THE PRODUCTIVITY EFFECT OF INNOVATION: EVIDENCE FROM EASTERN EUROPE AND CENTRAL ASIA

By

Anastasiya Drul

Submitted to

Central European University

Department of Economics

In partial fulfillment of the requirements for the degree of Master of Arts in Economics

Supervisor: Professor Ádám Reiff

Budapest, Hungary

2017

Abstract

This study addresses the importance of innovation in the production process and analyses its effect on firm performance. The research is conducted for Eastern European and Central Asian firms, using micro-level data that altogether covers three inconsecutive years (2001, 2004 and 2008). In order to empirically investigate the relationship between firm productivity, measured by firm's total factor productivity (TFP), and innovation, proxied by four variables: amount of spending on R&D, dummy variables capturing an introduction of new products, investment in R&D and upgrade of an existing product during the last three years, FE procedure is used in Stage I estimation for two models, with Cobb-Douglas and translog production functions, respectively. Further, pooled OLS regressions for multiple model specifications are estimated in Stage II. Numbers of firm and industry characteristics that affect productivity are effectively controlled for. Final Stage II model specification suggests that there is a positive and statistically significant link between TFP and innovation spending, however economically marginal, while the innovation dummy for R&D activities over the last three years proves to have both statistically and economically significant positive effect on firm's productivity.

Acknowledgements

First of all, I would like to express the sincere gratitude to my supervisor, Professor Ádám Reiff, for his time, patience, valuable advises and constant assistance throughout the entire thesis writing process. Also, I would like to thank CEU and the Department of Economics for the exceptionally great academic environment; my family and friends for their help and support.

Table of contents

List of Tables

List of Figures

Introduction

Typing "innovation" or "why innovate" in Google search generates over 500 million results, which translates into approximately one internet user out of seven asking the searching engine that question. Why is our interest in innovation so genuinely extensive? Mainly because the concept of innovation is often viewed from a strategic management perspective. Innovation is considered to be a reaction to an inevitable change, which gives a privilege of staying on the top and riding the change rather than simply accepting and following it. Given the sufficiently long track record of humanity's interest in innovation and high current level of attention this topic receives, I decided to investigate its importance by means of economic research.

The nature of innovation, studied in a conjuncture with economic development and growth, has changed greatly over the last decades and is no longer considered to be an exclusive "first world" activity (Fagerberg et al., 2009). The word innovation or technological change is now meant to capture more than purely extraordinary high-tech inventions inherent to leading market players who heavily invest in R&D. The broader perspective allowed new series of research to emerge that now include a wider range of industries and countries of interest.

The main purpose of this research is to assess the importance of innovation in the production process and to quantify the effect from different types of innovation on productivity growth. So far, the majority of existing studies on innovation have been conducted for the scope of one country. In this regard, my study stands out since I investigate the relationship between innovation and productivity for numerous countries of Eastern Europe and Central Asia. This novelty is particularly possible due to the comparable enterprise survey by World Bank on firm-level innovation activities in developing economies.

To investigate the relationship of interest I employ two-stage estimation procedure, with fixed effects and pooled OLS methodologies applied to Stage I and Stage II regressions, respectively. Final stage models prove R&D expenditures to have a positive statistically significant effect on firm's productivity that, however, economically does not differ from zero. Results show that innovation gains its economic significance through another measure – a binary variable on R&D activities over the latest three years – as it successfully accounts for a time-delay due to the learning process.

The sections of the thesis are organized as follows. Section I presents an overview of the existing literature, main theoretical and empirical concepts used in the research on innovation and productivity. Section II introduces determinants of firm's productivity other than innovation and provides the description of the dataset used in the research. In Section III employed methodology is described. Empirical findings from the regressions are discussed in Section IV. Finally, Section V concludes the study.

I. Literature overview

For ages, production of optimal performance has remained among the top questions of interest for economists. In particular, the channels through which this optimal stage may be achieved have been studied for further policy implications, and the force of invention and technological change has been inevitably among them.

"The idea that technological progress facilitated economic growth and improved welfare was appreciated long before economists became concerned with quantifying its impact" (Cohen and Levin, 1989). Classical economists Ricardo (1821) and Mill (1848) acknowledge the importance of technological progress for economic growth. Later Marx devotes a lot of attention to the analysis of technological change and arrives at the conclusion of its great importance for capitalist production (Marx and Engels, 1867).

A fundamental problem many studies on innovation faced over a prolonged period of time was a lack of data that would adequately measure the contribution of innovation to the production process. In 1960's, patents were usually used as a proxy for innovative output of countries and industries. However, this measurement of productivity presented some significant problems to the proper econometrical usage. Firstly, many patents never reach a point of being actually exploited; secondly, differences in filling regulations limit researchers from using patents for cross-country comparison; thirdly, cross-industry differences, like preferences for copyright instead of patent, put additional constraints on between-industry comparison (Cohen and Levin, 1989). Furthermore, patents, as a measurement of innovativeness, had little implication in the firm level research.

In 1974, the Financial Accounting Standards Board[1] established a new accounting standard that required all public and private companies in the U.S. that report financial statements

[1]Financial accounting Standards Board – is a private, non-profit organization that establishes standards of accounting and financial reporting in the U.S.

also disclose their R&D expenses. The above improvement of the financial reporting practices took place in other developed countries outside of the U.S., which consequently led to better data availability on innovation activities at an enterprise level. Further, higher quality disclosure on firm's innovation activities shifted a focus from country/industry level research to one which took a firm as a unit of analysis.

Generally, all firm level studies on inventive activities and their effect on productivity can be divided into two categories based on the estimation methodology approach. The first category adopts a modified version of the Cobb-Douglas production function model with three input parameters, namely capital, labour and knowledge (or R&D) capital. Improvements in technology captured by knowledge capital variable in the model are most commonly measured as deflated R&D expenditures (Vivero, 2002) or the rate of return to R&D (Wakelin, 2001). One of the main advantages of this approach is that it does not assume constant returns to scale.

However, this estimation procedure is not useful for this particular research given I am interested in quantifying the effect of innovation while also controlling for other policy variables. Hence my research is conducted in accordance with the following approach. It comprises two stages of estimation: Stage I focuses on the computation of firm's total factor productivity, and Stage II is devoted solely to the estimation of the effects innovation and other firm/industry/country specific variables have on productivity improvements.

Odagiri and Iwata (1986) successfully apply latter to the investigation of an impact R&D expenditures have on the rate of total factor productivity increase on a sample of listed Japanese manufacturing companies. In Stage I they estimate the rate of change in total factor productivity as a function of the real value added, the number of employees, the real capital assets and the ratio of total employee compensations over value added. In Stage II the authors regress the rate of change in total factor productivity on the ratio of R&D expenditures to value added. Importantly, the independent variable of interest is measured by an average of ratios of R&D spending to value added in the first two years of each sample period, hence allowing for a lag of several years from the time of R&D spending to the time of actual productivity improvement. The study finds an evidence of positive effect from R&D spending, which is estimated in the range of 17%-20%, depending on the time period.

While developed part of the world has been continuously exercising effort to understand innovation forces, develop new frameworks and measures that would enhance the understanding of those since 1960's (Grilliches, 1984), the developing countries have cached up with this trend not that long time ago. The majority of research on developing countries covers the time period starting in 2000's. Benefiting this research are more comprehensive data collections of innovation variables constructed from various waves of innovation surveys conducted by local communities and international organizations.

Fazliogu et al (2016) conduct a research on productivity changes driven by different "typologies of innovation" for Turkish manufacturing firms. Exploiting the rich dataset, the authors assess the effect of more than one variable that represent firm's innovative activities. In particular, they differentiate between innovative outputs and innovative inputs, by defining inputs as internal/external R&D expenditures, investments in innovative machinery or acquisition of external knowledge; and outputs as process/product/service innovation, organizational or marketing innovation. Given a wider set of innovation variables, authors' findings are more granular and applicable for policy authorities. They find innovative outputs have more direct effect on productivity improvements than innovative inputs, which is likely due to the time-delays. Also, authors confirm the difference between in-house and outsourced R&D: given internal R&D activities contribute to the growth of firm's tacit knowledge they consequently benefit productivity more. For the output innovations, they find a descending structure of productivity improvements starting from process innovations yielding the highest effect and then gradually decreasing for organizational, product/service and finally marketing innovations.

As mentioned previously, the skeleton of my study is consistent with the second estimation approach. Importantly, this research is unique in terms of the variable selection for Stage II estimation, which is described in details in the next section.

II. Data

The main objective of this research is to quantify the connection between firm's productivity and its innovativeness in countries of Eastern Europe and Central Asia (ECA)[2], using a firm-level panel dataset that covers enterprises from ECA countries for three inconsecutive yeas (2001, 2004 and 2008).

The main data source is the World Bank Group Enterprise Surveys Database[3], a firm-level survey of a representative sample of an economy's private sector, which was specifically requested for the purpose of this economic research. It covers a broad range of business environment topics including access to finance, corruption, infrastructure, competition, and performance measures. The full data sets are available to researchers and include all questions from the surveys at the firm level. All variables used in the estimation process are listed in Table 1. Table of descriptive statistics on the variables can be found in Appendix.

Stage I variables are used for TFP estimation, while Stage II variables are used to define the effect of innovation on productivity. The main independent variable of interest, the level of innovation by a firm, is hardly observable; for this reason, I use four proxy variables to capture its effect on productivity. The first is R&D USD, measured as dollar spending on research and development activities including wages and salaries of R&D personnel, materials, R&D related education and training costs at the end of a fiscal year.

[2]Albania, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Georgia, Hungary, Kazakhstan, Kyrgyz Republic, Latvia, Lithuania, Moldova, Montenegro, Poland, Republic of Macedonia, Romania, Russian Federation, Serbia, Slovak Republic, Slovenia, Tajikistan, Ukraine and Uzbekistan
[3] http://www.enterprisesurveys.org//data

Table 1 Notation and Definition of Variables

	Variable	Explanation
Stage I	Y	The firm's total annual sales at the end of a fiscal year, USD (000's)
	K	Cost for the firm to repurchase all of its machinery, land and buildings at the end of fiscal year, USD (000's)
	L	Number of permanent, full-time workers employed at the firm at the end of fiscal year
Stage II	R&D USD	Amount spent on research and development either in-house or contracted with other companies in the last fiscal year, USD (000's)
	R&D1	=1 if firm 'introduced new products or services in the last 3 years', 0 otherwise
	R&D2	=1 if firm 'invested in research and development in the last three years', 0 otherwise
	R&D3	=1 if firm 'upgraded an existing product line or service in the last 3 years', 0 otherwise
	Corruption	=1 if corruption is not considered an obstacle to current firm's operations, 0 otherwise
	Finance	=1 if access to finance is not considered an obstacle to current firm's operations, 0 otherwise
	Licence	=1 if process of obtaining business licencing and permits is considered as a severe obstacle to as obstacle to current firm's operations, 0 otherwise
	Competition	=1 if pressure from domestic competition has a very important effect on firm's decisions, 0 otherwise
	Export	Percentage of the firm's sales attributed to direct/indirect exports
	K/L ratio	Capital-to-labour ratio
	Utilization	Level of utilization of facilities and manpower
	Skilled	Percentage of skilled production workers as share of permanent, full-time workers
	Size	Size of the firm (small, medium, large) with respect to number of full-time employees, dummy variable
	Year	Year dummy variable
	Country	Country dummy variable
	Industry	Industry dummy variable[4]

[4]Industries included in the study: food, textiles, garments, chemicals, plastics & rubber, non-metallic mineral products, basic metals, fabricate metal products, machinery and equipment, electronics, construction, other services, hotel and restaurants, transport, IT and other manufacturers.

The second is RD1, a dummy variable that captures introduction of new products or services by a firm over the last three years. The third is RD2, also a dummy variable that indicates firm's involvement in any kind of R&D activities over the last three years. And lastly, RD3, a dummy variable for an upgrade of an already existing product line over the last three years. Importantly, there are numerous other firm- or industry-specific variables that are proven to have effect on firm's productivity.

Recently, a line of research has emerged that focuses on exploring the causality between corruption and productivity. De Rosa et al (2010) argue there is no harmful evidence of the bribery on firm-level productivity; however, when controlling for EU membership the authors find that non-EU countries (where informal payments to the government are more common) experience negative impact from bribery. To proxy for corruption I use the data provided by firms that grades the "bribe tax" on the scale between "no impact" and "severe impact" on firm operations.

Given an extensive amount of research on how increased financial development positively affects output growth (Goldstein 1969, Levin 2005), I also include a dummy describing the level of ease of access to financing for new investments.

Djankov et al (2006) conduct a research on linkage between business regulatory environment and economic performance and find economies with more favourable business regulations have higher output growth prospects. To capture the business environment's effect on enterprise performance I use the dummy variable that defines whether business licencing and permits are an obstacle to proper conduct of current operations or not.

Taking into account empirical (Gort, 1999) and theoretical (Aghion et al, 2013) research on competition that agrees on its positive productivity impact, I include the dummy variable that evaluates the effect from domestic competition on firm's decision making.

In accordance with the hypothesis that variation in productivity is to some extent associated with variation in labour skills, suggested by Haskel (2005), I also include a share of skilled production workers in TFP estimation.

To test a relationship between the knowledge flow from the international market participants and internal improvements of the firm, I include export variable into Stage II regression. One may note that export and TFP variables are likely to be codetermined, and hence my estimation of Stage II on solely basis can be hampered by endogeneity. The simultaneity problem is usually solved by using a lagged independent variable as an instrument variable, however, data limitations are such that this procedure is impossible. The risk of export endogeneity is mitigated by Loecker (2007) who finds an evidence of productivity gains due to exporting, and not vice versa, for Slovenian firms. While his finding is somewhat inconsistent with previous research conducted for developed countries that fails to find an evidence of such causality (Bernard and Wagner, 1995), it must be substantially due to significant differences between developed and developing economies. Loecker's research is devoted to the analysis of a transition economy, while Bernard and Wagner's paper focused on a market economy of the U.S. Since the former also holds true for economies of other countries from my research sample, I argue that export endogeneity is not a concern in my research.

Importantly, reverse causality is not an issue in case of previously discussed independent variables since they are either defined at a country level (corruption, licence, finance); or at an industry level (competition), which inherently makes them exogenous.

Other variables, deemed important for the firm's productivity and included in the model as exogenous, are the following: a dummy indicator of firm's size defined by the full-time employee base; country of operations, the main area of activity in terms of sales; a year dummy; capital-to-labour ratio; and level of current output as a fraction of the maximum output possible using the firm's facilities/manpower at the time.

III. Methodology

In order to empirically investigate the main relationship of interest, namely the effect innovations have on the productivity, I estimate the following regression:

$$
\begin{aligned}
\ln(\mathrm{TFP}) = {} & \mathrm{const} + \propto_1 (\mathrm{RD\ USD}) + \propto_2 (\mathrm{RD1}) + \propto_3 (\mathrm{RD2}) + \propto_4 (\mathrm{RD3}) \\
& + \beta_1(\mathrm{Corruprion}) + \beta_2(\mathrm{Finance}) + \beta_3(\mathrm{Licence}) + \beta_4(\mathrm{Competition}) + \gamma_1\left(\frac{\mathrm{K}}{\mathrm{L}}\mathrm{Ratio}\right) \\
& + \gamma_2(\mathrm{Export}) + \gamma_3(\mathrm{Utilization}) + \gamma_1(\mathrm{Skilled}) \\
& + \sum_{s=1}^{2} \theta_{1s}(\mathrm{Size}) + \sum_{y=1}^{2} \theta_{2y}(\mathrm{Year}) + \sum_{1}^{26} \theta_{2c}(\mathrm{Country}) + \sum_{1}^{17} \theta_{2i}(\mathrm{Industry}) + \mathrm{u}
\end{aligned}
$$

The dependent variable from the above equation, the total factor productivity, can be tracked back to the seminal paper by Solow (1957), in which the author presents a basic model of annual aggregate output. Solow breaks down the growth of output into the growth of production factors and the efficiency in the utilization of the above (further referred to as total factor productivity), using the economic model based on a neoclassical production function:

$$Y(t) = A_t F(K_t, L_t)$$

where *Y(t)* stands for the total production of the economy in year *t*, usually measured as economy's GDP, *K(t)* and *L(t)* are capital and labour of the productive economy, respectively, measured through the combined value of the companies and number of people employed, and *A(t)* represents TFP, or technology, and appears in a Hicks neutral way.

Since the origins of TFP, the renewed interest in theoretical and empirical studies emerged with the increased availability of the firm-level data, allowing for a shift from the aggregate level estimation technics to individual level estimation studies (Van Beveren, 2010). The roadmap for the estimation procedure used in this research is the following:

In line with Solow's approach, firm-level productivity studies assume the output to be a function of the inputs and productivity. Firstly, I assume the production function takes the form of a general Cobb-Douglas production function with two input variables, analogously to the Solow model:

$$Y_{it} = A_{it}{K_{it}}^{\beta_1}{L_{it}}^{\beta_2}$$

where *Y* represents physical output of firm *i* in period *t*, *K* is a capital input and L is a labour input. While *Y, K, L* are observables, *A*, TFP or efficiency level of the firm, is unobservable. Taking logs of the above equations results in a linear production function:

$$y_{it} = \beta_0 + \beta_1 k_{it} + \beta_2 l_{it} + \varepsilon_{it}$$

where$ln(A_{it}) = \beta_0 + \varepsilon_{it}and\beta_0$measures the mean TFP across firms and over time; ε_{it} is the time and firm-specific deviation from TFP mean, which is further decomposed into predictable and unobservable components, v_{it}andu_{it}.Decomposition results in the following equation:

$$y_{it} = \beta_0 + \beta_1 k_{it} + \beta_2 l_{it} + v_{it} + u_{it}$$

where $w_{it} = \beta_0 + v_{it}$represents firm-level total factor productivity, and$u_{it}$is an i.i.d. component, standing for unexpected deviation from the mean. The firm's TFP is hence estimated as follows, and can be further used to evaluate the influence and impact of various policy variables directly at the firm level:

$$\widehat{w}_{it} = \hat{\beta}_0 + \hat{v}_{it} = y_{it} - (\widehat{\beta_1} k_{it} + \widehat{\beta_2} l_{it}).$$

Simple OLS least squares can be used to estimate the above equation, however, given possible endogeneity of input variables, this can lead to biased estimated coefficients. Fixed effects estimation procedure can be used as an effective way to overcome omitted variable bias, conditional on time-invariance of the firm-specific TFPs, w_{it}. The estimated regression is then as follows:

$$y_{it} = \beta_0 + \beta_1 k_{it} + \beta_2 l_{it} + w_i + u_{it}^q$$

The level of firm productivity is then obtained as a residual from the above functional relationship and is further used as depending variable in the assessment of innovation impact.

To make sure the correct production function specification is chosen in Stage I estimation, I also run first stage regression for transcendental logarithmic production function, or translog production function, in lieu of previously discussed Cobb-Douglas production function. According to Klacek et al (2007), this flexible function form of production function has several advantages over the Cobb-Douglas, since it does not necessarily assume unit elasticity of substitution between production factors or perfect competition in market for those. The generalized form of translog production function that takes into account *n* production factors, can be expressed as:

$$y_{it} = \alpha_0 + \sum_{i=1}^{n} \alpha_i x_{it} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij}\, x_{it} x_{jt} + \delta_{it},$$

where x_i's are the logarithms of production factors.

Another obvious advantage of using the production function in the translog form is that it allows to expand a linear relationship between the output and production inputs to non-linear one.

Simplifying the general form in accordance with previous specification and accounting for two inputs, namely *K* and *L,* the above function can be rewritten as:

$$y_{it} = \beta_0 + \beta_1 k_{it} + \beta_2 l_{it} + \gamma_{11} k_{it}^2 + \gamma_{22} l_{it}^2 + \gamma_{12} k_{it} l_{it} + \delta_{it}$$

Analogously to the Cobb-Douglas case, the translog production function cannot be estimated by simple OLS given its proneness to omitted variable bias, and hence the estimation procedure will follow same fixed effects methodology.

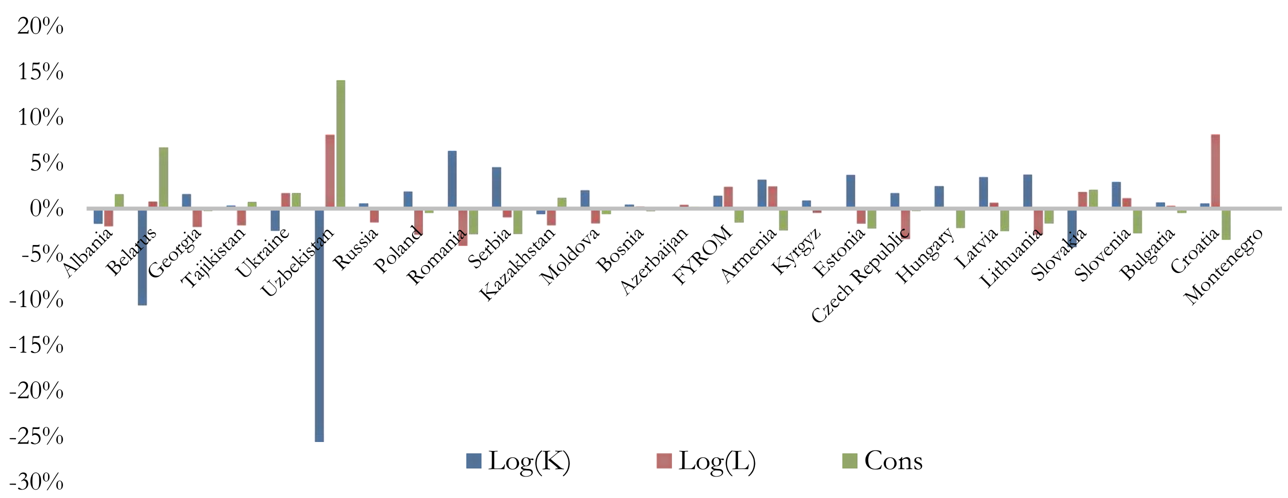
IV. Estimation results

Prior to choosing a model specification, I must address several limitations the data set introduces to the estimation process. Firstly, in Stage I estimation, firm's total annual sales are used as a dependent variable due to the absence of the output data. Usually, such an issue is solved simply by inclusion of the third input variable that would serve as a proxy for the cost of raw materials and intermediate goods used in production process. Given that data on materials' expenditure is very limited, I use a different approach in delimitation of the above mismatch. I exclude firms from the wholesale and retail industries from the sample due to their high inherent dependence on the material inputs and straightforward correlation between the amount of those and eventual sales.

Secondly, I perform a data check to decrease the likelihood of estimation discrepancies coming from the sample outliers. For this purpose, I use a simple Cobb-Douglas production function with two input parameters, Log (K) and Log (L), as Stage I regression model, and check its robustness to exclusion of country samples. Fixed effects estimation procedure is used.

As per results charted in Figure 1, I find that elimination of Uzbekistan from the sample causes 25%, 8% and 14% deviation in capital, labour and intercept estimates from the full sample estimates. In particular, I find that Uzbekistani firms have negative estimated returns to scale on labour and high estimated returns to scale on capital and materials, which consequently make Model 2 estimates on labour and capital underestimated and overestimated, respectively.

Figure 1 Deviations of estimated coefficients in case of one country exclusion



Also, as can be seen in Table 2, six other countries on the top of Uzbekistan – Russia, Serbia, Estonia, Czech Republic, Hungary and Slovenia – suffer from negative signs on one of the estimated returns to scale. These discrepancies might be due to the poor quality of data in the above countries, as proven by major difference in the number of observations per country and the number used in FE estimation.

Table 2 Country specific estimates

Log(Y)	*Uzbekistan*	*Russia*	*Serbia*	*Estonia*	*Czech Republic*	*Hungary*	*Slovenia*
Log(K)	0.8950***	**-0.4552**	**-0.0266**	**-0.0268**	**-0.0676**	**-0.0509**	0.1404***
Log(L)	**-0.6294**	0.8700*	0.94725***	0.9106*	0.82161***	1.0074	**-0.0765**
Cons	4.6229	16.2636	10.6979***	11.2320***	11.9369***	11.4823***	12.8402***
Total number of observations	593	1,646	617	468	639	878	509
Number of observation used in estimation	243	527	272	253	372	562	292

The primary motivation for this data cleaning exercise comes from the fact that the sample of 27 countries is likely to suffer from some heterogeneity. Partially it is mitigated by similarity of recovery patterns after the major political and economic disruptions of the 1990s in Eastern European and Central Asian countries. Elimination of Uzbekistan, which likely due to the data quality issues showed such disruptive results, benefits the estimation process through a more homogeneous sample.

Table 3 Preliminary Stage I estimation results

Log(Y)	*Model 1*	*Model 2*	*Model 3*
Log(K)	0.5138***	0.3178***	-0.463***
Log(L)	0.5391***	0.576***	0.4989
$Log(K)^2$			0.0285***
$Log(L)^2$			0.0037
Log(K)Log(L)			0.0054
Cons	4.768***	7.183***	12.31***

Note: Here and in all further tables asterisks denote significance levels (***: p<0.1%; **: p<1%; *: p<5%)

Further, I run three model specifications for the estimation of total factor productivity in Stage I on an adjusted sample: a benchmark model with the Cobb-Douglas production function estimated by pooled OLS method (Model 1), the Cobb-Douglas production function with fixed effects (Model 2) and the translog production function with fixed effects (Model 3). Estimated results are presented in Table 3.

Given Model 3 estimates have counterintuitive signs I check the translog production function for concavity. For this purpose I plot $F\left(Log(K), \overline{Log(L),}\right)$ and $F\left(\overline{Log(K)}, Log(L)\right)$, where *F* is the translog production function (visualisation of the above functions can be found in Appendix). Both functions are convex which means the traslog production function is misspecified in this particular setup. Hence I decide to proceed with Model 2 in further estimation.

Model 2 suggest that output elasticities of capital and labour equal 0.32 and 0.58, respectively. More comprehensive output of Model 2 results is presented in Table 4.

As noted in Table 4, Stage I regression is estimated based on *6,070* observations which fall under one of *5,595* groups, based on *panel_id* variable. This means that the fixed effects estimation procedure does not effectively eliminate all omitted variable bias that is captured by time-invariant country specific unobservable error terms.

Table 4 Final Stage I estimation results

Fixed-effects regression

Number of observations = 6070
Number of groups = 5595

Log(Y)	***Coefficient***	***Std. Err.***	***P>t***
Log(K)	0.3178	0.0245	0.0000
Log(L)	0.5760	0.0743	0.0000
Cons	7.1832	0.3789	0.0000

F test that all u_i=0: F(5594, 473) = 1.60
Prob> F = 0.0000

A large number of groups and a sizable number of missing observations for different firms over two or three years result in an average number of observations per group of 1.1, which effectively translates into 10% of total observations having data for at most two years. The remaining 90% of the sample observations are hence estimated as usual cross-sectional data with varying time index. As discussed in the methodology section, using cross-sectional OLS approach may result in biased estimates of model inputs. The risk of having unfeasible estimates is somewhat mitigated by the fact that above results are very much in line with the literature, hence I conclude that even if the bias is present in the above regression it is marginal in terms of estimated outputs. Consequently, the magnitude it might bring to the estimates is not economically significant and therefore acceptable.

Log(TFP) is further defined as the predicted residual from the above model and used as dependent variable in Stage II estimation. Stage II methodology follows a pooled OSL method, given the data construction limits the effective usage of more advanced techniques. Apart from the fact that the data sample is strongly unbalanced which puts constraints on fixed effects procedure, the main dummy variables of interest – namely RD1, RD2 and RD3 –do not change for majority of firms. Given the lack of year over year variation in the above variables and also limitation on inclusion of time-invariant industry or country specific variable into fixed effects estimation procedure, I decide to proceed with OLS methodology. Being aware of the possible omitted variable problem OLS estimation technique may cause, I deliberately control for as many possible explanatory exogenous variables as data collection allows.

Another point to note, since there are four variables in the above regression which are meant to capture the effect of innovation on productivity, there might arise an issue of multicollinearity, given it is likely that firms engage in different innovation activities following similar patterns (process innovation or product innovation).

Table 5 Correlation matrix of innovation inputs

	RD USD	RD1	RD2	RD3
RD USD	1			
RD1	0.041	1		
RD2	0.0529	0.3237	1	
RD3	0.0273	0.4794	0.2915	1

To identify any high correlation that further might cause coefficients to be poorly estimated, I examine the correlations between the above four innovation variables to detect a high level of association. As per results presented in Table 5, one can conclude that there is no evidence of harmful bivariate correlation between RD USD variable and any of the remaining binary innovation variables, while there is a much stronger correlation among binary variables themselves. This must be mainly due to the fact that RD USD captures current spending, while other RD variables represent innovative engagement over the last three years. The correlation between and among binary innovation variables RD1, RD2, and RD3 may result in an unfavourable increase in the variance of the regression coefficients, I include only one dummy innovation variable to the regression at a time to insure against any multicollinearity issues.

The results from Stage II regressions are presented for numerous model specifications in Table 6 (a) – (c). The first relevant thing to note is that as I add exogenous explanatory variables to the models of interest, the signs of estimated parameters mainly remain consistent as they gain or lose their statistical and/or economic significance.

In particular, research and development spending captured by RD USD variable appears positive and statistically significant in each of the considered models, however, its economic significance is decreasing in a number of independent variables (inclusion of year, country and industry dummies have the strongest decreasing effect).

As can be seen in Table 6 (a), the introduction of new products or services, RD1, loses its statistical significance as I start to control for other than innovation variables (Model B – Model F).

Table 6 (a) Estimated results for different model specifications with RD1

Log(TFP)	*Model A*	*Model B*	*Model C*	*Model D*	*Model E*	*Model F*
RD USD	4.1e-08***	4.1e-08***	4.0e-08***	3.6e-08***	3.0e-08***	2.9e-08***
RD1	.12**	0.0532	0.0326	-0.0558	-0.0219	-0.0308
Corruption		-0.0516	-0.0582	-0.0533	-.1027**	-.1011**
Finance		0.0372	0.0226	0.0692	0.064	0.0659
Licence		.814***	.7855***	0.2237	.4693*	.4887*
Competition		.1903***	.195***	.1697***	0.031	0.0245
K/L Ratio		-.0133*	0.0125	0.004	-.0457***	-.045***
Export		.0038***	.0026***	.0018*	-5.90E-04	-8.40E-06
Utilization		.0027**	.0029**	.0044***	.0017*	.0017*
Skilled		0.0739	0.0351	-0.0306	-.1846**	-.1686**
Size						
Medium			.1437**	.1543**	.143**	.1165*
Large			.4221***	.4163***	.3842***	.3373***
Year						
2004				.3881***	.3796***	.4391***
2008				.9994***	1.04***	1.085***
Country	No	No	No	No	Yes	Yes
Sector	No	No	No	No	No	Yes
Cons	-.1351***	-.3924***	-.6354***	-1.003***	-0.2107	-.4052**

While its sign in the final model specification may seem counterintuitive, it can be economically reasoned in the following way: usually the introduction of the new products/services is accompanied by higher marketing and advertising costs, which consequently results in a negative effect on productivity. Notably, as I control for more dummy variables which are not firm specific the negative effect becomes more pronounced.

Results presented in Table 6 (b) suggest that any research and development activity has a strongly statistically significant positive effect on productivity, which also compares well to the estimated coefficient on RD USD. These two innovation measures are capturing the same activity, with two differences: time lag in RD2 and different measurements of two.

Table 6 (b) Estimated results for different model specifications with RD2

Log(TFP)	*Model A*	*Model B*	*Model C*	*Model D*	*Model E*	*Model F*
RD USD	4.0e-08***	4.0e-08***	4.0e-08***	3.6e-08***	3.0e-08***	2.9e-08***
RD2	.2544***	.253***	.1971***	.3312***	.1949***	.1571**
Corruption		-0.0543	-0.0593	-0.0519	-.1004**	-.0985**
Finance		0.0472	0.0325	0.0784	0.068	0.0669
Licence		.718**	.7134**	0.2257	.474*	.4898*
Competition		.2039***	.2048***	.1677***	0.0329	0.0257
K/L Ratio		-0.0025	.0165*	0.006	-.0437***	-.0434***
Export		.0033***	.0024**	.0016*	-6.90E-04	-5.50E-05
Utilization		.0033***	.0034***	.0044***	.0017*	.0017*
Skilled		0.0642	0.0334	-0.0384	-.183**	-.1641**
Size						
Medium			.1133*	0.1008	.1158*	.0957*
Large			.3546***	.3223***	.3344***	.2991***
Year						
2004				.5893***	.4957***	.5342***
2008				.9421***	1.009***	1.066***
Country	No	No	No	No	Yes	Yes
Sector	No	No	No	No	No	Yes
Cons	-.1916***	-.5727***	-.7361***	-1.267***	-.3639**	-.5461***

Unlike RD2, RD3 that captures any upgrade of an existing product or service line, and is estimated in Table 6 (c), consistently has a positive sign. In line with the previous economic reasoning for RD1 estimates, it suggests that marketing spending are not that or at all common in the event of the product upgrade. Consequently, the coefficient on RD3 captures the correct causality effect, however is not statistically significant.

Final Stage II model specifications with varying RD dummies, are presented in Table 7. Apart from variables listed in the table, "F" specification also controls for country and sector dummies. Please refer to Appendix for the full table of estimated coefficients.

As previously discussed, final Stage II estimates suggest that research and development spending has a positive and statistically very significant impact on a firm's productivity.

Table 6 (c) Estimated results for different model specifications with RD3

Log(TFP)	*Model A*	*Model B*	*Model C*	*Model D*	*Model E*	*Model F*
RD USD	4.1e-08***	4.1e-08***	4.0e-08***	3.6e-08***	3.0e-08***	2.9e-08***
RD3	.1708***	.1111*	.0964*	0.0211	0.0138	0.0094
Corruption		-0.0508	-0.057	-0.0513	-.1019**	-.0997**
Finance		0.0441	0.0292	0.0721	0.0662	0.0679
Licence		.8035***	.7722***	0.2209	.4683*	.4879*
Competition		.1863***	.1913***	.1675***	0.0306	0.0244
K/L Ratio		-0.0119	0.0139	0.0046	-.0455***	-.0448***
Export		.0037***	.0025***	.0017*	-6.30E-04	-3.30E-05
Utilization		.0027**	.0029**	.0044***	.0016*	.0017*
Skilled		0.0712	0.0337	-0.027	-.1844**	-.1681**
Size						
Medium			.1437**	.1491**	.1396**	.1128*
Large			.4189***	.4094***	.3801***	.3329***
Year						
2004				.3932***	.38***	.4401***
2008				.9779***	1.031***	1.076***
Country	No	No	No	No	Yes	Yes
Sector	No	No	No	No	No	Yes
Cons	-.1902***	-.4417***	-.6862***	-1.045***	-0.2253	-.4266**

Economically the effect is not strongly significant, since the estimated coefficient translates only into 0.00029 basis point increase in average TFP driven by $100k increase in the amount of R&D spending, other things being equal. The result makes economic sense, given the high cost of R&D activities, high percentage of failures and usually prolonged implementation period of successfully developed new products or services. In line with RD USD, RD2 is also positive and statistically significant. On average, firm's engagement in any kind of R&D activities over three-year horizon results in a 0.16 percentage point increase in its productivity, ceteris paribus.

The corruption variable appears with a negative sign, meaning firms that do not consider corruption an obstacle to their operations are performing 10 basis points worse than others. Keeping in mind that countries from my sample are listed with an average corruption perception index[5] of 20 by Transparency International, the above should be interpreted as follows. It is highly likely that the enterprise which reports the corruption as such that doesn't have an adverse effect on firm's performance is actually engaged in some kind of bribery activity, consequently obtained sign is consistent with the literature. Licence falls under same economic interpretation and, in line with corruption, is weakly statistically significant.

Table 7 Final Stage II estimation results

Log(TFP)	*Model F:* RD1	*Model F:* RD2	*Model F:* RD3
RD USD	2.9e-08***	2.9e-08***	2.9e-08***
RDi, i=1,3	-0.0308	.1571**	0.0094
Corruption	-.1011**	-.0985**	-.0997**
Finance	0.0659	0.0669	0.0679
Licence	.4887*	.4898*	.4879*
Competition	0.0245	0.0257	0.0244
K/L Ratio	-.045***	-.0434***	-.0448***
Export	-8.40E-06	-5.50E-05	-3.3E-05
Utilization	.0017*	.0017*	.0017*
Skilled	-.1686**	-.1641**	-.1681**
Size			
Medium	.1165*	.0957*	.1128*
Large	.3373***	.2991***	.3329***
Year			
2004	.4391***	.5342***	.4401***
2008	1.085***	1.066***	1.076***
Cons	-.4052**	-.5461***	-.4266**

The finance variable has a positive, however statistically insignificant estimates. Competition's estimates are in line with the literature in terms of the sign, however also lack statistical significance in this particular model setup.

Utilization estimates have a positive and weakly significant effect on productivity. It means that as the level of utilization increases by 10 percentage points, for instance from 50% to 60%,

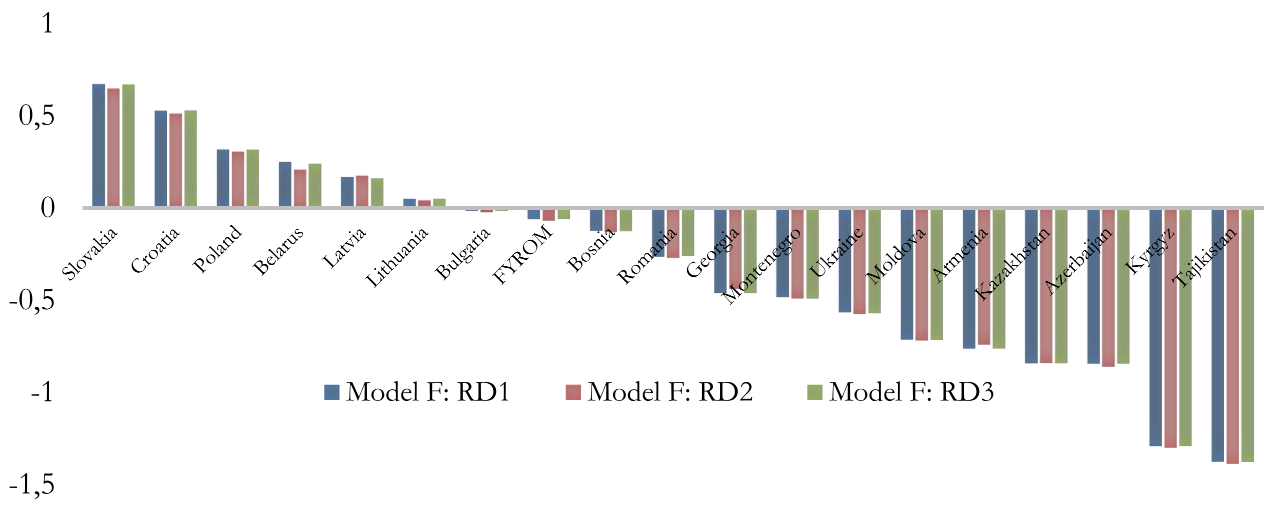
[5] 0 – highly corrupt ,100 – very clean

productivity increase on average by 1.7 basis points. Intuitively it means that firm's current output is lower when compared to the maximum output possible, conditional on using same facilities and manpower. The firms' average inability to utilize resources to the fullest consequently results in negative estimates on the number of skilled production workers and capital-to-labour ratio.

Positive and more importantly increasing estimates on size dummies advocate in favour of large enterprises' ability to exercise economies of scale. Time dummies indicate the productivity increased over the time period covered in this research.

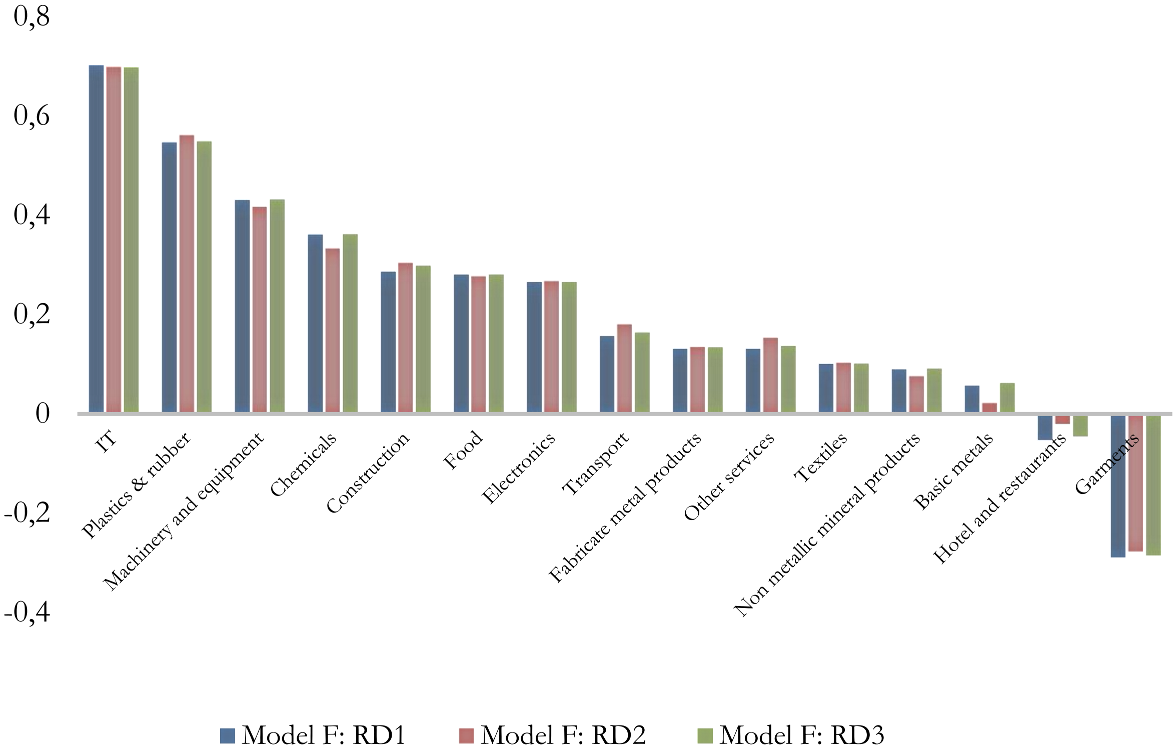
Not surprisingly, top highest country effects are attributed to the member states of the European Union and Belarus, one of the most industrially developed states by percentage of GDP at the time of the dissolution of the Soviet Union. Albania is used here as a reference group for country dummies.

Figure 2 Country effects



When controlling for industry effects and using "other manufacturing" industry as a reference group, I find that all industries perform better in terms of productivity than those which fall under other manufacturers category, with an exception of hotel& restaurants and garments industries that have negative estimates. IT industry is leading with the highest fixed effect of 0.7.

Figure 3 Industry effects



V. Conclusions

This thesis contributes to a better understanding of firm's innovativeness and of its effect on productivity. Usage of different proxies for innovation measurement allowed me to estimate the effect from various typologies of innovation. The overall picture from the analysis confirms that innovation has a positive effect on firm's productivity and importantly this effect gains economic significance as time passes. It means that while spending on research and development activities have an economically marginal immediate effect on TFP, albeit statistically significant and positive, its lagged effect is much stronger. This time delay can be attributed to the learning activities which are essential in the process of translating innovation into productivity gains. The relationship between productivity growth and output innovation is found to be extremely sensitive to the inclusion of country, industry and year dummies. As such, final stage models find no evidence of positive effect on productivity improvements from output innovation, measured by an introduction of new products/services or by an upgrade of existing ones.

Importantly, the results of this study suggest that the role of research and development in Eastern European and Central Asian countries is similar to the previous findings for more developed countries, like the U.S. or Japan. From a policy perspective, this finding may contribute to the promotion of innovation for the enhancement of productivity at the country level.

Bibliography

Aghion, P., Akcigit, U., & Howitt, P. (2013). What Do We Learn From Schumpeterian Growth Theory?

Bernard, A., & Jensen, J. B. (1997). Exceptional Exporter Performance: Cause, Effect, or Both?

Beveren, I. V. (2010). Total Factor Productivity Estimation: A Practical Review. Journal of Economic Surveys,26(1), 98-128.

Bound, J., Cummins, C., Griliches, Z., Hall, B., & Jaffe, A. (1982). Who Does R&D and Who Patents?

Cieślik, A., Gauger, I., &Michałek, J. J. (2017). Determinants of Productivity of Ukrainian Firms. Comparative Economic Research,20(1).

Crepon, B., Duguet, E., &Mairesse, J. (1998). Research, Innovation, and Productivity: An Econometric Analysis at the Firm Level.

Dalgic, B., Fazlioglu, B., &Yereli, A. B. (2016) The Effect of Innovation on Productivity:Evidence from Turkish ManufacturingFirms. Munich Personal RePEc Archive, 75773.

Djankov, S., Mcliesh, C., &Ramalho, R. M. (n.d.). Regulation and Growth. SSRN Electronic Journal

Fagerberg, J., Srholec, M., & Verspagen, B. (2010). Innovation and Economic Development. Handbook of the Economics of Innovation, Volume 2 Handbook of the Economics of Innovation, 833-872.

Gort, M., & Sung, N. (1999). Competition And Productivity Growth: The Case Of The U.s. Telephone Industry. Economic Inquiry, 37(4), 678-691.

Griliches, Z. (1984). R&D, Patents and Productivity.

Hall, B. (2011). Innovation and Productivity.

Loecker, J. D. (2007). Do exports generate higher productivity? Evidence from Slovenia. Journal of International Economics, 73(1), 69-98

Odagiri, H., & Iwata, H. (1986). The impact of R&D on productivity increase in Japanese manufacturing companies. Research Policy, 15(1), 13-19.

Rosa, D. D., Gooroochurn, N., &Gorg, H. (2010). Corruption and Productivity: Firm-Level Evidence from The BEEPS Survey. Policy Research Working Papers

Roth, R. (2010). Marx on technical change in the critical edition. The European Journal of the History of Economic Thought, 17(5), 1223-1251.

Schmalensee, R., &Willig, R. D. (1989). Handbook of industrial organization. Amsterdam: North-Holland.

Solow, R. M. (1957). Technical Change and the Aggregate Production Function. The Review of Economics and Statistics, 39(3), 312.

Tossy, T. (2017). Information technology integration for socio-economic development. Hershey, PA: IGI Global.

Vivero, R. L. (2002). The impact of process innovations on firms' productivity growth: the case of Spain. Applied Economics, 34(8), 1007-1016.

Wakelin, K. (2001). Productivity growth and R&D expenditure in UK manufacturing firms. Research Policy, 30(7), 1079-1090.

APPENDIX

Table A1 Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Log(Y)	10709	13.389	2.371	6.907	25.964
Log(K)	6460	12.976	2.432	4.628	27.121
Log(L)	13809	3.424	1.563	0	9.305
R&D USD	4473	191,490.8	4,431,566	0	2.00E+08
RD1	12325	0.450	0.498	0	1
RD2	8105	0.290	0.454	0	1
RD3	8105	0.704	0.456	0	1
Corruption	13361	0.320	0.467	0	1
Finance	13470	0.321	0.467	0	1
Licence	13572	0.024	0.154	0	1
Competition	12235	0.271	0.444	0	1
Export	13838	14.182	29.146	0	100
K/L ratio	6443	4.729	2.893	1.376	21.291
Utilizatin	11415	76.755	25.047	0	100
Skilled	11387	0.519	0.310	0	100

Figure A1 Production function with labour fixed

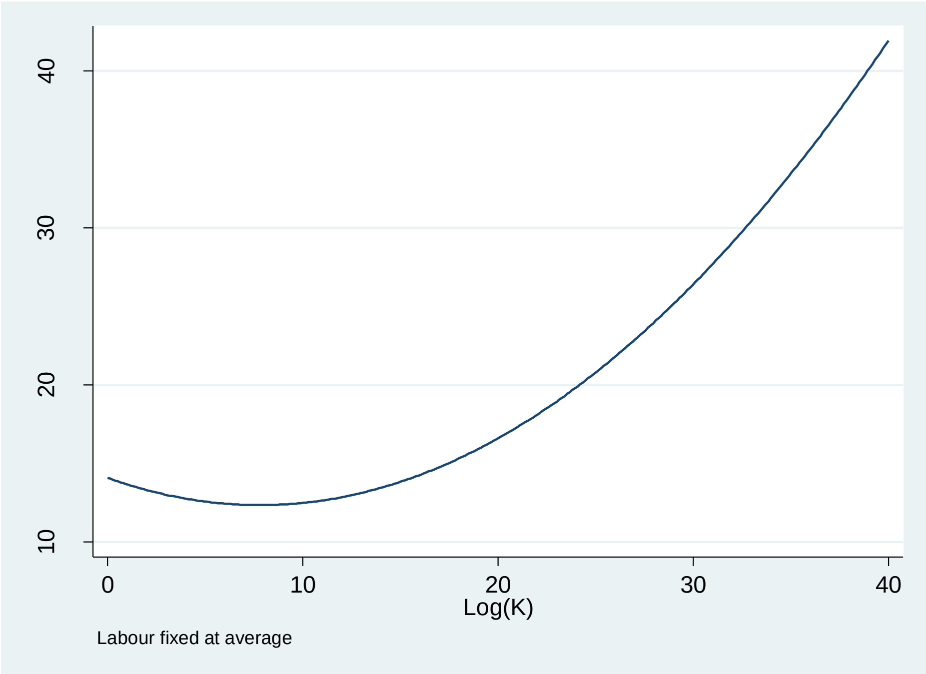


Figure A2 Production function with capital fixed

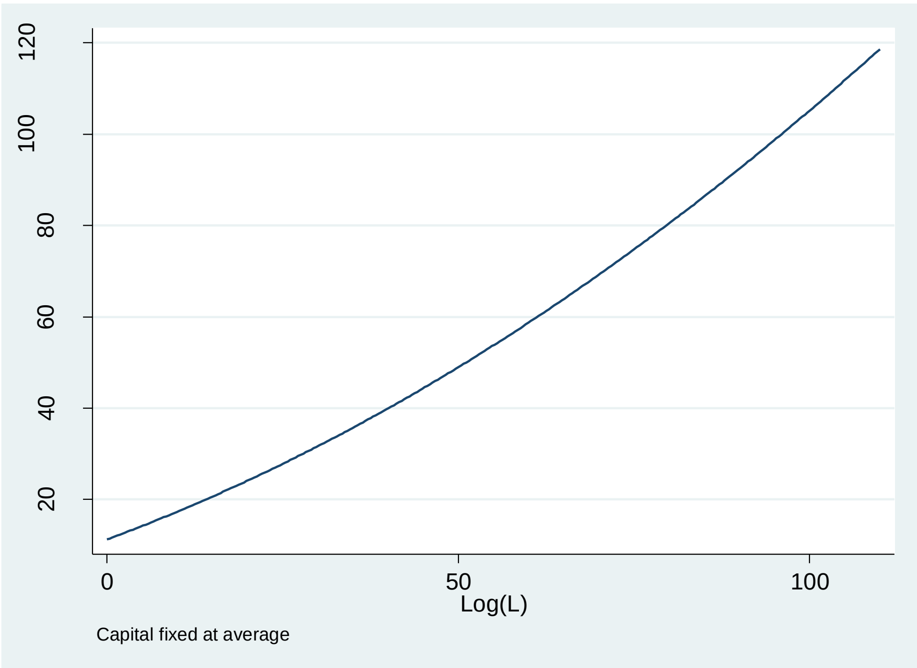


Table A2. Final Stage II estimation results, full output

Log(TFP)	Model F: RD1	Model F: RD2	Model F: RD3
RD USD	2.9e-08***	2.9e-08***	2.9e-08***
RDi, i=1,3	-0.0308	.1571**	0.0094
Corruption	-.1011**	-.0985**	-.0997**
Finance	0.0659	0.0669	0.0679
Licence	.4887*	.4898*	.4879*
Competition	0.0245	0.0257	0.0244
K/L Ratio	-.045***	-.0434***	-.0448***
Export	-8.40E-06	-5.50E-05	-0.000033
Utilization	.0017*	.0017*	.0017*
Skilled	-.1686**	-.1641**	-.1681**
Size			
Medium	.1165*	.0957*	.1128*
Large	.3373***	.2991***	.3329***
Year			
2004	.4391***	.5342***	.4401***
2008	1.085***	1.066***	1.076***
Country			
Belarus	0.2502	0.2085	0.2422
Georgia	-.4593***	-.438***	-.4601***
Tajikistan	-1.375***	-1.386***	-1.377***
Ukraine	-.5651***	-.5746***	-.5705***
Poland	.3179**	.3061**	.3185**
Romania	-.2623*	-.2693*	-.2596*
Kazakhstan	-.8411***	-.8413***	-.8413***
Moldova	-.7123***	-.718***	-.7139***
Bosnia	-0.1223	-0.1304	-0.1254
Azerbaijan	-.8428***	-.8599***	-.842***
FYROM	-0.0586	-0.0672	-0.0596
Armenia	-.7613***	-.7395***	-.7614***
Kyrgyz	-1.29***	-1.3***	-1.289***
Latvia	0.1676	0.1751	0.1623
Lithuania	0.0509	0.0416	0.0497
Slovakia	.6719***	.6486***	.67***
Bulgaria	-0.0134	-0.0227	-0.0156
Croatia	.5282***	.5121***	.5296***
Montenegro	-0.484	-0.4906	-0.4906
Sector			
Food	.2805***	.2762***	.2803***
Textiles	0.1008	0.1028	0.1011
Garments	-.2887***	-.2774**	-.2851***
Chemicals	.3608*	.3326*	.3616*
Plastics & rubber	.5465***	.5604***	.5487***

Non-metallic mineral products	0.0898	0.0752	0.0908
Basic metals	0.0567	0.0212	0.0618
Fabricate metal products	0.131	0.1338	0.1343
Machinery and equipment	.4305***	.4162***	.4317***
Electronics	0.2656	0.2665	0.2658
Construction	.2859***	.3037***	.2985***
Other services	0.131	.1526*	0.1367
Hotel and restaurants	-0.0519	-0.0205	-0.0448
Transport	0.1567	.1795*	0.164
IT	.7017***	.6982***	.6973***
Cons	-.4052**	-.5461***	-.4266**