	.4131324	.0750475	5.50	0.000	.2660181	.5602468
nongov	.0490968	.0131493	3.73	0.000	.0233205	.0748731
city	.4050379	.016003	25.31	0.000	.3736676	.4364082
town	.2794040	.0170118	16.42	0.000	.246056	.312752
Kyiv	.2784637	.0294517	9.45	0.000	.22073	.3361975
KyivOblast	.2129355	.0407804	5.22	0.000	.1329943	.2928768
_cons	6.335410	.0742978	85.27	0.000	6.189765	6.481054

Table 18. Mincer-type Regression, OLS with Robust Errors, Russia, 2000

Cook-Weisberg test for heteroskedasticity, Ho: Constant variance
chi2(1) = 64.51, Prob > chi2 = 0.0000

Regression with robust standard errors

Number of obs	=	3910
F(18, 3891)	=	66.31
Prob > F	**=**	**0.0000**
R-squared	**=**	**0.2280**
A. R-squared	**=**	**0.2245**
Root MSE	=	.95102

lnW	Coef.	Rob. Std. Err.	t	P>|t|	[95% Conf.	Interval]
chigh	**1.632361**	.3304380	4.94	**0.000**	.9845125	2.280209
ihigh	**1.303740**	.3372251	3.87	**0.000**	.6425854	1.964895
fhigh	**1.516700**	.3398672	4.46	**0.000**	.8503649	2.183034
ptu	**1.113913**	.3320670	3.35	**0.001**	.4628709	1.764954
ssec	**1.237784**	.3313934	3.74	**0.000**	.5880629	1.887505
csec	**1.054140**	.3325248	3.17	**0.002**	.4022005	1.70608
fsec	**.7205302**	.3373176	2.14	**0.033**	.0591941	1.381866
prim	.6497490	.3488657	1.86	**0.063**	-.0342279	1.333726
exp	.0682215	.006774	10.07	**0.000**	.0549407	.0815024
exp2	-.0007795	.000091	-8.56	**0.000**	-.000958	-.000601
age	-.0360805	.0055124	-6.55	0.000	-.046888	-.0252729
sxmr	.5141640	.0313625	16.39	0.000	.4526754	.5756525
foreign	.3158308	.1264703	2.50	0.013	.0678764	.5637852
nongov	.2845894	.0317188	8.97	0.000	.2224023	.3467765
city	.4901747	.0390363	12.56	0.000	.4136412	.5667083
town	.0933198	.0403463	2.31	0.021	.0142178	.1724218
Moscow	.2093768	.075047	2.79	0.005	.0622416	.3565121
MoscowOblast	.5638378	.0641932	8.78	0.000	.4379824	.6896933
_cons	5.777256	.3683168	15.69	0.000	5.055143	6.499368

Table 19. Heckman Selection Model, Ukraine, 2000

Heckman selection model	Number of obs	=	8838
(regression model with sample selection)	Censored obs	=	1414
	Uncensored obs	=	7424
	Wald chi2(19)	=	3402.00
Log likelihood = -9092.56	**Prob > chi2**	**=**	**0.0000**

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
lnW						
chigh	**.8747178**	.0435143	20.10	**0.000**	.7894313	.9600044
ihigh	**.6267515**	.0587615	10.67	**0.000**	.511581	.7419219
fhigh	**.767603**	.0518436	14.81	**0.000**	.6659914	.8692145
ptu	**.5429524**	.0441819	12.29	**0.000**	.4563574	.6295474
ssec	**.6334053**	.04336	14.61	**0.000**	.5484212	.7183894
csec	**.4184351**	.0441494	9.48	**0.000**	.3319038	.5049663
fsec	**.3185852**	.0473962	6.72	**0.000**	.2256905	.41148
prim	**.3584468**	.0829596	4.32	**0.000**	.195849	.5210446
exp	**.0357849**	.0027132	13.19	**0.000**	.0304671	.0411028
exp2	**-.000644**	.0000476	-13.53	**0.000**	-.0007373	-.0005507
age	-.0084007	.0017748	-4.73	0.000	-.0118793	-.0049221
sex	.0708123	.0262276	2.70	0.007	.0194071	.1222176
sxmr	.3065897	.027358	11.21	0.000	.252969	.3602104
foreign	.3536783	.0919774	3.85	0.000	.1734058	.5339507
nongov	.0350352	.0126038	2.78	0.005	.0103323	.0597381
city	.479563	.0161996	29.60	0.000	.4478124	.5113136
town	.3016342	.0170827	17.66	0.000	.2681528	.3351157
Kyiv	.297238	.0292026	10.18	0.000	.2400021	.354474
KyivOblast	.2148495	.0401605	5.35	0.000	.1361364	.2935627
_cons	6.125311	.0627072	97.68	0.000	6.002407	6.248215
select						
chigh	.8131058	.0440191	18.47	0.000	.72683	.8993817
ihigh	.6745312	.0752344	8.97	0.000	.5270745	.821988
fhigh	.536924	.057788	9.29	0.000	.4236616	.6501864
ptu	.3490201	.0431226	8.09	0.000	.2645013	.4335389
ssec	.4394966	.0410866	10.70	0.000	.3589684	.5200248
csec	.117754	.0407215	2.89	0.004	.0379413	.1975666
exp	.0954048	.003873	24.63	0.000	.0878138	.1029958
exp2	-.0009124	.0000665	-13.71	0.000	-.0010428	-.0007819
age	-.0400106	.002666	-15.01	0.000	-.0452358	-.0347853
sex	-.1133919	.0337544	-3.36	0.001	-.1795494	-.0472345
sxmr	.607034	.0352953	17.20	0.000	.5378565	.6762116
city	.4828195	.0236039	20.46	0.000	.4365566	.5290824
town	.2020403	.0235042	8.60	0.000	.1559728	.2481078
Kyiv	.3079812	.0512053	6.01	0.000	.2076207	.4083417
KyivOblast	.1621971	.0537033	3.02	0.003	.0569406	.2674536
_cons	.5640302	.0730922	7.72	0.000	.4207721	.7072884
/athrho	1.380958	.0678071	20.37	0.000	1.248059	1.513858
/lnsigma	-.4873973	.0120797	-40.35	0.000	-.5110731	-.4637215
rho	.8811656	.0151581			.8477384	.9076212
sigma	.614223	.0074196			.5998515	.6289387
lambda	.5412322	.0148781			.5120717	.5703926

Wald test of indep. eqns. (rho = 0): chi2(1) = 414.77 **Prob > chi2 = 0.0000**

Table 20. Heckman Selection Model, Russia, 2000

```
Heckman selection model                         Number of obs      =      4681
(regression model with sample selection)        Censored obs       =       771
                                                Uncensored obs     =      3910
                                                Wald chi2(18)      =   1237.00
Log likelihood = -7247.982                      Prob > chi2        =    0.0000
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
lnW						
chigh	**1.674142**	.3279494	5.10	**0.000**	1.031373	2.316911
ihigh	**1.341283**	.3349375	4.00	**0.000**	.6848177	1.997748
fhigh	**1.559366**	.3374418	4.62	**0.000**	.8979925	2.22074
ptu	**1.151764**	.3297232	3.49	**0.000**	.5055182	1.798009
ssec	**1.277962**	.3289394	3.89	**0.000**	.6332524	1.922671
csec	**1.089566**	.3300891	3.30	**0.001**	.4426036	1.736529
fsec	**.7567543**	.3349958	2.26	**0.024**	.1001746	1.413334
prim	**.686032**	.3462202	1.98	**0.048**	.0074529	1.364611
exp	**.0695513**	.0067185	10.35	**0.000**	.0563833	.0827192
exp2	**-.0007876**	.0000907	-8.68	**0.000**	-.0009654	-.0006098
age	-.0368486	.0054726	-6.73	0.000	-.0475746	-.0261225
sxmr	.5150029	.031279	16.46	0.000	.4536972	.5763086
foreign	.3152523	.1261677	2.50	0.012	.0679681	.5625366
nongov	.2841643	.0316474	8.98	0.000	.2221365	.346192
city	.4977496	.0392056	12.70	0.000	.4209081	.5745912
town	.0997563	.0404013	2.47	0.014	.0205712	.1789414
Moscow	.2050358	.0749934	2.73	0.006	.0580514	.3520201
MoscowOblast	.5638818	.0640422	8.80	0.000	.4383613	.6894022
_cons	5.731187	.3665503	15.64	0.000	5.012762	6.449612
select						
chigh	1.612603	.1658282	9.72	0.000	1.287586	1.93762
ihigh	1.318535	.1639384	8.04	0.000	.9972218	1.639849
fhigh	1.716222	.2250253	7.63	0.000	1.27518	2.157263
ptu	1.361021	.1481567	9.19	0.000	1.070639	1.651403
ssec	1.480995	.1531703	9.67	0.000	1.180787	1.781204
csec	1.219311	.1527905	7.98	0.000	.9198473	1.518775
fsec	1.293133	.1583008	8.17	0.000	.9828695	1.603397
prim	1.260697	.2262286	5.57	0.000	.8172967	1.704097
exp	.0719835	.0085909	8.38	0.000	.0551456	.0888213
exp2	-.0003183	.000143	-2.23	0.026	-.0005986	-.000038
age	-.0448469	.006125	-7.32	0.000	-.0568517	-.0328421
sex	-.3120276	.074617	-4.18	0.000	-.4582743	-.1657809
sxmr	.3130527	.0766559	4.08	0.000	.1628098	.4632955
city	.4973546	.0606459	8.20	0.000	.3784908	.6162184
town	.4038739	.0551522	7.32	0.000	.2957775	.5119702
Moscow	-.3067478	.1373419	-2.23	0.026	-.5759329	-.0375626
/athrho	.0459908	.0229327	2.01	0.045	.0010436	.090938
/lnsigma	-.0523013	.0146239	-3.58	0.000	-.0809636	-.023639
rho	.0459584	.0228842			.0010436	.0906882
sigma	.9490429	.0138787			.9222273	.9766382
lambda	.0436165	.0217538			.0009797	.0862532

Wald test of indep. eqns. (rho = 0): chi2(1) = 4.02 **Prob > chi2 = 0.0449**

Appendix E

Human Capital Stock Per Capita Indices

Table 21. Human Capital Indices, by Sector, Ukraine, 2000*

Code	Sector	two genders			men			women		
		EdC	ExC	Index	EdC	ExC	Index	EdC	ExC	Index
1	Industry (including manufacturing)	0.5942	0.4677	2.8918	0.5925	0.4647	2.8783	0.5967	0.4721	2.9119
2	Agriculture, hunting, forestry and fishing	0.5541	0.4655	2.7720	0.5123	0.4616	2.6481	0.5635	0.4709	2.8135
3	Construction	0.5898	0.4464	2.8185	0.5758	0.4390	2.7586	0.6389	0.4683	3.0261
4	Transport, storage and communication services	0.5857	0.4552	2.8320	0.5812	0.4493	2.8024	0.5964	0.4678	2.8986
5	Wholesale and retail trade, public catering and related services	0.6026	0.3865	2.6886	0.5965	0.3722	2.6345	0.6064	0.3949	2.7219
6	Business services (mediation**)	0.6740	0.3672	2.8327	0.6828	0.3373	2.7736	0.6585	0.4127	2.9187
7	IT services (computer related)	0.6926	0.4218	3.0478	0.7554	0.3228	2.9395	0.6430	0.4710	3.0467
8	Housing, public utilities and related consumer services	0.5824	0.4671	2.8564	0.5801	0.4676	2.8511	0.5850	0.4666	2.8622
9	Health care, physical training, sports and social work	0.6192	0.4610	2.9452	0.6786	0.4721	3.1604	0.6071	0.4585	2.9025
10	Education	0.7129	0.4609	3.2341	0.7559	0.4899	3.4757	0.7010	0.4490	3.1580
11	Culture and arts	0.6979	0.4439	3.1326	0.7334	0.4637	3.3106	0.6833	0.4344	3.0577
12	Science	0.7506	0.4953	3.4761	0.7533	0.4968	3.4904	0.7488	0.4938	3.4647
13	Finance, credit, insurance, legal services	0.6722	0.4574	3.0945	0.6926	0.4529	3.1440	0.6637	0.4592	3.0739
14	Government	0.6985	0.4475	3.1454	0.7323	0.4374	3.2210	0.6868	0.4507	3.1190
15	Defense and maintenance of public order (army, police etc.)	0.6421	0.4013	2.8388	0.6568	0.3831	2.8287	0.5949	0.4492	2.8407
16	Other and Non-classified	0.6823	0.4600	3.1341	0.6973	0.4729	3.2229	0.6585	0.4337	2.9808

Source: own calculations, based on data from Derzhkomstat.

Notes: *EdC stands for the educational component of the index ($\sum_{j=1}^{q} \beta_{1j}\eta_{ji}$), ExC denotes the experience component of the index ($\beta_2 \overline{Exp_i} + \beta_3 \overline{Exp_i}^2$).

**Brokerage etc.

Table 22. Human Capital Indices, by Sector, Russia, 2000*

Code	Sector	two genders			men			women		
		EdC	ExC	Index	EdC	ExC	Index	EdC	ExC	Index
1	Industry (including manufacturing)	1.2184	1.0926	10.0838	1.2176	1.1294	10.4546	1.2196	1.0307	9.4905
2	Agriculture, hunting, forestry and fishing	1.1275	1.1017	9.2924	1.1152	1.1217	9.3633	1.1403	1.0806	9.2153
3	Construction	1.1476	1.0859	9.3321	1.1488	1.0792	9.2815	1.1394	1.1297	9.6700
4	Transport, storage and communication services	1.1442	0.9985	8.5219	1.1262	0.9855	8.2627	1.2270	1.0565	9.8115
5	Wholesale and retail trade, public catering and related services	1.2215	0.8577	7.9978	1.2601	0.7347	7.3503	1.2097	0.8934	8.1909
6	Business services (mediation)	1.2427	1.1184	10.6032	1.3180	0.8985	9.1743	1.2239	1.1662	10.9145
7	IT services (computer related)	1.4080	0.7530	8.6798	1.5302	0.6784	9.1032	1.3032	0.8143	8.3101
8	Housing, public utilities and related consumer services	1.2113	1.1103	10.1917	1.2449	1.0792	10.2171	1.1844	1.1344	10.1634
9	Health care, physical training, sports and social work	1.3484	1.0192	10.6721	1.4855	1.0635	12.7946	1.3248	1.0114	10.3412
10	Education	1.5099	1.0621	13.0913	1.5762	1.1866	15.8437	1.4998	1.0414	12.6939
11	Culture and arts	1.3377	1.0131	10.4938	1.3306	0.9821	10.1020	1.3413	1.0287	10.6978
12	Science	1.4910	1.0457	12.6380	1.5212	0.9951	12.3827	1.4738	1.0734	12.7711
13	Finance, credit, insurance, legal services	1.3920	1.0483	11.4765	1.4962	1.0036	12.1796	1.3810	1.0529	11.4032
14	Government	1.4181	1.0254	11.5133	1.4100	1.1718	13.2218	1.4189	1.0099	11.3453
15	Defense and maintenance of public order (army, police etc.)	1.3608	0.7760	8.4721	1.3490	0.7483	8.1444	1.4561	0.9817	11.4469
16	Other and Non-classified	1.1518	0.9259	7.9857	1.1518	0.5622	5.5507	1.1518	1.0760	9.2790

Source: own calculations, based on data from RLMS.

**Note*: EdC stands for the educational component of the index ($\sum_{j=1}^{q} \beta_{1j}\eta_{ji}$), ExC denotes the experience component of the index ($\beta_2 \overline{Exp_i} + \beta_3 \overline{Exp_i}^2$).

Table 23. Human Capital Indices, by Economic Status, Ukraine, 2000*

Category	two genders			men			women		
	EdC	ExC	Index	EdC	ExC	Index	EdC	ExC	Index
employees	0.6131	0.4608	2.9268	0.6030	0.4619	2.9005	0.6224	0.4599	2.9514
employers/entrepreneurs/self-employed	0.5886	0.3785	2.6303	0.5839	0.3718	2.6006	0.5974	0.3907	2.6860
unemployed	0.5222	0.3392	2.3665	0.5035	0.3453	2.3370	0.5374	0.3341	2.3905
whole sample	0.5965	0.4413	2.8232	0.5864	0.4417	2.7957	0.6063	0.4410	2.8498

Source: own calculations, based on data from Derzhkomstat.

**Note:* EdC stands for the educational component of the index ($\sum_{j=1}^{q} \beta_{1j}\eta_{ji}$), ExC denotes the experience component of the index ($\beta_2 \overline{Exp_i} + \beta_3 \overline{Exp_i}^2$).

Table 24. Human Capital Indices, by Economic Status, Russia, 2000*

Category	two genders			men			women		
	EdC	ExC	Index	EdC	ExC	Index	EdC	ExC	Index
employees	1.2525	1.0367	9.8677	1.2138	1.0490	9.6103	1.2875	1.0255	10.1050
employers/entrepreneurs/self-employed	1.1950	1.0376	9.3241	1.2056	1.0001	9.0759	1.1820	1.0819	9.6201
unemployed	1.1690	0.8007	7.1688	1.1417	0.8365	7.2293	1.1981	0.7618	7.0982
whole sample	1.2332	1.0011	9.3402	1.2004	1.0095	9.1150	1.2646	0.9930	9.5601

Source: own calculations, based on data from RLMS.

**Note:* EdC stands for the educational component of the index ($\sum_{j=1}^{q} \beta_{1j}\eta_{ji}$), ExC denotes the experience component of the index ($\beta_2 \overline{Exp_i} + \beta_3 \overline{Exp_i}^2$).

Table 25. Human Capital, Income, and Age Rankings, Ukraine, 2000*

Code	Sector	Average Earned Income, UAH	Average Age	Rank of Index	Rank of Income	Rank of EdC	Rank of Age
1	Industry (including manufacturing)	2409.97	41.71	9	7	12	5
2	Agriculture, hunting, forestry and fishing	1473.84	42.12	15	14	16	3
3	Construction	2447.69	40.25	14	6	13	11
4	Transport, storage and communication services	2282.07	40.38	13	9	14	9
5	Wholesale and retail trade, public catering and related services	2209.01	36.36	16	11	11	15
6	Business services (mediation)	3544.46	35.20	12	1	7	16
7	IT services (computer related)	2758.05	38.85	7	4	5	13
8	Housing, public utilities and related consumer services	1928.29	42.36	10	12	15	2
9	Health care, physical training, sports and social work	1463.74	41.36	8	15	10	6
10	Education	1497.86	41.81	2	13	2	4
11	Culture and arts	1461.30	40.32	5	16	4	10
12	Science	2240.58	47.46	1	10	1	1
13	Finance, credit, insurance, legal services	2888.27	40.89	6	2	8	8
14	Government	2361.55	39.95	3	8	3	12
15	Defense and maintenance of public order (army, police etc.)	2749.49	36.40	11	5	9	14
16	Other and Non-classified	2808.18	41.32	4	3	6	7

Source: own calculations, based on data from Derzhkomstat.

**Note:* EdC stands for the educational component of the index ($\sum_{j=1}^{q} \beta_{1j}\eta_{ji}$). Rankings are in descending order.

Table 26. Human Capital, Income, and Age Rankings, Russia, 2000*

Code	Sector	Average Earned Income, Roubles	Average Age	Rank of Index	Rank of Income	Rank of EdC	Rank of Age
1	Industry (including manufacturing)	2111.74	40.22	9	9	11	4
2	Agriculture, hunting, forestry and fishing	1467.74	40.34	11	13	16	3
3	Construction	2595.54	39.47	10	3	14	6
4	Transport, storage and communication services	2199.31	37.23	13	8	15	12
5	Wholesale and retail trade, public catering and related services	1836.45	35.61	15	10	10	14
6	Business services (mediation)	1827.51	40.81	6	11	9	2
7	IT services (computer related)	2584.19	32.31	12	4	4	15
8	Housing, public utilities and related consumer services	1519.52	40.98	8	12	12	1
9	Health care, physical training, sports and social work	1229.21	39.16	5	16	7	10
10	Education	1379.38	40.20	1	15	1	5
11	Culture and arts	1431.12	37.90	7	14	8	11
12	Science	2427.75	39.28	2	5	2	8
13	Finance, credit, insurance, legal services	2285.98	39.16	4	7	5	9
14	Government	2371.12	39.47	3	6	3	7
15	Defense and maintenance of public order (army, police etc.)	2733.46	32.14	14	2	6	16
16	Other and Non-classified	3053.33	35.67	16	1	13	13

Source: own calculations, based on data from RLMS.

**Note:* EdC stands for the educational component of the index ($\sum_{j=1}^{q} \beta_{1j}\eta_{ji}$). Rankings are in descending order.

Appendix F

Testing Bivariate Normality Assumption in Tobit II

The upshot is that the issue remains unsettled.
William H. Greene. Econometric Analysis (4^{ed}, p. 934)

According to the theory, consistency of the maximum likelihood estimators directly depends on the validity of the "distributional" assumption, i.e., postulated probability density function of the error term.

Therefore, much attention in the literature is paid to testing the specification of such popular models as probit and logit, in particular, to verifying the validity of the chosen functional form of the distribution: standard normal or logistic, respectively. The derivation and calculation of the test statistics is relatively simple and is described by Verbeek (2000).

In contrast, the specification of Tobit II (Heckman selection model) is by and large not verified in empirical studies and the bivariate normality assumption continues to be taken for granted: "...the empirical literature on the subject continues to be dominated by Heckman's original model built around the joint normality distribution." (Greene 2000, 934). This may be explained by the fact that relevant tests have not yet been put into graduate textbooks and guides in advanced econometrics still being scattered among few highly technical articles.

Despite the fact that Greene (2000, 934) calls the issue of testing bivariate normality "unsettled" and recommends to use "some promising approaches based on robust and nonparametric estimators" for this purpose, it is still possible to conduct the relevant test staying within the "standard" testing methodology framework.

The two articles of interest are by Pagan and Vella (1989) and Lee (1984). Pagan and Vella's suggestion, which is borrowed from Gallant and Nychka

(1987) is to use a conditional moment test ("generalized" Lagrange multiplier test). Strictly speaking, however, the idea that is put forward by the researchers is applicable only when the estimates of the parameters of the selection model are obtained via Heckman's (1979) two-step procedure. The two-step methodology, in turn, is undesirable if the efficiency of the estimator is of a concern.

In contrast, Lee (1984) suggests that a classical LM test be used, in which the derivation of the test statistic is directly related to the loglikelihood function. As noted by Pagan and Vella (1989, S51) Lee's test "seems to have had little application". Partially, this is due to the complexity of the scores. More likely, however, this is because of the popularity of the two-step estimation procedure prior to the time when the algorithms of "direct" maximization of the non-trivial loglikelihood functions (such as in Tobit I and Tobit II) were implemented in popular econometrics software.

Since in my thesis I rely on ML estimation of the parameters of Tobit II model, which under the hypothesis of bivariate normality produces more efficient estimates, Lee's suggestion appears to me the only feasible way to conduct the specification test. In what follows I reproduce the major findings of Lee (1984), explicitly write out all necessary formulas, and provide some explanations.

Let the model be:

$$y_i^* = \mathbf{x}_i'\boldsymbol{\beta} + \sigma u_i, \qquad (35)$$

$$g_i^* = \mathbf{z}_i'\boldsymbol{\gamma} - \varepsilon_i, \qquad (36)^1$$

$$y_i = y_i^* \text{ if } g_i^* > 0 \Leftrightarrow I_i = 1, \qquad (37)$$

$$y_i \text{ is not observed if } g_i^* \le 0 \Leftrightarrow I_i = 0, \qquad (38)$$

where I_i indicates whether y_i is observed for observation *i*.

[1] Non-conventional "minus" sign is important, since it allows not to impose an (additional) assumption of the symmetric shape of the "true" bivariate and marginal probability density functions.

u_i and ε_i are assumed to be independent of both $\mathbf{x}_i$ and $\mathbf{z}_i$ and to be identically distributed according to some joint probability density function (pdf). In standard Tobit II (Heckman selection) model, this joint pdf is postulated to be bivariate normal, $b(u,\varepsilon)$, i.e., $(u,\varepsilon) \sim NIID(0,0,1,1,\rho)$: normally, independently, and identically distributed with zero means, unit variances, and correlation coefficient, ρ[2].

Suppose that the "distributional" assumption does not hold with respect to the functional form of the bivariate distribution. Then, the "true" bivariate pdf[3], $h(u,\varepsilon)$, may be approximated (via a type of Taylor series) as:

$$h(u,\varepsilon) = b(u,\varepsilon) + \sum_{r+s\geq 3} \kappa_{rs} \frac{1}{r!s!} H_{rs}(u,\varepsilon) b(u,\varepsilon), \qquad (39)$$

where H_{rs} are bivariate Hermite polynomials (which include the partial derivatives of the pdf), κ_{rs} are "coefficients" (in fact, some functions of the cumulants of u and ε). Intuitively, κ_{rs} represent (central) moments $E(u^j \varepsilon^k),\ j+k \geq 3$.[4]

Given (39), the marginal pdf of ε is:

$$f(\varepsilon) = \phi(\varepsilon) + \sum_{s=1}^{\infty} \kappa_{0s} \frac{1}{s!} H_s(\varepsilon)\phi(\varepsilon), \qquad (40)$$

---

[2] Evidently, ρ is the negative of the estimated correlation coefficient between u and $-\varepsilon$ (which is reported by Stata and some other popular econometrics software packages that employ the conventional form of (F2) with the "plus" sign).

[3] Hereinafter, the word combination "given $\mathbf{X}$ and $\mathbf{Z}$" will be omitted.

[4] Note that the parameters of the bivariate pdf are identified by 5 (central) moments: two means, two variances, and a correlation coefficient, i.e., $1 \leq j+k \leq 2$.

where H_s is a univariate Hermite polynomial of order s: $H_s = \left[(-1)^s \frac{d^s \phi(\varepsilon)}{d\varepsilon^2} \right] \Big/ \phi(\varepsilon), s = 0,1,2.... H_0 \equiv 1$[5] , $\phi(\cdot)$ is standard normal pdf.

Without loss of generality, for the purpose of the specification test, and upper bound of 4 may be *r*+*s*. Then, $h(u,\varepsilon)$ and its margin $f(\varepsilon)$ represent the so called Type AA surface, where κ_{rs} are:

$$\begin{aligned}
\kappa_{30} &= \mu_{30}, \\
\kappa_{03} &= \mu_{03}, \\
\kappa_{21} &= \mu_{21}, \\
\kappa_{12} &= \mu_{12}, \\
\kappa_{40} &= \mu_{40} - 3, \\
\kappa_{04} &= \mu_{04} - 3, \\
\kappa_{31} &= \mu_{31} - 3\rho, \\
\kappa_{13} &= \mu_{13} - 3\rho, \\
\kappa_{22} &= \mu_{22} - 2\rho^2 - 1, \\
\mu_{jk} &= E(u^j \varepsilon^k).
\end{aligned} \tag{41}$$

If the "true" pdf is bivariate normal, then, as it follows from the properties of the (central) moments of the bivariate pdf, all κ_{rs} in (41) are zero. Testing the hypothesis of zero κ_{rs} is, thus, equivalent to testing the "distributional" assumption and is the core of the specification test.

Let $\psi(y,1)$ be the joint density function of y^* and $g^* : g^* > 0$. It follows from (39) that (relevant subscripts are omitted):

[5] Greene (2000, 94). Formulas for bivariate Hermite polynomials are complex. For the purpose of the test, they will be after some simplification written out explicitly below.

$$\psi(y,1)=\frac{1}{\sigma}\int_{-\infty}^{\mathbf{z'\gamma}}h(\frac{y-\mathbf{x'\beta}}{\sigma},\varepsilon)d\varepsilon= \\ =s\int_{-\infty}^{\mathbf{z'\gamma}}b(u,\varepsilon)\left[1+\sum_{3\le r+s\le 4}\kappa_{rs}\frac{1}{r!s!}H_{rs}(u,\varepsilon)\right]d\varepsilon, \tag{42}$$

where $s=1/\sigma$ and $u=(y-\mathbf{x'\beta})/\sigma=s(y-\mathbf{x'\beta})$ (these (re)definitions will be used below).

Given (40), it is easy to show that the probability of observing I_i=1, i.e., probability of selection, is (omitting subscripts):

$$F(\mathbf{z'\gamma})=\Phi(\mathbf{z'\gamma})-\sum_{s=3}^{4}\kappa_{0s}\frac{1}{s!}H_{s-1}(\mathbf{z'\gamma})\phi(\mathbf{z'\gamma}), \tag{43}$$

where $\Phi(\cdot)$ is standard normal probability mass function (cumulative).

Denoting all κ_{rs} as $\boldsymbol{\kappa}$ vector the loglikelihood function is:

$$\ln L(\boldsymbol{\beta},s,\boldsymbol{\gamma},\rho,\boldsymbol{\kappa})=\sum_{i=1}^{N}\ln L_i=\sum_{i=1}^{N}\{I_i\ln\psi(y_i,1)+(1-I_i)\ln[1-F(\mathbf{z}_i'\boldsymbol{\gamma})]\}, \tag{44}$$

where N is the size of the whole sample, which includes observations with y_i observed (corresponding to the first term in the sum in (44)) and where it is not observed (second term). Note that maximization of (original) (44) with respect to $s=1/\sigma$ must produce same results as maximization of (redefined) (44) with respect to σ.

Given that the parameters of the model are estimated with the normality assumption imposed *a priori*, an obvious feasible candidate for the specification test is a Lagrange multiplier test (score test).

As follows from Verbeek (2000, 163-165), the test statistic is:

$$\xi_{LM}=\mathbf{i'S(S'S)}^{-1}\mathbf{S'i}, \tag{45}$$

where $\mathbf{i}$ is a (Nx1) vector of ones. **S** matrix is the matrix of scores of (44) evaluated at restricted parameter estimates (with bivariate normality assumption imposed):

$$S = \left. \begin{pmatrix} \frac{\partial \ln L_1}{\partial \boldsymbol{\beta}'} & \frac{\partial \ln L_1}{\partial s} & \frac{\partial \ln L_1}{\partial \rho} & \frac{\partial \ln L_1}{\partial \boldsymbol{\gamma}'} & \frac{\partial \ln L_1}{\partial \boldsymbol{\kappa}'} \\ \vdots & & \ddots & & \vdots \\ \frac{\partial \ln L_N}{\partial \boldsymbol{\beta}'} & \frac{\partial \ln L_N}{\partial s} & \frac{\partial \ln L_N}{\partial \rho} & \frac{\partial \ln L_N}{\partial \boldsymbol{\gamma}'} & \frac{\partial \ln L_N}{\partial \boldsymbol{\kappa}'} \end{pmatrix} \right|_{(\hat{\boldsymbol{\beta}}, \hat{s}, \hat{\rho}, \hat{\boldsymbol{\gamma}}, \boldsymbol{\kappa}=\mathbf{0})} , \quad (46)$$

where, as argued above, $\hat{\rho}$ is the negative of the correlation coefficient estimated by Stata and other popular econometrics software, $\hat{s}$ is the inverse of estimated "sigma", and all other estimates are equal to their estimated counterparts.

Under H_0: $\boldsymbol{\kappa} = \mathbf{0}$ (the "true" joint distribution is bivariate normal) ξ_{LM} should follow the χ^2 distribution with nine degrees of freedom (note that there are exactly nine κ_{rs}). Note also that under H_0 the marginal distribution of ε is automatically univariate normal by properties of the bivariate normal pdf.

The main issue is the derivation of the scores and estimation of the elements of the matrix **S**. As shown by Lee (1984, 851), the scores evaluated under H_0 are (suppressing hats for convenience):

$$\frac{\partial \ln L_i}{\partial \boldsymbol{\beta}'}\Big|_{H_0} = \frac{s\mathbf{x}_i' I_i}{(1-\rho^2)} \left[u_i - \rho \int_{-\infty}^{\mathbf{z}_i'\boldsymbol{\gamma}} \varepsilon b^*(\varepsilon \mid u_i) d\varepsilon \right] \quad (47)$$

$$\frac{\partial \ln L_i}{\partial s}\Big|_{H_0} = \frac{I_i}{s} \left[1 - \frac{u_i}{1-\rho^2} \left(u_i - \rho \int_{-\infty}^{\mathbf{z}_i'\boldsymbol{\gamma}} \varepsilon b^*(\varepsilon \mid u_i) d\varepsilon \right) \right] \quad (48)$$

$$\frac{\partial \ln L_i}{\partial \rho}\Big|_{H_0} = I_i \left[\begin{array}{l} -\frac{\rho}{(1-\rho^2)^2} u_i^2 + \frac{\rho}{1-\rho^2} + \frac{1+\rho^2}{(1-\rho^2)^2} u_i \int_{-\infty}^{\mathbf{z_i'\gamma}} \varepsilon b^*(\varepsilon | u_i) d\varepsilon - \\ -\frac{\rho}{(1-\rho^2)^2} \int_{-\infty}^{\mathbf{z_i'\gamma}} \varepsilon^2 b^*(\varepsilon | u_i) d\varepsilon \end{array} \right] \quad (49)$$

$$\frac{\partial \ln L_i}{\partial \boldsymbol{\gamma}'}\Big|_{H_0} = \mathbf{z}_i' \left[I_i b^*(\mathbf{z}_i'\boldsymbol{\gamma} | u_i) - (1-I_i)\phi_*(\mathbf{z}_i'\boldsymbol{\gamma}) \right] \quad (50)$$

$$\frac{\partial \ln L_i}{\partial \kappa_{jk}}\Big|_{H_0} = \frac{I_i}{j!k!} \int_{-\infty}^{\mathbf{z_i'\gamma}} H_{jk}(u_i, \varepsilon) b^*(\varepsilon | u_i) d\varepsilon, (j,k) \neq (0,3),(0,4) \quad (51)$$

$$\frac{\partial \ln L_i}{\partial \kappa_{03}}\Big|_{H_0} = \frac{1}{6}\left[I_i \int_{-\infty}^{\mathbf{z_i'\gamma}} H_{03}(u_i, \varepsilon) b^*(\varepsilon | u_i) d\varepsilon + (1-I_i) H_2(\mathbf{z}_i'\boldsymbol{\gamma})\phi_*(\mathbf{z}_i'\boldsymbol{\gamma}) \right] \quad (52)$$

$$\frac{\partial \ln L_i}{\partial \kappa_{04}}\Big|_{H_0} = \frac{1}{24}\left[I_i \int_{-\infty}^{\mathbf{z_i'\gamma}} H_{04}(u_i, \varepsilon) b^*(\varepsilon | u_i) d\varepsilon + (1-I_i) H_3(\mathbf{z}_i'\boldsymbol{\gamma})\phi_*(\mathbf{z}_i'\boldsymbol{\gamma}) \right], \quad (53)$$

where $b^*(\varepsilon | u_i) = \frac{1}{\sqrt{1-\rho^2}} \frac{\phi\left(\frac{\varepsilon - \rho u_i}{\sqrt{1-\rho^2}}\right)}{\Phi\left(\frac{\mathbf{z}_i'\boldsymbol{\gamma} - \rho u_i}{\sqrt{1-\rho^2}}\right)}$, $\phi_*(\mathbf{z}_i'\boldsymbol{\gamma}) = \frac{\phi(\mathbf{z}_i'\boldsymbol{\gamma})}{1-\Phi(\mathbf{z}_i'\boldsymbol{\gamma})}$.

Let the typical row of the matrix **S** be (note that the first and the fourth elements are rows, all other elements are scalars):

$$\left(S1\mathbf{x} \quad S2 \quad S3 \quad S4\mathbf{z} \quad S5 \quad S6 \quad S7 \quad S8 \quad S9 \quad S10 \quad S11 \quad S12 \quad S13 \right). \quad (54)$$

Expanding the bivariate Hermite polynomials and employing the formulas provided by Lee (1984), it may be shown that:

$$S1\mathbf{x} = \mathbf{x}_i' \frac{sI_i}{(1-\rho^2)} \left[u_i - \rho AA \right] \quad (55)$$

$$S2 = \frac{I_i}{s}\left[1 - \frac{u_i}{1-\rho^2}\left(u_i - \rho AA\right)\right] \tag{56}$$

$$S3 = I_i\left[-\frac{\rho}{(1-\rho^2)^2}u_i^2 + \frac{\rho}{1-\rho^2} + \frac{1+\rho^2}{(1-\rho^2)^2}u_i AA - \frac{\rho}{(1-\rho^2)^2}BB\right] \tag{57}$$

$$S4\mathbf{z} = \mathbf{z}_i'\left[I_i \frac{1}{\sqrt{1-\rho^2}} \frac{\phi\left(\frac{\mathbf{z}_i'\boldsymbol{\gamma} - \rho u_i}{\sqrt{1-\rho^2}}\right)}{\Phi\left(\frac{\mathbf{z}_i'\boldsymbol{\gamma} - \rho u_i}{\sqrt{1-\rho^2}}\right)} - (1-I_i)\frac{\phi(\mathbf{z}_i'\boldsymbol{\gamma})}{1-\Phi(\mathbf{z}_i'\boldsymbol{\gamma})}\right] \tag{58}$$

$$\begin{aligned} S5 = {} & \frac{1}{6(1-\rho^2)^3}\left[I_i u_i^3 + (1-I_i)EE\right] - \frac{\rho}{2(1-\rho^2)^3}\left[I_i u_i^2 AA + (1-I_i)FF\right] + \\ & + \frac{\rho^2}{2(1-\rho^2)^3}\left[I_i u_i BB + (1-I_i)\rho GG\right] - \frac{\rho^3}{6(1-\rho^2)^3}\left[I_i CC + (1-I_i)GG\right] \end{aligned} \tag{59}$$

$$\begin{aligned} S6 = {} & -\frac{\rho}{2(1-\rho^2)^3}\left[I_i u_i^3 + (1-I_i)EE\right] + \frac{1+2\rho^2}{2(1-\rho^2)^3}\left[I_i u_i^2 AA + (1-I_i)FF\right] - \\ & - \frac{\rho(2+\rho^2)}{2(1-\rho^2)^3}\left[I_i u_i BB + (1-I_i)\rho GG\right] + \frac{\rho^2}{2(1-\rho^2)^3}\left[I_i CC + (1-I_i)GG\right] \end{aligned} \tag{60}$$

$$\begin{aligned} S7 = {} & \frac{\rho^2}{2(1-\rho^2)^3}\left[I_i u_i^3 + (1-I_i)EE\right] - \frac{2\rho+\rho^3}{2(1-\rho^2)^3}\left[I_i u_i^2 AA + (1-I_i)FF\right] + \\ & + \frac{1+2\rho^2}{2(1-\rho^2)^3}\left[I_i u_i BB + (1-I_i)\rho GG\right] - \frac{\rho}{2(1-\rho^2)^3}\left[I_i CC + (1-I_i)GG\right] \end{aligned} \tag{61}$$

$$\begin{aligned} S8 = {} & -\frac{\rho^3}{6(1-\rho^2)^3}\left[I_i u_i^3 + (1-I_i)EE\right] + \frac{\rho^2}{2(1-\rho^2)^3}\left[I_i u_i^2 AA + (1-I_i)FF\right] - \\ & - \frac{\rho}{2(1-\rho^2)^3}\left[I_i u_i BB + (1-I_i)\rho GG\right] + \frac{1}{6(1-\rho^2)^3}\left[I_i CC + (1-I_i)GG\right] \end{aligned} \tag{62}$$

$$\begin{aligned} S9 = {} & \frac{1}{24(1-\rho^2)^4}\left[I_i u_i^4 + (1-I_i)HH - 3\right] - \frac{\rho}{6(1-\rho^2)^4}\left[I_i u_i^3 AA + (1-I_i)II - 3\rho\right] + \\ & + \frac{\rho^2}{4(1-\rho^2)^4}\left[I_i u_i^2 BB + (1-I_i)JJ - 2\rho^2 - 1\right] - \\ & - \frac{\rho^3}{6(1-\rho^2)^4}\left[I_i u_i CC + (1-I_i)\rho KK - 3\rho\right] + \frac{\rho^4}{24(1-\rho^2)^4}\left[I_i DD + (1-I_i)KK - 3\right] \end{aligned} \tag{63}$$

$$S10 = -\frac{\rho}{6(1-\rho^2)^4}\left[I_i u_i^4 + (1-I_i)HH - 3\right] + \frac{1+3\rho^2}{6(1-\rho^2)^4}\left[I_i u_i^3 AA + (1-I_i)II - 3\rho\right] -$$

$$-\frac{\rho+\rho^3}{2(1-\rho^2)^4}\left[I_i u_i^2 BB + (1-I_i)JJ - 2\rho^2 - 1\right] + \tag{64}$$

$$+\frac{3\rho^2+\rho^4}{6(1-\rho^2)^4}\left[I_i u_i CC + (1-I_i)\rho KK - 3\rho\right] - \frac{\rho^3}{6(1-\rho^2)^4}\left[I_i DD + (1-I_i)KK - 3\right]$$

$$S11 = \frac{\rho^2}{14(1-\rho^2)^4}\left[I_i u_i^4 + (1-I_i)HH - 3\right] - \frac{\rho+\rho^3}{7(1-\rho^2)^4}\left[I_i u_i^3 AA + (1-I_i)II - 3\rho\right] +$$

$$+\frac{1+4\rho^2+\rho^4}{14(1-\rho^2)^4}\left[I_i u_i^2 BB + (1-I_i)JJ - 2\rho^2 - 1\right] - \tag{65}$$

$$-\frac{\rho+\rho^3}{7(1-\rho^2)^4}\left[I_i u_i CC + (1-I_i)\rho KK - 3\rho\right] + \frac{\rho^2}{14(1-\rho^2)^4}\left[I_i DD + (1-I_i)KK - 3\right]$$

$$S12 = -\frac{\rho^3}{6(1-\rho^2)^4}\left[I_i u_i^4 + (1-I_i)HH - 3\right] + \frac{3\rho^2+\rho^4}{6(1-\rho^2)^4}\left[I_i u_i^3 AA + (1-I_i)II - 3\rho\right] -$$

$$-\frac{\rho+\rho^3}{2(1-\rho^2)^4}\left[I_i u_i^2 BB + (1-I_i)JJ - 2\rho^2 - 1\right] + \tag{66}$$

$$+\frac{1+3\rho^2}{6(1-\rho^2)^4}\left[I_i u_i CC + (1-I_i)\rho KK - 3\rho\right] - \frac{\rho}{6(1-\rho^2)^4}\left[I_i DD + (1-I_i)KK - 3\right]$$

$$S13 = \frac{\rho^4}{24(1-\rho^2)^4}\left[I_i u_i^4 + (1-I_i)HH - 3\right] - \frac{\rho^3}{6(1-\rho^2)^4}\left[I_i u_i^3 AA + (1-I_i)II - 3\rho\right] +$$

$$+\frac{\rho^2}{4(1-\rho^2)^4}\left[I_i u_i^2 BB + (1-I_i)JJ - 2\rho^2 - 1\right] - \tag{67}$$

$$-\frac{\rho}{6(1-\rho^2)^4}\left[I_i u_i CC + (1-I_i)\rho KK - 3\rho\right] + \frac{1}{24(1-\rho^2)^4}\left[I_i DD + (1-I_i)KK - 3\right]$$

Tedious integral calculus and highly involved algebra are necessary to show that in formulas (55) – (67):

$$AA = \rho u_i - \frac{\sqrt{1-\rho^2}\,\phi\left(\dfrac{\mathbf{z}_i'\boldsymbol{\gamma} - \rho u_i}{\sqrt{1-\rho^2}}\right)}{\Phi\left(\dfrac{\mathbf{z}_i'\boldsymbol{\gamma} - \rho u_i}{\sqrt{1-\rho^2}}\right)} \tag{68}$$

$$BB = \rho u_i AA + (1-\rho^2) - \frac{\sqrt{1-\rho^2}(\mathbf{z}_i'\boldsymbol{\gamma})\phi\left(\frac{\mathbf{z}_i'\boldsymbol{\gamma} - \rho u_i}{\sqrt{1-\rho^2}}\right)}{\Phi\left(\frac{\mathbf{z}_i'\boldsymbol{\gamma} - \rho u_i}{\sqrt{1-\rho^2}}\right)} \tag{69}$$

$$CC = \rho u_i BB + 2(1-\rho^2)AA - \frac{\sqrt{1-\rho^2}(\mathbf{z}_i'\boldsymbol{\gamma})^2\phi\left(\frac{\mathbf{z}_i'\boldsymbol{\gamma} - \rho u_i}{\sqrt{1-\rho^2}}\right)}{\Phi\left(\frac{\mathbf{z}_i'\boldsymbol{\gamma} - \rho u_i}{\sqrt{1-\rho^2}}\right)} \tag{70}$$

$$DD = \rho u_i CC + 3(1-\rho^2)BB - \frac{\sqrt{1-\rho^2}(\mathbf{z}_i'\boldsymbol{\gamma})^3\phi\left(\frac{\mathbf{z}_i'\boldsymbol{\gamma} - \rho u_i}{\sqrt{1-\rho^2}}\right)}{\Phi\left(\frac{\mathbf{z}_i'\boldsymbol{\gamma} - \rho u_i}{\sqrt{1-\rho^2}}\right)} \tag{71}$$

$$EE = \left[3\rho - \rho^3 + \rho^3(\mathbf{z}_i'\boldsymbol{\gamma})^2\right]\frac{\phi(\mathbf{z}_i'\boldsymbol{\gamma})}{1-\Phi(\mathbf{z}_i'\boldsymbol{\gamma})} \tag{72}$$

$$FF = \left[1 + \rho^2 + \rho^2(\mathbf{z}_i'\boldsymbol{\gamma})^2\right]\frac{\phi(\mathbf{z}_i'\boldsymbol{\gamma})}{1-\Phi(\mathbf{z}_i'\boldsymbol{\gamma})} \tag{73}$$

$$GG = \frac{\left[2 + (\mathbf{z}_i'\boldsymbol{\gamma})^2\right]\phi(\mathbf{z}_i'\boldsymbol{\gamma})}{1-\Phi(\mathbf{z}_i'\boldsymbol{\gamma})} \tag{74}$$

$$HH = 3 + \rho^2\left[3(2-\rho^2) + \rho^2(\mathbf{z}_i'\boldsymbol{\gamma})^2\right](\mathbf{z}_i'\boldsymbol{\gamma})\frac{\phi(\mathbf{z}_i'\boldsymbol{\gamma})}{1-\Phi(\mathbf{z}_i'\boldsymbol{\gamma})} \tag{75}$$

$$II = 3\rho + \rho\left[3 + \rho^2(\mathbf{z}_i'\boldsymbol{\gamma})^2\right](\mathbf{z}_i'\boldsymbol{\gamma})\frac{\phi(\mathbf{z}_i'\boldsymbol{\gamma})}{1-\Phi(\mathbf{z}_i'\boldsymbol{\gamma})} \tag{76}$$

$$JJ = 1 + 2\rho^2 + \left[1 + 2\rho^2 + \rho^2(\mathbf{z}_i'\boldsymbol{\gamma})^2\right](\mathbf{z}_i'\boldsymbol{\gamma})\frac{\phi(\mathbf{z}_i'\boldsymbol{\gamma})}{1-\Phi(\mathbf{z}_i'\boldsymbol{\gamma})} \tag{77}$$

$$KK = \frac{(\mathbf{z}_i'\boldsymbol{\gamma})^3\phi(\mathbf{z}_i'\boldsymbol{\gamma}) + 3(\mathbf{z}_i'\boldsymbol{\gamma})\phi(\mathbf{z}_i'\boldsymbol{\gamma}) - 3\Phi(\mathbf{z}_i'\boldsymbol{\gamma})}{1-\Phi(\mathbf{z}_i'\boldsymbol{\gamma})} \tag{78}$$

Formulas (55) – (78) provide the explicit form of the typical row of the matrix **S** and are sufficient for the calculation of ξ_{LM}. Given the complexity of the formulas hardly any doubts remain about the reason for systematic failure of most empirical researchers to try to test the "distributional" assumption in Tobit II models.

$$BB = \rho u_i AA + (1-\rho^2) - \frac{\sqrt{1-\rho^2}\,(\mathbf{z}_i' f)\,\phi\left(\frac{\mathbf{z}_i' f - \rho u_i}{\sqrt{1-\rho^2}}\right)}{\Phi\left(\frac{\mathbf{z}_i' f - \rho u_i}{\sqrt{1-\rho^2}}\right)} \tag{69}$$

$$CC = \rho u_i BB + 2(1-\rho^2)AA - \frac{\sqrt{1-\rho^2}\,(\mathbf{z}_i' f)^2\,\phi\left(\frac{\mathbf{z}_i' f - \rho u_i}{\sqrt{1-\rho^2}}\right)}{\Phi\left(\frac{\mathbf{z}_i' f - \rho u_i}{\sqrt{1-\rho^2}}\right)} \tag{70}$$

$$DD = \rho u_i CC + 3(1-\rho^2)BB - \frac{\sqrt{1-\rho^2}\,(\mathbf{z}_i' f)^3\,\phi\left(\frac{\mathbf{z}_i' f - \rho u_i}{\sqrt{1-\rho^2}}\right)}{\Phi\left(\frac{\mathbf{z}_i' f - \rho u_i}{\sqrt{1-\rho^2}}\right)} \tag{71}$$

$$EE = \left[3\rho - \rho^3 + \rho^3(\mathbf{z}_i' f)^2\right]\frac{\phi(\mathbf{z}_i' f)}{1-\Phi(\mathbf{z}_i' f)} \tag{72}$$

$$FF = \left[1 + \rho^2 + \rho^2(\mathbf{z}_i' f)^2\right]\frac{\phi(\mathbf{z}_i' f)}{1-\Phi(\mathbf{z}_i' f)} \tag{73}$$

$$GG = \frac{\left[2 + (\mathbf{z}_i' f)^2\right]\phi(\mathbf{z}_i' f)}{1-\Phi(\mathbf{z}_i' f)} \tag{74}$$

$$HH = 3 + \rho^2\left[3(2-\rho^2) + \rho^2(\mathbf{z}_i' f)^2\right](\mathbf{z}_i' f)\frac{\phi(\mathbf{z}_i' f)}{1-\Phi(\mathbf{z}_i' f)} \tag{75}$$

$$II = 3\rho + \rho\left[3 + \rho^2(\mathbf{z}_i' f)^2\right](\mathbf{z}_i' f)\frac{\phi(\mathbf{z}_i' f)}{1-\Phi(\mathbf{z}_i' f)} \tag{76}$$

$$JJ = 1 + 2\rho^2 + \left[1 + 2\rho^2 + \rho^2(\mathbf{z}_i' f)^2\right](\mathbf{z}_i' f)\frac{\phi(\mathbf{z}_i' f)}{1-\Phi(\mathbf{z}_i' f)} \tag{77}$$

$$KK = \frac{(\mathbf{z}_i' f)^3\phi(\mathbf{z}_i' f) + 3(\mathbf{z}_i' f)\phi(\mathbf{z}_i' f) - 3\Phi(\mathbf{z}_i' f)}{1-\Phi(\mathbf{z}_i' f)} \tag{78}$$

Formulas (55) – (78) provide the explicit form of the typical row of the matrix **S** and are sufficient for the calculation of ξ_{LM}. Given the complexity of the formulas hardly any doubts remain about the reason for systematic failure of most empirical researchers to try to test the "distributional" assumption in Tobit II models.

WORKS CITED

Ashenfelter, Orley, Alan Krueger. 1994. Estimates of the Economic Return to Schooling from a New Sample of Twins. *The American Economic Review*, Vol. 84, Issue 5 (Dec.): 1157-1173.

Barro, Robert J., Jong-Wha Lee. 2000. *International Data on Educational Attainment Updates and Implications.* NBER Working Paper Series. Working Paper #7911. Available at http://papers.nber.org/

Barro, Robert J., Jong-Wha Lee. 1997. *Schooling Quality in a Cross Section of Countries.* NBER Working Paper Series. Working Paper #6198. Available at http://papers.nber.org/

Becker, Gary S. 1993a. *Human Capital. A Theoretical and Empirical Analysis with Special Reference to Education.* 3d ed. Chicago and London: The University of Chicago Press.

Becker, Gary S. 1993b. Nobel Lecture: The Economic Way of Looking at Behavior. *Journal of Political Economy*, Vol. 101, Issue 3: 385-409.

Bessen, Jim. 1997. *Productivity Adjustments and Learning-by-Doing as Human Capital.* Economic Studies from US Census. Working Paper. Available at http://econpapers.hhs.se/scripts/redir.pl?u=http://www.ces.census.gov/paper.php?paper=101609&h=RePEc:wop:censes:97-17

Betts, Julian R. 1999. *Returns to Quality of Education*. World Bank Economics of Education Series. Available at http://www.worldbank.org/education/economicsed/research/econseries/betts.htm

Bils, Mark, Peter Klenow. 2000a. Does Schooling Cause Growth? *The American Economic Review*, Vol. 90, Number 5 (December): 1160-1183

Bils, Mark, Peter Klenow. 2000b. *Quantifying Quality Growth.* University of Rochester Working Paper. Available at http://www.econ.rochester.edu/Faculty/BilsPapers/BKQ7.pdf

Bowles, Samuel, Herbert Gintis. 1975. The Problem with Human Capital Theory–A Marxian Critique. *The American Economic Review,* Vol. 65, Issue 2, Papers and Proceedings of the Eighty-seventh Annual Meeting of the American Economic Association (May): 74-82.

Chiswick, Barry R. 1997. *Interpreting the Coefficient of Schooling in the Human Capital Earnings Function*. World Bank Economics of Education Series. Available at http://www.worldbank.org/education/economicsed/research/econseries/chiswick.pdf

Gallant, Ronald A., Douglas W. Nychka. 1987. Semi-Nonparametric Maximum Likelihood Estimation. *Econometrica*, Vol. 55, Issue 2 (Mar.): 363-390.

Greene, William H. 2000. *Econometric Analysis*. 4th ed. Upper Saddle River, New Jersey: Prentice-Hall, Inc.

Griliches, Zvi. 1996. *Education, Human Capital, and Growth: A Personal Perspective*. NBER Working Paper Series. Working Paper #5426. Available at http://www.nber.org

Hanushek, Eric A., Javier A. Luque. 2001. *Efficiency and Equity in Schools around the World*. World Bank Economics of Education Series. Available at http://www.worldbank.org/education/economicsed/research/econseries/hanushek.pdf

Heckman, James J. 1979. Sample Selection Bias as a Specification Error. *Econometrica*, Vol. 47, Issue 1 (Jan.): 153-162.

Khan, Shahrukh R. 1991. The Income Redistributional Impact of Financing Higher Education in Pakistan. *World Development*, Vol. 19, No. 9 : 1241-1246.

Klazar, Stanislav, Milan Sedmihradsky, Alena Vancurova. 2001. Returns to Education in the Czech Republic. *International Tax and Public Finance*, 8: 609-620.

Lee, Lung-Fei. 1984. Tests for the Bivariate Normal Distribution in Econometric Models with Selectivity. *Econometrica*, Vol. 52, Issue 4 (Jul.): 843-863.

Leschenko, Natalie. 2001. Returns to Education in Ukraine. *EERC MA Thesis Paper*.

Maddala, G. S. 1999. *Limited-dependent and Qualitative Variables in Econometrics*. Cambridge, UK: Cambridge University Press.

Mankiw, N. Gregory, David Romer, and David. N. Weil. 1992. A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics* 107 (May): 407-437.

Mincer, Jacob. 1958. Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, Vol. 66: 281-302.

Mincer, Jacob. 1974. *Schooling, Experience and Earnings*. New York: Columbia University Press (for NBER).

Mincer, Jacob. 1994. *The Production of Human Capital and The Lifecycle of Earnings: Variations on a Theme.* NBER Working Paper Series. Working Paper #4838. Available at http://www.nber.org

Mulligan, Casey B., Xavier Sala-i-Martin. 1995a. *A Labor-Income-Based Measure of the Value of Human Capital: An Application to the States of the United States*. NBER Working Paper Series. Working Paper #5018. Available at http://www.nber.org

Mulligan, Casey B., Xavier Sala-i-Martin. 1995b. *Measuring Aggregate Human Capital.* NBER Working Paper Series. Working Paper #5016. Available at http://www.nber.org

Nesterova, Daria, Klara Sabirianova. 1999. *Investment in Human Capital under Economic Transformation in Russia*. EERC Russia Working Paper Series. Working Paper #99/04E. Available at http://www.eerc.ru/publications/workpapers/WP_99-04e.pdf

Pagan, Adrian, Frank Vella. 1989. Diagnostic Test for Models Based on Individual Data: A Survey. *Journal of Applied Econometrics*, Vol. 4, Issue Supplement: Special Issue on Topics in Applied Econometrics (Dec.): S29-S59.

Parsons, Donald O. 1974. The Cost of School Time, Foregone Earnings, and Human Capital Formation. *The Journal of Political Economy,* Vol. 82, Issue 2, Part 1 (Mar. – Apr.): 251-266.

Psacharopoulos, George. 1994. Returns to Investment in Education: A Global Update. *World Development* 22: 1325-1343.

Psacharopoulos, George. 1995. *The Profitability of Investment in Education: Concepts and Methods*. The World Bank Working Paper Series. Available at http://www.worldbank.org

Ritzen, Jozef M., Donal R. Winkler. 1977. The Production of Human Capital Over Time. *The Review of Economics and Statistics,* Vol. 59, Issue 4 (Nov.): 427-437.

Sabirianova, Klara. 2001. *The Great Human Capital Reallocation: A Study of Occupational Mobility in Transitional Russia*. EERC Russia Working Paper Series. Working Paper #2K/11E. Available at http://www.eerc.ru/publications/workpapers/WP_2K-11e.pdf

Shyshkina, Anastasiya. 2001. The Returns to Education in Ukraine. *EERC MA Thesis Paper*.

Spence, Michael. 1973. Job Market Signaling. *Quarterly Journal of Economics*, Vol. 87, Issue 3 (Aug.): 355-374.

Stata User's Guide Release 6. 1999. College Station, Texas: Stata Press.

Sunde, Uwe. 2001. *Human Capital Accumulation, Education and Earnings Inequality*. The Institute for the Study of Labor (IZA). IZA Discussion Paper No. 310. Available at http://papers.ssrn.com/paper.taf?abstract_id=273042

Verbeek, Marno. 2000. *A Guide to Modern Econometrics*. Chichester, West Sussex, England: John Wiley & Sons, Ltd.

Weisbrod, Burton A. 1961. The Valuation of Human Capital. *The Journal of Political Economy,* Vol. 69, Issue 5 (Oct.): 425-436.

Woessmann, Ludger. 2000. *Specifying Human Capital: A Review, Some Extensions, and Development Effects*. Kiel Institute of World Economics. Working Paper No. 1007. Available at http://www.uni-kiel.de/ifw/pub/kap/2000/kap1007.pdf

Supplementary Materials

(with the exception of the Annotation (in Ukrainian) none of the following documents is a part of the thesis)

Thesis Annotation (in Ukrainian)

Thesis Presentation Notes

Thesis Presentation Handouts

State Examination Committee Vote

А Н О Т А Ц І Я
кваліфікаційної роботи

Тема: " Розміщення людського капіталу в перехідних економіках України та Росії"

Автор: Жилєвський Олександр Едуардович

Науковий керівник: Доктор Джаффар А. Мугаль

Захищена "___" червня 2002 р.

Короткий зміст роботи

Зважаючи на ключову теоретичну та практичну роль поняття людського капіталу, в роботі використано базові концептуальні засади, розроблені Бекером, Мінсером й іншими, на прикладі України та Росії. Модифіковано методологічні підходи Муллігана та Сала-і-Мартіна, Білза та Кленова й сконструйовано індекс людського капіталу на душу населення. На основі даних Опитування домогосподарств Держкомстату України та Російського поздовжнього контрольного опитування обраховано серію значень індексу з метою визначити та порівняти розміщення людського капіталу за сектором зайнятості, соціально-економічним статусом і статтю в 2000 році. Встановлено, що в обох країнах сектори освіти та науки мають найвищий рівень людського капіталу, а сектор торгівлі – найнижчий. Підприємці та самозайняті є порівняно менш кваліфікованими, ніж наймані працівники, що притаманно країнам, які розвиваються. Безробітні є найменш кваліфікованими, а жінки в більшості випадків мають вищий рівень людського капіталу, ніж чоловіки.

Presentation Notes

[Slide 1] *Introduction*

{Dear Chairman and} Dear members of the State Examination Committee!
Dear colleagues!
I am glad to present you my thesis on "Human Capital Allocation in the Transitional Economies of Ukraine and Russia".
Before I begin, I'd like to express my deep gratitude to Dr. Ghaffar Mughal who supervised me throughout the research.

[Slide 2] *Main Task*

The primary goal of the thesis is to determine and to compare the current allocation of human capital in Ukraine and Russia by sector of employment, economic status, and gender of labor force members and to offer related policy implications.

[Slide 3] *Contents*

In my presentation, I will answer what motivated me to undertake the research. Then, I will outline the key aspects of human capital theory, discuss the estimation methodology, present the results and summarize my findings.

[Slide 4] *Motivation*

Education is primarily an investment good. This statement has won recognition from the laity. Meanwhile, human capital theory has gained many supporters within academia.

The success of the theory stems from its fruitful use on the micro level, where it has helped to comprehend many aspects of observed behavior regarding education, health, occupational choice and mobility. On the macro level, the theory has facilitated understanding why development patterns vary across countries.

Transitional context makes the research a challenging intellectual endeavor. Transitional economies are in the process of *the great human capital reallocation*. This phenomenon is characterized by a discontinuous increase in occupational mobility, massive job destruction and creation, and dramatic swings in demand for various skills.

To a lesser extent, in this study, I was motivated by the scarcity of relevant literature on human capital allocation in transitional countries. Research work on the sectoral allocation of human capital within a particular economy appears to be virtually absent. Still, understanding the sectoral allocation of human capital is crucial if government is targeting a certain branch of the economy. Besides, it allows to make proximate predictions about the country's future performance.

[Slide 5] *Theory I: Mincer Regression*

Mincer earnings function relating the natural log of wage to educational attainment and labor-market experience lies at the heart of the human capital theory.
In my opinion, in the light of this theory, wage may be decomposed into four components:
raw wage,
human capital,
innate abilities and other non-human-capital characteristics components,
plus the purely random component.

[Slide 6] ***Theory II: Human Capital Stock per Capita***

Literature suggests several ways of assessing human capital stock, which range from very rough specifications to more refined proxies. Of the latter ones, two modern approaches that have solid theoretical foundation are due to Mulligan and Sala-i-Martin (1995) and Bils and Klenow (2000).

Defining human capital as the quality-adjusted stock of labor, Mulligan and Sala-i-Martin proceed to develop the Labor-income-based measure of human capital, where $\theta(s)$ is the quality parameter of a worker with s years of schooling, $w(s)$ is her wage, $\eta(s)$ is the share of such individuals in the labor force, and $w(0)$ is the wage of an unskilled laborer.

Mulligan and Sala-i-Martin's approach, which employs Mincer regression to estimate the wage of the unskilled, is similar to Bils and Klenow specification also shown on the slide. Here, e is the base of the natural log, $f(s)$ is some function of schooling, and $p(a\text{-}s)$ is a function of experience, as a denotes age.

In my thesis, I discuss the merits and demerits of each approach and choose Bils and Klenow specification as more conceptually appealing.

For estimation purposes, I specify the function of schooling to have a conventional quasi-linear form and the function of experience to have a conventional quadratic form.

Using the formulas for the rates of return to schooling I proceed to a "ready-to-be-computed" formula for the human capital stock per capita, where β's are to be estimated from Mincer regression.

[Slide 7] ***Conditional vs. Unconditional Returns***

The calculation of human capital indices would be straightforward, if the estimation were not fraught with the following pitfall.

Many researchers calculate the returns for the sub-sample of the employed. However, if there is a bias in selecting the employed out of labor force, the non-randomness of the sub-sample precludes us from extending the results to the unemployed and labor force as a whole.

Moreover, the returns by OLS or IV are not designed to capture the effect of education and experience on the *probability* of being employed.

Still, as evidence shows, the impact of education and experience on wages stems from two effects:

first, increase in probability of having positive earnings *per se*,

second, increase in earnings given that a person is already employed.

For this idea to be crystal clear, let me introduce the notions of conditional and unconditional returns, which are already implicit in many studies, but are not named as such.

So, conditional returns measure the effect of education and experience on earnings provided that an individual is already employed.

Unconditional returns additionally take into account the impact of skills on observing a positive wage *per se* and, thus, measure the "full" impact of education and experience on earnings.

As the standard Tobit II procedure makes use of the probability of observing wage and corrects for the sample selection bias, returns by Tobit II are unconditional and applicable to the sub-samples of the employed and unemployed.

To tentatively check for the difference in conditional and unconditional returns, I suggest estimating a simple logit model, where the probability of having job is supposed to depend on individual's schooling, experience, demographic characteristics and place of residence.

[Slide 8] ***Model: OLS vs. Tobit II***
As argued, the parameters of interest may be obtained by estimating the Tobit II model, also known as Heckman's selection model. The model consists of two equations:
the wage equation, relating the log of earnings to educational attainment, experience, demographic characteristics, enterprise ownership, and place of residence dummies;
and the selection equation, where the probability of being selected into the pool of the employed depends on the variables affecting the reservation wage and the chance of finding job.
Tobit II model is closed by imposing a standard bivariate normality assumption on the joint distribution of the error terms of the two equations.
For comparability with other studies, I also estimate the wage equation by OLS to obtain conditional returns.

[Slide 9] ***Econometric Digression***
Apart from the conventional tests for the significance of coefficients, the main tests to conduct after estimating Tobit II model are as follows.
First, it is necessary to check whether the correlation coefficient between the error terms of the wage and the selection equations is different from zero. In this case, the sample selection bias arises and OLS results do not apply to the labor force.
Second, since the consistency of the Tobit II estimator, as any other maximum likelihood estimator, depends on the validity of the distributional assumption, it is crucially important to test for specification error. I should emphasize that most researchers do not even try to check Tobit II's specification, but I take the issue seriously and provide the highly non-trivial derivation of the test statistic in Appendix F of the thesis. The test checks whether the third- and fourth-order central moments of the observed joint distribution behave as if it were, indeed, bivariate normal.

[Slide 10] ***Datasets***
Let me now proceed to the empirical section.
The data for the research for Ukraine comes from the micro-file of the Derzhkomstat's survey conducted in the year 2000.
For Russia, the data is taken from the micro-file of RLMS Round IX and confidential files of the survey.
The number of observations in the original micro-files as well as the number of observations in the cleaned samples are shown in the table.

[Slide 11] ***Sectors and Socio-economic Groups***
In my thesis, I follow the Derzhkomstat's classification of the branches of the economy, and on the basis of the information about primary job I assign each employed to one of the shown fifteen sectors or to the non-classified category.
Next, each labor force member is supposed to hold exactly one economic status:
to be an employee,
to be an employer, entrepreneur, or self-employed,
or to be unemployed.

[Slide 12] ***Logit Results***
Estimation of the logit model reveals that schooling and experience positively affect the chance of being employed in both samples.
This preliminary test suggests a difference between conditional and unconditional returns.

[Slide 13] ***OLS Results***
On this slide, I present the results of OLS estimation of the wage equation.
As evident, with the exception of primary education in Russian sample, the coefficients of interest are rather precisely estimated. All coefficients have expected sign, including negative one for the square of the experience, and possess reasonable magnitude.

[Slide 14] ***Tobit II: Wage Equation***
Tobit II estimation of the wage equation has similar features. Here, all coefficients of interest are significant at a conventional level, they also have expected sign and reasonable magnitude.
As shown in the bottom, for both samples the hypothesis of zero correlation between the error terms of the wage and the selection equation is rejected. This implies the presence of the sample selection bias and invalidates returns by OLS.
It should be also said that in conducting the specification test, I failed to reject the hypothesis of bivariate normality.

[Slide 15] ***Returns***
On the basis of OLS and Tobit II estimations, I have calculated the rates of return to one year schooling at each particular level. Though returns to human capital investments are <u>not</u> the primary focus of my research, several issues are worth noting.
First, unconditional returns for most educational levels are higher than conditional returns. This reveals the positive effect of skills on the probability of having job.
Second, diminishing returns to education are not evident in Ukrainian data, but manifest themselves in Russian sample.
Third, the rates in Ukraine are fairly moderate, smaller than world averages. Still, they fall into the range reported by other researchers. Russian rates are approximately 1.5 times higher and close to the pattern of advanced economies. This may reflect a more favorable state of the Russian labor market in the year 2000 and signify the relative severity of the problem of skill misallocation in Ukraine.
Fourth, I do rectify the results of previous estimations for Ukraine, in particular, done last year in EERC MA theses. I show that returns are definitely not negative and that the return to incomplete higher education is far below those for complete and fundamental higher. This is intuitively and theoretically plausible as being a reflection of the "sheepskin effect".
In addition, I have estimated the experience-earnings profiles, which appear to be more favorable in Russia.

[Slide 16] ***Results: Human Capital Indices***
Given Tobit II estimates, I have obtained sectoral allocation of human capital stock per capita and allocation of human capital by economic status and gender. This is the main result of my thesis.
The calculated indices allowed me to compare the sectoral rankings by human capital stock, average income, and average age.
So as not to clutter the presentation with many tables, I directly proceed to the discussion of the obtained results and policy implications.

[Slide 17] ***Discussion I***

As Russian rates of return are unambiguously higher, Ukraine demonstrates lower human capital indices in absolute magnitude.

In both samples, the unemployed are the least skilled. Unexpectedly, the group of employers, entrepreneurs, and self-employed has lower stock of human capital than the employees. This bears some resemblance to the pattern of developing countries. There, the less-skilled, to make for their living, have to undertake low-productivity self-employment jobs as better options (in terms of wage and gaining valuable experience) are not available.

Women have somewhat higher stock of human capital than men. This phenomenon is observed for each economic status with the exception of unemployed women in Russia.

[Slide 18] ***Discussion II***

In the two countries, the sectors with highest human capital stock per capita are science and education, as expected. The sector where relatively low skilled individuals are concentrated is wholesale and retail trade, which is rather surprising. Sectoral allocation of human capital by gender is more diversified in Russia.

Inspection of income and age rankings of sectors reveals alarming prospects of Ukraine's future. Education and science are not attracting younger labor force members because of, most likely, very low remuneration.

Possibly, the two sectors will shrink and lose their top positions during the next decade as their employees start retiring on a large scale. This may have strong negative implications for the national economy's long-term growth, as education and science are the sectors where human capital is generated and new knowledge is created.

Probably, immediate corrective effort by the government is needed. A possible suggestion would be the change in the wage-setting schemes in the sectors to encourage the young to pursue careers there.

[Slide 19] ***Thank you***
Having expressed my concerns about Ukraine's future, I would like to thank you for your attention.

Theory

Mincer Earnings Function

$$\ln W_{S\,,Exp} = \ln w_0 + \sum_{j=1}^{q} r_j S_j + \gamma_1 Exp + \gamma_2 Exp^2 + V$$

W – Wage
w_0 – compensation for "raw labor"
r_j – rate of return to one year of schooling at educational level j
S_j – number of years of schooling at level j
Exp – experience
$\gamma_1 \geq 0,\ \ \gamma_2 \leq 0$ are supposed "to capture" positive, but decreasing returns to experience
V – error term
q – number of educational levels.

Conceptual view of wage within the human capital theory framework:

$$W = \underbrace{w_0}_{\text{raw labor component of wage}} \cdot \underbrace{e^{f(\textit{education},\ \textit{experience},\ \textit{tenure},\ \textit{other HC components})}}_{\text{human capital component of wage}} \cdot \underbrace{e^{g(\textit{innate abilities},\ \textit{static/dynamic characteristics})}}_{\text{innate abilities and other non-human capital characteristics' component of wage}} \cdot \underbrace{e^{\varepsilon}}_{\text{purely random component of wage, error term}}$$

Human Capital Stock Per Capita

Mulligan and Sala-i-Martin (1995):

$$h = \int_0^\infty \theta(s)\eta(s)ds = \int_0^\infty \frac{w(s)\eta(s)ds}{w(0)}$$

$\theta(s)$ – efficiency parameter according to which a worker with s years of schooling contributes to the stock,
$w(s)$ – wage of a worker with s years of schooling,
$\eta(s)$ – share of individuals with s years of schooling.

Bils and Klenow (2000):

$$h(a,s) = e^{f(s)+p(a-s)}$$

$f()$ – function of schooling, s
$p()$ – function of experience (a is age).

May be expanded as (per capita human capital stock in a cohort):

$$h = e^{\sum_{j=1}^{q} r_j \bar{S}_j + \gamma_1 \overline{Exp} + \gamma_2 \overline{Exp}^2} \text{; (easy to show)} \Rightarrow\ h_i = e^{\sum_{j=1}^{q} \beta_{1j}\eta_{ji} + \beta_2 \overline{Exp_i} + \beta_3 \overline{Exp_i}^2}$$

$\bar{S}_j$ – average years of schooling at level j
$\overline{Exp}$ – average experience
η_{ji} is the proportion of individuals with *highest* educational level j in sub-sample i.

Estimation Methodology

Conditional vs. Unconditional Returns

Conditional returns measure the effect of education and experience on earnings provided that an individual is already employed.
Unconditional returns additionally take into account the impact of skills on observing a positive wage per se. Unconditional rates of returns, thus, measure the "full" impact of education and experience on earnings and apply both to the sub-sample of employed and to the one of unemployed.

Preliminary Test for Difference of Conditional and Unconditional Returns: ***Logit Model***

$$y = \text{logit}(\alpha_0 + \alpha_1 S + \alpha_2 Exp + \sum_{l=1}^{L} \alpha_{3l} Dem_l + \sum_{k=1}^{K} \alpha_{4k} Res_k)$$

y=1 if an individual is employed, 0 otherwise
S – # of years of schooling
Exp – # of years of working experience
Dem_l's: *age* (in years), *sex* (1 if male), *sxmr* (1 if married man)
Res_k's: (residence dummies): *city*, *town, Kyiv/Moscow*.

Conditional Returns: ***OLS***

$$\ln W = \beta_0 + \sum_{j=1}^{q} \beta_{1j} Ed_j + \beta_2 Exp + \beta_3 Exp^2 + \sum_{l=1}^{L} \beta_{4l} Dem_l + \sum_{p=1}^{P} \beta_{5p} Prop_p + \sum_{k=1}^{K} \beta_{6k} Res_k + \varepsilon$$

Ed_j*'s* (educational dummies) are: *chigh* (for complete higher education as the highest educational level obtained), *ihigh* (incomplete higher), *fhigh* (fundamental higher), *ptu* (vocational training), *ssec* (specialized secondary), *csec* (complete secondary), *fsec* (fundamental secondary), *prim* (primary), and uneducated are the benchmark category (*q* is the total number of educational levels);

$Prop_p$*'s* (ownership type dummies) are: *foreign* (foreign or joint-stock enterprises), *nongov* (domestic non-governmental enterprises/organizations), state-owned enterprises are the benchmark;

Res_k's (place of residence dummies) additionally include: *KyivOblast* and *MoscowOblast*.
Specification for Russia does not include *sex* (found to be insignificant in various specifications).

Unconditional Returns: ***Tobit II*** (Heckman selection):
The first equation (wage equation):

$$\ln W = \beta_0 + \sum_{j=1}^{q} \beta_{1j} Ed_j + \beta_2 Exp + \beta_3 Exp^2 + \sum_{l=1}^{L} \beta_{4l} Dem_l + \sum_{p=1}^{P} \beta_{5p} Prop_p + \sum_{k=1}^{K} \beta_{6k} Res_k + \varepsilon_1$$

The second equation (selection equation):

$$y = \delta_0 + \sum_{j=1}^{q} \delta_{1j} Ed_j + \delta_2 Exp + \delta_3 Exp^2 + \sum_{l=1}^{L} \delta_{4l} Dem_l + \sum_{k=1}^{K} \delta_{5k} Res_k + \varepsilon_2$$

To insure convergence several specifications of selection equation must be tried. In particular, I had to exclude *prim* and *fsec* for Ukraine, *MoscowOblast* and constant term for Russia.

"Distributional" assumption:

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \sim NIID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & 1 \end{pmatrix} \right)$$

Results

Data

Ukraine: "The Survey on Households' Standards of Living", Derzhkomstat, 2000.
Russia: Russia Longitudinal Monitoring Survey, Round IX, 2000, RLMS confidential files.
# of observations in Ukrainian sample: 8838; in Russian sample: 4681.

Logit Model

Ukraine:

variable	Coefficient	P-value	Marginal effect	P-value
sch_years	**0.2053**	**0.000**	**0.0210**	**0.000**
exp	**0.1186**	**0.000**	**0.0122**	**0.000**
age	-0.0663	0.000	-0.0068	0.000
sex	-0.3655	0.000	-0.0376	0.000
sxmr	1.0571	0.000	0.1007	0.000
city	0.8196	0.000	0.0802	0.000
town	0.2466	0.001	0.0244	0.001
Kyiv	0.4434	0.039	0.0390	0.014

Russia:

variable	Coefficient	P-value	Marginal effect	P-value
sch_years	**0.0792**	**0.000**	**0.0096**	**0.000**
exp	**0.1102**	**0.000**	**0.0133**	**0.000**
age	-0.0803	0.000	-0.0097	0.000
sex	-0.6031	0.000	-0.0736	0.000
sxmr	0.6098	0.000	0.0704	0.000
city	0.8638	0.000	0.0923	0.000
town	0.7073	0.000	0.0783	0.000
Moscow	-0.5612	0.027	-0.0813	0.060

OLS

	Ukraine		Russia	
Variable	Coefficient	P-value	Coefficient	P-value
chigh	**0.7614**	**0.000**	**1.6324**	**0.000**
ihigh	**0.5030**	**0.000**	**1.3037**	**0.000**
fhigh	**0.6779**	**0.000**	**1.5167**	**0.000**
ptu	**0.4999**	**0.000**	**1.1139**	**0.001**
ssec	**0.5724**	**0.000**	**1.2378**	**0.000**
csec	**0.4210**	**0.000**	**1.0541**	**0.002**
fsec	**0.3515**	**0.000**	**0.7205**	**0.033**
prim	**0.3634**	**0.000**	**0.6497**	**0.063**
exp	**0.0154**	**0.000**	**0.0682**	**0.000**
exp2	**-0.0004**	**0.000**	**-0.0008**	**0.000**
age	-0.0005	0.755	-0.0361	0.000
sex	0.1332	0.000	–	–
sxmr	0.1937	0.000	0.5142	0.000
foreign	0.4131	0.000	0.3158	0.013
nongov	0.0491	0.000	0.2846	0.000
city	0.4050	0.000	0.4902	0.000
town	0.2794	0.000	0.0933	0.021
Kyiv/Moscow	0.2785	0.000	0.2094	0.005
Kyiv/MoscowOblast	0.2129	0.000	0.5638	0.000
_cons	6.3354	0.000	5.7773	0.000
Number of obs.	7424		3910	
R^2	0.2283		0.2280	
P-value of F-stat.	0.0000		0.0000	

Tobit II

	Ukraine				Russia			
	Wage		Selection		Wage		Selection	
Variable	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
chigh	**0.8747**	**0.000**	0.8131	0.000	**1.6741**	**0.000**	1.6126	0.000
ihigh	**0.6268**	**0.000**	0.6745	0.000	**1.3413**	**0.000**	1.3185	0.000
fhigh	**0.7676**	**0.000**	0.5369	0.000	**1.5594**	**0.000**	1.7162	0.000
ptu	**0.5430**	**0.000**	0.3490	0.000	**1.1518**	**0.000**	1.3610	0.000
ssec	**0.6334**	**0.000**	0.4395	0.000	**1.2780**	**0.000**	1.4810	0.000
csec	**0.4184**	**0.000**	0.1178	0.004	**1.0896**	**0.001**	1.2193	0.000
fsec	**0.3186**	**0.000**	–	–	**0.7568**	**0.024**	1.2931	0.000
prim	**0.3584**	**0.000**	–	–	**0.6860**	**0.048**	1.2607	0.000
exp	**0.0358**	**0.000**	0.0954	0.000	**0.0696**	**0.000**	0.0720	0.000
exp2	**-0.0006**	**0.000**	-0.0009	0.000	**-0.0008**	**0.000**	-0.0003	0.026
age	-0.0084	0.000	-0.0400	0.000	-0.0368	0.000	-0.0448	0.000
sex	0.0708	0.007	-0.1134	0.001	–	–	-0.3120	0.000
sxmr	0.3066	0.000	0.6070	0.000	0.5150	0.000	0.3131	0.000
foreign	0.3537	0.000	–	–	0.3153	0.012	–	–
nongov	0.0350	0.005	–	–	0.2842	0.000	–	–
city	0.4796	0.000	0.4828	0.000	0.4977	0.000	0.4974	0.000
town	0.3016	0.000	0.2020	0.000	0.0998	0.014	0.4039	0.000
Kyiv/Moscow	0.2972	0.000	0.3080	0.000	0.2050	0.006	-0.3067	0.026
Kyiv/MoscowOblast	0.2148	0.000	0.1622	0.003	0.5639	0.000	–	–
_cons	6.1253	0.000	0.5640	0.000	5.7312	0.000	–	–
Censored obs.	1414				771			
Uncensored obs.	7424				3910			
P-value of Wald chi2	0.0000				0.0000			
rho=0: P-value of chi2	**0.0000**				**0.0449**			

Returns to Education (to one year of schooling at each particular level):

Level	Conditional		Unconditional	
	Ukraine	Russia	Ukraine	Russia
Complete higher +	6.25	10.21	8.37	10.32
Incomplete higher	2.27	6.90	5.75	6.95
Fundamental higher	6.62	11.86	9.00	12.04
Vocational training	4.77	12.98	7.21	13.03
Specialized secondary	5.62	13.18	8.01	13.28
Complete secondary	3.77	18.54	5.43	18.50
Fundamental secondary	4.20	8.49	3.81	8.92
Primary	9.37	19.06	9.24	20.12

Income-experience profile (marginal effects)

Conditional returns

Unconditional returns

***Normalized Human Capital Indices*: by Sector**

Ukraine

Code	Sector	total	men		women	
		NT	NT	NM	NT	NW
1	Industry	1.02	1.02	1.03	1.03	1.02
2	Agriculture +	0.98	0.94	0.95	1.00	0.99
3	Construction	1.00	0.98	0.99	1.07	1.06
4	Transport +	1.00	0.99	1.00	1.03	1.02
5	Trade +	0.95	0.93	0.94	0.96	0.96
6	Business services	1.00	0.98	0.99	1.03	1.02
7	IT services	1.08	1.04	1.05	1.08	1.07
8	Housing +	1.01	1.01	1.02	1.01	1.00
9	Health care +	1.04	1.12	1.13	1.03	1.02
10	Education	1.15	1.23	1.24	1.12	1.11
11	Culture and arts	1.11	1.17	1.18	1.08	1.07
12	Science	1.23	1.24	1.25	1.23	1.22
13	Finance +	1.10	1.11	1.12	1.09	1.08
14	Government	1.11	1.14	1.15	1.10	1.09
15	Defense +	1.01	1.00	1.01	1.01	1.00
16	Non-classified	1.11	1.14	1.15	1.06	1.05

Russia

Code	Sector	total	men		women	
		NT	NT	NM	NT	NW
1	Industry	1.08	1.12	1.15	1.02	0.99
2	Agriculture +	0.99	1.00	1.03	0.99	0.96
3	Construction	1.00	0.99	1.02	1.04	1.01
4	Transport +	0.91	0.88	0.91	1.05	1.03
5	Trade +	0.86	0.79	0.81	0.88	0.86
6	Business services	1.14	0.98	1.01	1.17	1.14
7	IT services	0.93	0.97	1.00	0.89	0.87
8	Housing +	1.09	1.09	1.12	1.09	1.06
9	Health care +	1.14	1.37	1.40	1.11	1.08
10	Education	1.40	1.70	1.74	1.36	1.33
11	Culture and arts	1.12	1.08	1.11	1.15	1.12
12	Science	1.35	1.33	1.36	1.37	1.34
13	Finance +	1.23	1.30	1.34	1.22	1.19
14	Government	1.23	1.42	1.45	1.21	1.19
15	Defense +	0.91	0.87	0.89	1.23	1.20
16	Non-classified	0.85	0.59	0.61	0.99	0.97

by Status:

Ukraine

Category	total	men		women	
	NT	NT	NM	NT	NW
employees	1.04	1.03	1.04	1.05	1.04
employers +	0.93	0.92	0.93	0.95	0.94
unemployed	0.84	0.83	0.84	0.85	0.84
whole sample	1.00	0.99	1.00	1.01	1.00

Russia

Category	total	men		women	
	NT	NT	NM	NT	NW
employees	1.06	1.03	1.05	1.08	1.06
employers +	1.00	0.97	1.00	1.03	1.01
unemployed	0.77	0.77	0.79	0.76	0.74
whole sample	1.00	0.98	1.00	1.02	1.00

Rankings: Ukraine

Code	Sector	Rank of Index	Rank of Income	Rank of EdC	Rank of Age
1	Industry	9	7	12	5
2	Agriculture +	15	14	16	3
3	Construction	14	6	13	11
4	Transport +	13	9	14	9
5	Trade +	16	11	11	15
6	Business services	12	1	7	16
7	IT services	7	4	5	13
8	Housing +	10	12	15	2
9	Health care +	8	15	10	6
10	Education	2	13	2	4
11	Culture and arts	5	16	4	10
12	Science	1	10	1	1
13	Finance +	6	2	8	8
14	Government	3	8	3	12
15	Defense +	11	5	9	14
16	Non-classified	4	3	6	7

Rankings: Russia

Code	Sector	Rank of Index	Rank of Income	Rank of EdC	Rank of Age
1	Industry	9	9	11	4
2	Agriculture +	11	13	16	3
3	Construction	10	3	14	6
4	Transport +	13	8	15	12
5	Trade +	15	10	10	14
6	Business services	6	11	9	2
7	IT services	12	4	4	15
8	Housing +	8	12	12	1
9	Health care +	5	16	7	10
10	Education	1	15	1	5
11	Culture and arts	7	14	8	11
12	Science	2	5	2	8
13	Finance +	4	7	5	9
14	Government	3	6	3	7
15	Defense +	14	2	6	16
16	Non-classified	16	1	13	13

Discussion

- The pool of the employed is **not** a random draw from the labor force. The sample selection bias implies that returns by OLS are not applicable to the whole labor force.
- Unconditional rates are higher than conditional ones. Diminishing returns are not observed for Ukraine, but are evident for Russian sample. Ukrainian rates are fairly moderate. Russian rates, on average, are approximately 1.5 times higher and even close to the pattern of OECD.
- Russia has higher cumulative returns. The experience-earnings profile is more "favorable" in Russia. Consequently, human capital indices are lower in Ukraine.
- According to the calculated indices, of the three socio-economic groups, the unemployed are the least skilled. Unexpectedly, the group of employers, entrepreneurs, and self-employed has lower stock of human capital per capita than the employees.
- Women have higher stock of human capital per capita than men. This phenomenon is observed for all socio-economic groups with the exception of unemployed women in Russia.
- The sectors with the highest human capital stock are science and education. Trade has the lowest index. Sectoral distribution of human capital conditional on gender is more diversified in Russia.
- Inspection of income and age rankings reveals gloomy prospects of Ukraine's future. Education and science are not attracting younger labor force members because of, most likely, very low remuneration. The two sectors may shrink and lose their top positions in 8-13 years. This may have negative implications for the economy's growth and development. Probably, immediate corrective effort by the government is needed.

State Examination Committee Vote

Date: June 13th, 2002
NaUKMA, Kyiv

Members of the State Examination Committee:

- Dr. Serhiy Korablin (Chairman), Institute of Economic Forecasting at the National Academy of Sciences of Ukraine
- Dr. Roy J. Gardner, Indiana University and Stanford University, EERC Visiting Professor
- Dr. Stefan Lutz, EERC Visiting Professor
- Dr. Lilia Maliar, University of Alicante (Spain), EERC Visiting Professor
- Dr. Serguei Maliar, University of Alicante (Spain), EERC Visiting Professor
- Dr. Ghaffar Mughal (Thesis Advisor), EERC Visiting Professor

External Reader:

Dr. Christopher Waller, University of Kentucky

Committee Secretary:

Yaroslava Naimushyna

Vote by secret ballot:

NaUKMA Grade (out of 100): 99
State Grade: excellent
EERC Grade: A