WIFO

ÖSTERREICHISCHES INSTITUT
FÜR WIRTSCHAFTSFORSCHUNG

WORKING PAPERS

Innovation, Competition and Productivity

Firm Level Evidence for Eastern Europe and Central Asia

Klaus S. Friesenbichler, Michael Peneder

516/2016

ÖSTERREICHISCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG
AUSTRIAN INSTITUTE OF ECONOMIC RESEARCH

Innovation, Competition and Productivity

Firm Level Evidence for Eastern Europe and Central Asia

Klaus S. Friesenbichler, Michael Peneder

WIFO Working Papers, No. 516

April 2016

Abstract

We investigate the drivers of firm level productivity in catching-up economies by jointly estimating its relationship to innovation and competition using data from the EBRD-WB Business Environment and Enterprise Performance Survey (BEEPS) in Eastern Europe and Central Asia. The findings confirm an inverted-U shaped impact of competition on R&D. Both competition and innovation have a simultaneous positive effect on labour productivity in terms of either sales or value added per employee, as does a high share of university graduates and foreign ownership. Further positive impacts come from firm size, exports, or population density. Innovation and foreign ownership appear to be the strongest drivers of multifactor productivity.

E-mail address: Klaus.Friesenbichler@wifo.ac.at, Michael.Peneder@wifo.ac.at
2016/097/W/0

Medieninhaber (Verleger), Hersteller: Österreichisches Institut für Wirtschaftsforschung • 1030 Wien, Arsenal, Objekt 20 • Tel. (43 1) 798 26 01-0 • Fax (43 1) 798 93 86 • http://www.wifo.ac.at/ • Verlags- und Herstellungsort: Wien

Innovation, competition and productivity: firm level evidence for Eastern Europe and Central Asia

Klaus Friesenbichler,* Michael Peneder**

April 10, 2016

Abstract

We investigate the drivers of firm-level productivity in catching-up economies by jointly estimating its relationship to innovation and competition using data from the EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS) in Eastern Europe and Central Asia. The findings confirm an inverted-U shaped impact of competition on R&D. Both competition and innovation have a simultaneous positive effect on labour productivity in terms of either sales or value added per employee, as does a high share of university graduates and foreign ownership. Further positive impacts come from firm size, exports, or population density. Innovation and foreign ownership appear to be the strongest drivers of multifactor productivity.

JEL Codes: O12, O25, O31, O33, L22

Key Words: innovation, competition, productivity, development, transition economies, simultaneous system

* Austrian Institute of Economic Research, Arsenal Obj. 20, A-1030 Vienna, Austria, Tel.: + 43-1-798 26 01 296; E-mail: Klaus.Friesenbichler@wifo.ac.at
** Corresponding author: Austrian Institute of Economic Research, Arsenal Obj. 20, A-1030 Vienna, Austria, Tel.: + 43-1-798 26 01 480; E-mail: Michael.Peneder@wifo.ac.at

1 Introduction

After a quarter century of systemic change, the countries of Central Eastern Europe as well as Central Asia and the Caucasus have largely completed their transition to a market economy. However, substantial differences between these countries are evident and likely to persist. Currently, in some countries market oriented reform efforts are lacking altogether, and in others they have lost public support in the aftermath of the financial crisis. Hence the European Bank for Reconstruction and Development concluded that many economies seem to be 'stuck in transition' (EBRD 2013). Stalling reform efforts have macroeconomic implications, since they are thought to impede the shift of emerging economies from an investment-based to an innovation-driven growth regime. Competition and regulatory policies may facilitate such a switch (Acemoglu et al., 2006).

The reluctant reform implementation implies a slow-down of the industrial dynamism, which has long been the engine of the catching-up process. Innovation and competition – the bottom-up antithesis to a top-down command economy – are generally regarded as key drivers of productivity growth (Kornai, 1992). But the two also constitute an uneasy pairing, where we still lack a precise understanding of how they affect each other, and how they jointly affect productivity. Theoretical models have demonstrated a variety of mechanisms and possible outcomes which depend, among other factors, on the static vs. dynamic nature of the game, whether R&D expenditures are modeled as fixed or variable costs, the mode of competition, the nature of innovations, or the structure of rewards (De Bondt and Vandekerckhove, 2012; Peneder, 2012).

Due to the limited availability of data, many intricate aspects of these models are typically outside the reach of empirical applications (Cohen, 2010). However, even for the estimation of more robust average relationships there is much contradicting evidence on their sign and shape. This also holds for the presently analysed countries. For instance, Carlin et al. (2004) demonstrated that the presence of a few rivals can be more conducive to firm performance than the presence of many competitors. In contrast, Gorodnichenko et al. (2010) challenged findings of a negative effect of competition on innovation and provided weak support for an inverted-U relationship between innovation and growth. But the ambiguity of results may also reflect different stages of development. For example, if in the early phases of transition new firms focus on less contested markets, but later aim to expand into the larger and more competitive ones, the role of innovation as a driver of competitive advantage also changes (Dutz and Vagliasindi, 2000; Hare and Turley, 2013).

While the bulk of empirical studies on competition and innovation focuses on developed market economies, research on transition economies is scarce. One notable exception has been the EBRD Transition Report 2014 on 'Innovation in Transition'. Using similar data as the current analysis, its findings are especially relevant to our research. Estimating the CDM-model of Crepon et al (1998), it reports a positive impact of innovation on labour productivity. In a separate chapter, descriptive charts picture an inverted-U relationship between competition and the share of innovative firms (EBRD, 2014, p. 35 ff., p. 61).

The data originate from the fifth round of the *Business Environment and Enterprise Performance Survey* (BEEPS V), fielded in 2012 and 2013 by the European Bank for Reconstruction and Development (EBRD) and the World Bank (WB). Our dataset covers Central and Eastern Europe (CEE) as well as Central Asia and Caucasus (CAC).

Starting from the model developed in Peneder and Woerter (2014), we first set-up a simultaneous system of three equations to disentangle the mutual impact of innovation and competition. Similar to the CDM model we apply separate equations to estimate R&D expenditures as an input and innovation as its output. In the CDM-approach causality then goes straight from research to innovation, and from innovation to productivity. Different from our analysis, it does not consider competition as an endogenous determinant of innovation and productivity. In contrast, Peneder and Woerter (2014) accounted for the reverse causality between competition and innovation, but did not address productivity.

Our objective is to advance the agenda by integrating the productivity equation in the simultaneous system of Peneder and Woerter and apply it to the large sample of firms in Eastern Europe and Central Asia. We thus can test the simultaneous impact of competition and innovation on productivity within a comprehensively structured model: First, the innovation *opportunity* function determines the impact of competition on the firm's research expenditures. Second, the innovation *production* function captures the transmission from research expenditures to innovation outcome. Third, the innovation *impact* function shows how innovation outcomes affect the number of competitors perceived by the individual firm. This set of three endogenous equations forms our *basic system*.

The endogenous system is then extended by adding an equation for the squared component of competition and another to determine the extensive margin of R&D (i.e. the probability to conduct any R&D expenditures). In the final step, we add productivity as a strictly dependent variable and estimate it for three alternative indicators. The broadest measure is total sales per employee, which has the advantage of preserving the initial sample of observations. Sample size substantially decreases for the alternative measures of value added per employee and even more so for multi factor productivity (MFP).

We apply three-stage least-square estimations (3SLS) with sectoral taxonomies of technological regimes as the main exclusion restrictions to identify the system. The taxonomies account for the repeated concern about the relationship between innovation and competition being dominated by the specific technological and market environment (Cohen 2010; Gilbert 2006).

Our main findings are as follows:

- First, the estimations confirm an inverted-U shaped impact of competition on R&D. At low levels of initial rivalry, more competitors incite firms to do (more) R&D, but at a diminishing rate. Hence, the largest incentives for R&D appear at intermediate levels of competition. Among the exogenous variables, a higher share of university graduates, direct and indirect exports raise both margins of R&D. A larger firm size increases the extensive margin, while age tends to decrease the intensive margin.

– Second, spending more on R&D consistently increases the probability to innovate successfully. So does a higher share of university graduates, direct and indirect exports, larger firm size and foreign ownership for most specifications.

– Third, successful innovations consistently reduce the perceived number of competitors. Conversely, a higher share of university graduates, firm size, direct and indirect exports, population density and better regulatory quality tend to raise their number. Evidence for the impact of appropriability conditions is mixed and calls for further study.

– Fourth, competition and innovation have a simultaneous yet independent positive impact on productivity for sales and value added per employee. We can only see their separate effects, when controlling for their endogeneity in the simultaneous system. In the single regressions with no account of their mutual causality, the competition variable becomes insignificant.

– Fifth, among the exogenous variables education, foreign ownership and own exports are the most robust drivers of firm-level productivity. Impeding the diffusion of new technologies, better appropriability conditions *ceteris paribus* have a negative impact on total sales and value added per employee.

The basic endogenous system is robust also for smaller samples in separate estimations for the CEE and CAC regions. In contrast, the extended and full system, while theoretically preferable, show to be more sensitive to sample size. But once we control for the endogeneity between competition and innovation and use the total sample of firms, the alternative specifications make little difference for the final productivity estimates.

The remainder of this paper is organized as follows. In Section 2 we explain the theoretical rationales and develop a stylized structural model. In Section 3 we discuss the data and explain the variables. Section 4 explores the different specifications for estimation. Section 5 presents the empirical findings. In Section 6 we summarize and conclude.

2 Theoretical framework

2.1 The inverted-U effect

The recent surge of interest in the impact of competition on innovation feeds on the growing popularity of the *inverted-U* hypothesis by Aghion et al. (2005). The inverted-U implies that neither perfect competition nor a monopoly can provide the optimal market environment, and that instead some intermediate degree of rivalry is most conducive to innovation. Aghion et al. (2005) distinguish between the firms' pre- and post-innovation rents and relate them to their relative proximity to the technological frontier. The 'rent dissipation effect' involves a negative impact of competition on post-innovation rents, since competition is expected to be high, even if the firm successfully innovates. A positive 'escape competition effect' occurs if competition reduces pre-innovation rents more than post-innovation rents, thereby raising the incremental returns to innovation. Their key prediction is that the positive 'escape effect' dominates at low levels of competition, while the negative 'dissipation effect' is stronger at high levels of competition. The trade-off depends on the technological characteristics of the industry, in particular the technological distance between firms.

The framework of Aghion et al. (2005) and other game theoretical models cannot easily be transposed to the cross-sectional micro-econometric setting of our analysis. The majority of the firms in our sample do not operate within an environment that is consistent with the specific assumptions applied therein. Even if they did, we do not have the data to verify them and discriminate our observations accordingly.[1] The many duopoly models clearly do not apply, since the vast majority of our firms has more than one rival. It is even hard to justify applying predictions from more general oligopolistic models, as on average and across countries about 44 percent of the firms in our sample report having more than 25 competitors and another 27 percent report having between 6 and 25 competitors for their principal product (Table A1 in the Annex). Moreover, many of the game theoretical models specifically refer to process innovations, whereas the majority of innovating firms in our sample report having introduced novel products. It is precisely from game theory that we have learned just how sensitive predictions are with respect to these assumptions.

Because of its straightforward intuition and good match of variables with the available data, the older decision-theoretical model of Kamien and Schwartz (1976) provides another appealing analytic setting for our case. In a nutshell, they modeled an innovation race, where competition enters the decision problem of a firm i in the form of a subjective belief about the exogenous (and positive) hazard h, which is the probability of preempting innovations by a rival. Without additional knowledge about the innovation strategies and capabilities of competitors, firms assign equal probabilities of innovation to each of these and the constant $1/h$ depicts the expected introduction time of a rival innovation. Within this information setting, the hazard h directly relates to the perceived number of competitors C_i. Maximizing the expected net return of R&D, greater rivalry raises the risk of preemption. Up to a certain degree, the competitive threat spurs own R&D. But when competition gets increasingly intense and the cost of defense rises, lower returns from cheaper imitation become more attractive than risky returns from own innovation. As a consequence, firms are more cautious and invest less in R&D when competition is too intense. In this regard, the different models of Aghion et al (2005) and Kamien and Schwarz (1976) are observationally equivalent and form a solid basis for our empirical investigation.

[1] While game-theoretical models apply to very specific markets with a few well-defined competitors, we should realistically assume that most firms in our sample have only limited knowledge of the precise information set and intricate strategic aspects of their rivals' choices.

2.2 Causal structure

We are interested in the impact of competition and innovation on productivity. But competition and innovation are highly endogenous, simultaneously affecting each other with causality going both ways.[2] In order to test and identify independent effects of competition and innovation on firm productivity, we must first understand the underlying causal structure and aim to solve the endogenous relationships simultaneously.

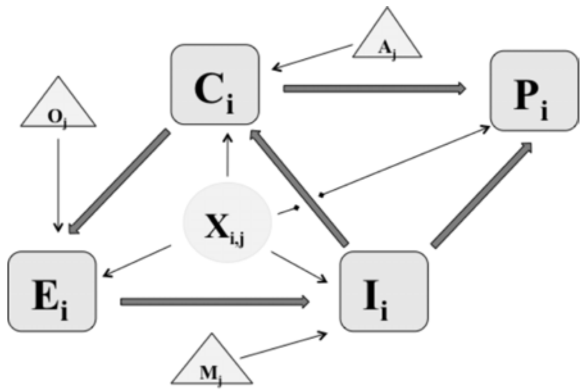
Similar to Crepon et al (1998) we distinguish between research expenditures and innovation outcome. We thus start with three endogenous variables that are connected in a system of circular causation by three equations: First we formulate an innovation *opportunity* function to test how the number of competitors in the principal product C_i affects the firm's expenditures on R&D E_i. Second, an innovation *impact* function addresses the reverse causation from innovation outcome I_i to the number of competitors C_i. To close the system, we add an innovation *production* function, which relates research expenditures E_i to innovation outcome I_i. This forms the basic causal structure of the endogenous system, to which we later add the productivity equation (Figure 1).

Since the endogenous variables are simultaneously determined by different equations, one cannot directly identify the true parameters for each equation from its observed values. Instead one must introduce exclusion restrictions Z_i that are correlated with the endogenous variable but not with the error term. Creating counterfactual observations understood to be exogenous, these shift one equation in order to trace out the other. There is, however, more than a technical side to it. We mainly use sectoral proxies of so called 'technological regimes'. Their choice is motivated by the repeated complaint about the failure to account for different market and technological conditions in past studies of the relationship between innovation and competition.[3] Our variables about *opportunity* conditions (O_i), *appropriability* (A_i), and the *cumulativeness* of knowledge (M_i) thus reflect fundamental theoretical notions of innovation research.[4]

R&D expenditures

The innovation *opportunity* function specifies for firm i how competition affects research expenditures E_i and estimates the impact of the number of competitors C_i together with a vector of control variables X_i. Adding a nonlinear term C_i^2, we can test for an inverted-U relationship. The underlying rationale is that competition affects R&D incentives via the firm's changing beliefs about the probability of a rival introduction of an innovation. The model of Kamien and Schwartz (1976) directly relates that to the perceived number of competitors C_i. Consistent with the prediction of an inverted-U relationship, we expect a positive sign for β_1, and θ to be negative.

[2] See, for example Sutton (1991, 1998) or Martin (2012).
[3] For example, Gilbert (2006) and Cohen (2010) repeatedly stress this point.
[4] See Winter (1984), Malerba and Orsenigo (1993), and Breschi et al (2000).



NB: C_i = perceived number of competitors; E_i = research expenditures; I_i = innovation outcome; P_i = productivity; X_{ij} = control variables; O_j = opportunity conditions; A_j = appropriability conditions; M_j = cumulativeness of knowledge.

Figure 1: Basic structure of the model

$$E_i = \alpha_1 + \beta_1 C_i + \theta C_i^2 + \delta_1 X_i + \gamma_1 O_j + \upsilon_1 \quad (1)$$

A sectoral taxonomy of *opportunity conditions* O_j provides for the exclusion restriction. It was created from micro-data of the European Union's *Community Innovation Survey* (EU-CIS) and accounts for exogenous sectoral contingencies of R&D investments. It affects the probability of the individual firms to invest in own R&D, whereas its impact on the innovation outcome is only indirect, i.e. due to the variations in R&D expenditures. By *assumption* and consistent with the causal structure of the model, the variable is therefore not correlated with the error term in the innovation production function. The same applies to the R&D dummy variable, which in the basic specification controls for the initial selection into R&D performing and non-performing firms. Among further extensions, we will later introduce a separate equation to endogenize this extensive margin of R&D.

Innovation outcome

The innovation *production* function relates the innovation outcome I_i to the firm's R&D expenditures E_i and a vector of control variables X_i. In the Kamien - Schwartz model this corresponds to the assumption that more expenditures on R&D buy a sooner completion date and hence raise the probability to win the innovation race. That is, more R&D raises the probability of innovation success. Hence we expect the coefficient for R&D expenditures to be positive. We need this equation to close the system. In addition, the estimates will tell us about the impact of exogenous variables such as firm size, exports, age or foreign ownership on innovation success, conditional on the jointly determined level of R&D expenditures.

$$I_i = \alpha_2 + \beta_2 E_i + \delta_2 X_i + \gamma_2 M_j + \upsilon_2 \tag{2}$$

For the exclusion restriction, we apply a sectoral taxonomy depicting the *cumulativeness of knowledge* M_j, which is again derived from the micro-data of the EU-CIS. For a given status as R&D performer, we expect that increasing returns to knowledge creation have an impact on the probability of innovation success. Conversely, the impact on the intensity of competition can only be indirect depending on whether the innovation is indeed successful. As a consequence, the cumulativeness of knowledge at the sector level is not correlated with the error term in the innovation impact function. Similarly, the influence of increasing returns in knowledge creation on the R&D incentives is only indirect and depends on their impact on the probability of innovation success.

Competition

The innovation *impact* function captures the effect of the endogenous innovation outcome I_i and a vector of exogenous control variables X_i on the number of competitors C_i. It accounts for a quintessential assumption in many models of industrial organization or Schumpeterian endogenous growth:[5] Firms invest in R&D in order to earn a positive rent from market power, and hence to capture markets where competition is less intense and margins are higher.

$$C_i = \alpha_3 + \beta_3 I_i + \delta_3 X_i + \gamma_3 A_j + \upsilon_3 \tag{3}$$

As in the Kamien - Schwartz model, we assume that the rents from innovation depend on exogenous characteristics of markets and technology and can either be fully appropriated by the innovator, or diffuse among a mixed ecology of innovation leaders and followers, who imitate and on average earn lower returns. In an ordinal ranking of firms that either do not apply new technologies, adopt them from external sources, or innovate on their own, we expect that a higher degree of innovativeness decreases the number of competitors.

The main exclusion restriction applies with regard to a sectoral taxonomy of *appropriability* conditions A_j, also derived from the EU-CIS. We consider that the characteristic differences in the distribution of appropriability measures at the industry level reflect exogenous sectoral contingencies, which correlate with the perceived number of competitors by the individual firms. Furthermore, our model implies that the sectoral appropriability conditions affect innovation incentives only indirectly, that is if they have an influence on the intensity of competition. Consequently, they are uncorrelated with the error term in the innovation opportunity function. The same applies to population density and regulatory quality, which we assume to have a positive impact on competition, but exclude from the estimation of equation 1.

[5] See Aghion and Howitt (1992, 2009).

Reduced form

The above structural model depicts the economic intuition and rationale for our endogenous system. For the sake of completeness, one can also solve it in reduced form. To do so, first substitute equation (1) in equation (2):

$$I_i = \alpha_2 + \beta_2(\alpha_1 + \beta_1 C_i + \theta C_i^2 + \delta_1 X_i + \gamma_1 O_j + \upsilon_1) + \delta_2 X_i + \gamma_2 M_j + \upsilon_2$$

Next, substitute this expression in equation (3):

$$C_i = \alpha_3 + \beta_3[\alpha_2 + \beta_2(\alpha_1 + \beta_1 C_i + \theta C_i^2 + \delta_1 X_i + \gamma_1 O_j + \upsilon_1) + \delta_2 X_i + \gamma_2 M_j + \upsilon_2] + \delta_3 X_i + \gamma_3 A_j + \upsilon_3$$

Multiplication of terms yields:

$$\begin{aligned} C_i &= \alpha_3 + \alpha_2\beta_3 + \alpha_1\beta_2\beta_3 + \beta_1\beta_2\beta_3 C_i + \beta_2\beta_3\theta C_i^2 + \beta_2\beta_3\delta_1 X_i + \beta_2\beta_3\gamma_1 O_j \\ &\quad + \beta_2\beta_3\upsilon_1 + \beta_3\delta_2 X_i + \beta_3\gamma_2 M_j + \beta_3\upsilon_2 + \delta_3 X_i + \gamma_3 A_j + \upsilon_3 \end{aligned}$$

From this we can derive the following quadratic function:

$$\begin{aligned} \beta_2\beta_3\theta C_i^2 &+ (\beta_1\beta_2\beta_3 - 1)C_i + \beta_2\beta_3\delta_1 X_i + \beta_3\delta_2 X_i + \delta_3 X_i + \beta_2\beta_3\gamma_1 O_j \\ &+ \beta_3\gamma_2 M_j + \gamma_3 A_j + \alpha_1\beta_2\beta_3 + \alpha_2\beta_3 + \alpha_3 + \beta_2\beta_3\upsilon_1 + \beta_3\upsilon_2 + \upsilon_3 = 0 \end{aligned}$$

To simplify the following expressions, we substitute

$$a = \beta_2\beta_3\theta$$

$$b = (\beta_1\beta_2\beta_3 - 1)$$

Provided that a solution exists, $b^2 - 4ac \geq 0$ and the general formula of the roots of the quadratic polynomial $aC^2 + bC + c$ gives us the reduced form of the competition variable:

$$C_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (4)$$

In the next step, we substitute equation 2 into equation 3:

$$C_i = \alpha_3 + \beta_3(\alpha_2 + \beta_2 E_i + \delta_2 X_i + \gamma_2 M_j + \upsilon_2) + \delta_3 X_i + \gamma_3 A_j + \upsilon_3$$

$$C_i = \beta_2\beta_3 E_i + \alpha_3 + \alpha_2\beta_3 + \beta_3\delta_2 X_i + \delta_3 X_i + \beta_3\gamma_2 M_j + \gamma_3 A_j + \beta_3\upsilon_2 + \upsilon_3$$

or

$$\beta_2\beta_3 E_i = C_i - [\alpha_3 + \alpha_2\beta_3 + (\beta_3\delta_2 + \delta_3)X_i + \beta_3\gamma_2 M_j + \gamma_3 A_j + \beta_3\upsilon_2 + \upsilon_3]$$

Note that the coefficients β_2 and β_3 capture the endogenous effects of R&D on innovation, and of innovation on the number of competitors, respectively. The expression on the right hand side corresponds to the number of competitors C_i minus the intercept, which we substitute by the variable d:

$$d = \alpha_3 + \alpha_2\beta_3 + (\beta_3\delta_2 + \delta_3)X_i + \beta_3\gamma_2 M_j + \gamma_3 A_j + \beta_3\upsilon_2 + \upsilon_3$$

The reduced forms of research expenditures E and innovation impact I are then given by

$$E_i = \frac{\frac{-b\pm\sqrt{b^2-4ac}}{2a} - d}{\beta_2\beta_3} \tag{5}$$

$$I_i = \alpha_2 + \frac{\frac{-b\pm\sqrt{b^2-4ac}}{2a} - d}{\beta_3} + \delta_2 X_i + \gamma_2 M_j + \upsilon_2 \tag{6}$$

Productivity

Once we have established the endogenous system of competition, research expenditures and innovation, we are interested in their joint impact on the productivity performance P_i of the firms in our sample. We thereby assume that both competition C_i and innovation I_i affect productivity directly, whereas R&D expenditures E_i affect productivity only indirectly, that is if innovations I_i are successful.

$$P_i = \alpha_4 + \beta_4 C_i + \varphi I_i + \delta_4 X_i + \upsilon_4 \tag{7}$$

Among the independent variables, the vector of control variables X_i includes for example education, firm size, direct or indirect exports, foreign- and state ownership as well as population density. In addition, we control for the general income level in terms of the country's GDP per capita.

Note that different from the other equations, firm productivity is treated as a purely dependent outcome variable. Avoiding to add another endogenous equation and thereby overburden the capacity of the system, we assume that there are no reverse causal effects of productivity on its determinants, and any apparent interrelationships are captured through the mutual impacts between its endogenous drivers.

3 Data and variables

3.1 Data

The main data source of our analysis is the *Business Environment and Enterprise Performance Survey* (BEEPS V), which was jointly financed by the European Bank for Reconstruction and Development (EBRD) and the World Bank (WB). BEEPS provides internationally comparable, establishment-level quantitative information obtained from face-to-face interviews. The implemented questionnaire covers a broad range of variables on the particular business environment, including innovation activities, the degree of competition, and other determinants of firm-level productivity.

We use the surveys conducted in Central and Eastern Europe (CEE) as well as Central Asia and Caucasus (CAC). These transition economies have gradually increased their technological capabilities, and to varying degrees introduced pro-competitive policy measures that effectively led to less concentrated market structures and higher levels of competition (Friesenbichler et al., 2014).

The data cover the survey years 2012 for Russia and 2013 for the other countries. Monetary values are mostly for 2010 or 2011 as the last complete fiscal year[6] and were converted from local currency units into USD, using exchange rates from the World Bank Indicators (official exchange rate; local currency unit per US$, period average), which does not include Euro countries. Eurostat exchange rate data were used to calculate the USD values for establishments in Euro countries. The countries included are

– *Central and Eastern Europe* (CEE): Albania (359 observations), Belarus (360), Bosnia and Herzegovina (360), Bulgaria (293), Croatia (360), Estonia (273), Hungary (310), Kosovo (202), Latvia (336), Lithuania (270), Macedonia (360), Moldova (360), Montenegro (150), Poland (542), Romania (540), Russia (4,220), Serbia (359), Slovakia (268), Slovenia (270), and Ukraine (1,002).

– *Central Asia and Caucasus* (CAC): Armenia (360), Azerbaijan (390), Georgia (360); Kazakhstan (600), Kyrgyzstan (270), Tajikistan (359).

3.2 Dependent variables

The research and innovation indicators are in line with established definitions of private sector R&D expenditures for developing countries (OECD 2012; UNESCO 2010). The term 'firm' will be used instead of the less common 'establishment' to describe the surveyed unit. The following variables entered the analysis (code of questionnaire in parentheses):

– *Research.* Total expenditures are the sum of intramural and external R&D (ECAo17; ECAo19), divided by the annual sales (d2). For consistency with the other ordinal endogenous variables we assigned a value of one to firms that have no R&D-expenditures, two to firms with an R&D to sales ratio below 1.5 percent, three to firms with a ratio between 1.5 and 5 percent, and four to those above 5 percent. Since the first category of zero R&D comprises the mass of observations, we also created a dummy variable of whether the firms have any R&D expenditures or not. We interpret it as the *extensive margin* of R&D. When included in the estimation of the ordered categorical variable, we interpret the latter as its *intensive margin*.

[6]For more details see http://ebrd-beeps.com.

- *Innovation.* The variable classifies firms in terms of innovation outcome on an ordinal scale. Two adaptive and two creative types have been defined. First, adaptive firms (type I) are assigned the value one if they neither introduced a new product (h1), nor a new process (h3). We assign the value two to firms that introduced either kind of innovation, or both, if these were new to the firm but not new to the market (adaptive type II). Next we assign the value of three to creative firms (type I) that introduced innovations that were new to the market, but with a major contribution from external partners (ECAo5). A score of four was assigned to firms that launched an own innovation that was new to the market (h2; creative type II).

- *Competition.* Another ordinal variable was constructed from the question about the number of competitors for the principal product/service in the main market (e2b). The variable takes the value one if a firm is a monopoly, two if firms have reported one to five competitors, three for values between six and 25 competitors, and four if they perceived to have 26 or more competitors. The question also allowed for the option 'too many to count', to which also a four was assigned.

- *Productivity.* For most of the analysis we use total sales per employee, which is the crudest but best available measure from the surveys. To test for the robustness of our findings, we alter the productivity indicators to value added per employee and a proxy of multi factor productivity. However, the construction of these variables requires additional accounting data, which are not available for all firms and substantially reduce the sample.

  a. *Total sales per employee* is given by the establishment's total annual sales (d2) divided by the number of full time equivalents.[7] Despite apparent shortcomings this measure is frequently used in the literature. For our purposes it is the only measure, which provides us with a sufficient number of observations to comprehensively test our system of simultaneous equations. We thereby have to assume that the ample use of industry dummies and sectoral taxonomies eliminates systematic distortions between sectors.

  b. *Value added per employee* is defined as total annual sales (d2) minus intermediate inputs (n2e) and divided by the absolute number of full time equivalents. Intermediate inputs were calculated as the sum of total annual cost of raw materials and intermediate goods used in production, electricity and fuel. This variable is only available for manufacturing firms, since other firms were not asked about the replacement values of machinery and equipment and the total annual cost of raw materials or intermediate goods.

[7]Full time equivalents are defined as permanent workers (l1) plus temporary workers (l6) weighted by their average duration of employment (l8).

c. *Multi factor productivity* (MFP) is estimated by regressing value added on the replacement value of machinery and equipment, full time equivalents of labour inputs, industry dummies and a series of interactions of the industry dummies with the capital and the labour stock.[8] The estimations are conducted for each country separately. The interactions capture sector-specific effects and the time effects control for economy-wide productivity shocks. MFP is then defined as the residual plus the intercept. For the same reasons explained above, this variable is only available for manufacturing firms.

3.3 Independent variables

Technological Regimes

We use three taxonomies of so called *technological regimes* as our main exclusion restrictions. Peneder (2010) identified them at the level of NACE 2-digit industries.[9] The taxonomies summarize characteristic sectoral profiles in the distribution of individual firms with different innovation activities. They discard with the idea of indicating the behavior of a 'representative' firm, but instead point at the variety of micro behavior and their characteristic distribution within firm populations. Different from traditional approaches, firm-level variety is treated as a natural element of, and not an antagonism to sectoral classifications. In other words, innovative firms within so called low-tech industries are as much part of their characteristic distribution as are non-innovating firms within so called high-tech sectors.

Since the taxonomies capture particular technological characteristics at the industry level, we can consider them strictly exogenous to the choices of individual firms (other than entry, which is not covered by our data). Also the fact that they originate from a different data source, well documented and published before the current analysis was undertaken, enhances its credibility relative to exclusionary restrictions created *ad hoc* from the same survey data. One critical feature of any taxonomy is its persistence over time and between countries. One must expect that the propensity to invest in R&D or to apply patents changes in the process of development. Nevertheless, their relative importance in one industry as compared to others tends to be very persistent, not least because of the broad boundaries drawn within such classifications.[10]

[8] See Saliola and Seker (2011); Syverson (2011).

[9] The classifications were built from the micro-data of the EU-CIS, covering 78,000 individual firms from 22 European countries, among them the following CEE member states of the EU: Estonia, Latvia, Lithuania, Czech Republic, Slovakia, Hungary, Slovenia, Romania and Bulgaria. In a first step, firms were identified according to essential characteristics of technological regimes. In the second step, a combination of hierarchical and non-hierarchical statistical cluster methods was applied to classify industries by the distribution of the firm types previously identified.

[10] See Peneder (2010) for a detailed discussion of the theoretic rationales and statistical methods together with a full documentation and empirical validation of the resulting classifications.

- *Opportunity conditions.* The taxonomy captures the characteristic distribution of firms in terms of perceived opportunities for innovation as revealed by their intramural and external expenditures on R&D. The variable comprises four ordinal groups. We assume that it directly affects the probability distribution of the firms' R&D activities in the first innovation opportunity function, whereas its impact on innovation outcomes and competition are only indirect, which means a consequence of their R&D expenditures. It is therefore not included in the innovation production and the innovation impact functions.

- *Cumulativeness of knowledge.* It reflects the characteristic distribution of firms regarding increasing returns to knowledge creation within three ordinal categories. They were indirectly identified by distinguishing firms according to the relative importance of internal vs. external sources of knowledge and their role as either technological leaders with own innovations or followers.[11] Relating the sources of information with innovation success, this variable should directly affect the innovation production function. Again, this provides for a valid exclusion restriction, which affects competition or R&D incentives only indirectly, that is via the jointly determined impact on innovation outcomes.

- *Appropriability conditions.* The taxonomy captures a sector's characteristic distribution of firms regarding appropriability measures to protect the rents from innovation. It comprises five groups ordered by the relative importance of the different appropriability measures. Apart from formal intellectual property rights, the EU-CIS also contained information about the use of strategic measures, such as secrecy, lead-time, or complex design, which was therefore included. The direct impact should be on the intensity of competition, which in turn may indirectly affect the incentives and outcomes of innovation. This makes it a valid exclusionary restriction for the innovation impact function.

Other exogenous variables

We consistently use country and industry dummies for all equations. For the latter firms are assigned to 15 manufacturing and 9 service industries at the ISIC Rev. 3.1 two-digit level. In addition, the following variables originate again from the EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS V):

[11] If an innovative firm reported internal sources of knowledge to be more or as important as external sources, it was considered to operate under a regime of high cumulativeness. For firms adopting the technology of others, the rule was reversed, so that their knowledge environment is considered to be cumulative, if they report that internal sources of information are less important than external ones. Conversely, cumulativeness is low, if an innovative firm sources more information from external than from internal sources, or if an adopting firm reports that internal sources are more or as important than external sources.

- *Firm size.* We use full time equivalents to create dummy variables for the same size classes, which the sampling process considered.
- *Education (univ).* The fraction of the persons employed full time and who hold a university degree (ECAq69).
- *Exports.* Dichotomous variables were set for firms that export *indirectly* in the sense of selling domestically to a third party that exports (d3b), or export *directly* (d3c).
- *Age.* The difference between the survey year and the year in which the establishment began operations (b5).
- *Foreign ownership.* This dummy variable takes on the value one if at least one percent of the firm is owned by a private foreign individual, company or organisation (b2b). The cut-off reports some foreign ownership, not necessarily a controlling share (for which the number of affected observations would become much smaller). Alternative higher cut-offs were used for tests of robustness.
- *State ownership.* This dummy variable takes on the value one, if at least one percent of the firm is owned by the government or the state. The cut-off reflects the existence of some sort of state influence, not necessarily a controlling share (for which the number of affected observations would become much smaller). Alternative higher cut-offs were used for tests of robustness.
- *Political instability.* Survey response whether political instability is an obstacle to the current operations of the establishment (j30e), measured on a five part ordinal scale. It is introduced as an additional exclusionary restriction in an extended specification.

The following variables represent aggregate data for individual countries:

- *GDP pc.* Gross domestic product per capita in constant prices of 2005 from the World Development Indicators database provided by the World Bank.
- *Population density.* People per square kilometer of land area used taken from the World Development Indicators database provided by the World Bank.
- *Regulatory quality.* Extracted from the World Governance Indicators (Kaufmann et al, 2010) provided by the World Bank. The index captures data and perceptions of the ability of governments to implement sound policies and regulations and the extent of regulatory burden, with a focus, e.g., on the quality of legal systems and administration, competition and anti-trust policies, labour market policies, the tax system, or trade policy. The index is rescaled to values that range from zero to five.

4 The simultaneous system

In the presence of reverse causality, the endogenous explanatory variables are correlated with the error term. We therefore need an instrumental variable (IV) approach to ascertain that the first-stage residuals are uncorrelated with fitted values and covariates, thus avoiding biased and inconsistent estimates. Rather than specifying a distribution of the errors for ordered choice models that admit the prediction of the ordered endogenous outcomes (but for which the IV estimation is not applicable), we seek a linear probability model (LPM) to identify the linear approximation of causal effects (Angrist and Pischke, 2009).

The estimated coefficients depict the probability of a unit change of the dependent variable in response to a ceteris paribus unit increase of the explanatory variable. The linear approximation assumes a constant effect, which does not vary with the initial value of the dependent variable. While this calls for caution in the case of extreme values, one can nevertheless expect it to provide good estimates near the center of the distribution, i.e. close to the average values of the covariates (Wooldridge, 2010, p. 455; Menard, 2001). The case for the LPM is even stronger, if – as in our setting – most of the explanatory variables are also discrete and take on similarly few categories (ibid., p. 546). This mitigates the problem of potential predictions outside the range of admissible values.

We apply three-stage least-squares estimation (3SLS). In the first stage, the reduced form is estimated. In the second stage, the fitted values of the endogenous variables are used to estimate all the equations in the system. In the third stage, the residuals are used to estimate the cross-equation variances and co-variances, and generalized least-squares parameter estimates are obtained. Accounting for the cross-equation correlations, the 3SLS procedure yields more efficient parameter estimates than the 2SLS (Madansky, 1964).

Keeping an eye on the trade-off between efficiency and robustness of the endogenous system, we explore three different specifications to which we later add the productivity equation. The basic system uses the minimum number of equations needed to represent the causal structure of equations 1 to 3 in a rudimentary form. It is reproduced by the vector equation (8). On the left hand side we find the dependent variables for each of the n equations. On the right hand side, the first vector gives the constant intercept α, the second vector the endogenous variables with coefficients β, the third vector the exclusion restrictions with coefficients γ, the fourth vector the k common exogenous control variables, depicted by the coefficients δ^k, and finally the error terms υ. The subscripts denote whether the data refer to the level of individual firms i, industry j, or countries l.

Consistent with Crepon et al (1998)[12] we add an R&D dummy E_i^d in equation 1, which accounts for the fact that most firms have no R&D expenditures to report. The coefficient β_2 in the second equation should therefore be interpreted as an intensive margin of R&D, while assuming that the extensive margin is exogenous. This assumption will be relaxed in a later specification of the system.

[12] Note that they need the R&D dummy only when estimating the full sample with non-innovating firms.

Basic system:

$$\begin{pmatrix} E_i \\ I_i \\ C_i \\ P_i \end{pmatrix} = \begin{pmatrix} \alpha_1 + & \beta_1 C_i + \theta C_i^2 & +\gamma_1 E_i^d + \gamma_2 O_j & +\delta_1^k X_i^k & +\upsilon_1 \\ \alpha_2 + & \beta_2 E_i & +\gamma_3 M_j + \gamma_4 D_l & +\delta_2^k X_i^k & +\upsilon_2 \\ \alpha_3 + & \beta_3 I_i & +\gamma_5 A_j + \gamma_6 D_l + \gamma_7 R_l & +\delta_3^k X_i^k & +\upsilon_3 \\ \alpha_4 + & \beta_4 C_i + \beta_5 I_i & +\gamma_8 A_j + \gamma_9 D_l + \gamma_{10} G_l & +\delta_4^k X_i^k & +\upsilon_4 \end{pmatrix} \tag{8}$$

Among the other exogenous variables, population density D_l is an additional exclusion restriction. It is not included in equation 1, since there is no a priori reason to suspect an impact on the incentives for R&D. Similarly, regulatory quality R_l only appears in equation 3 with a presumed impact on the intensity of competition. In the final productivity equation GDP per capita G_l is added among the control variables. Competition and innovation outcome appear as explanatory variables. In contrast, R&D is not included since we expect it to affect productivity only indirectly, that is when it leads to successful innovations.

One problem with the basic model is that the first equation contains the nonlinear endogenous variable C_i^2 and thereby violates the linearity assumptions of the IV estimation. We therefore extend the model by adding a separate equation to explain C_i^2 endogenously. In equation 1 the coefficient accordingly changes from θ to β_2. We use the squared appropriability condition A_j^2 as exclusionary restriction first. Tests of the quality of the instruments led us to add another variable capturing the perception of political instability N_i as an obstacle to current operations in both equations for C_i and C_i^2.[13]

Extended system:

$$\begin{pmatrix} E_i \\ I_i \\ C_i \\ C_i^2 \\ P_i \end{pmatrix} = \begin{pmatrix} \alpha_1 + & \beta_1 C_i + \beta_2 C_i^2 & +\gamma_1 E_i^d + \gamma_2 O_j & +\delta_1^k X_i^k & +\upsilon_1 \\ \alpha_2 + & \beta_3 E_i & +\gamma_3 M_j + \gamma_4 D_l & +\delta_2^k X_i^k & +\upsilon_2 \\ \alpha_3 + & \beta_4 I_i & +\gamma_5 N_i + \gamma_6 A_j + \gamma_7 D_l + \gamma_8 R_l & +\delta_3^k X_i^k & +\upsilon_3 \\ \alpha_4 + & \beta_5 I_i & +\gamma_9 N_i + \gamma_{10} A_j^2 + \gamma_{11} D_l + \gamma_{12} R_l & +\delta_4^k X_i^k & +\upsilon_4 \\ \alpha_5 + & \beta_6 C_i + \beta_7 I_i & +\gamma_{13} A_j + \gamma_{14} D_l + \gamma_{15} G_l & +\delta_5^k X_i^k & +\upsilon_5 \end{pmatrix} \tag{9}$$

The mass of observations comprises firms with no expenditures on R&D. When we added an R&D dummy variable to the first equation of the basic specification, we treated the extensive margin of R&D as purely exogenous. In a final step, we want to endogenize the R&D dummy in a separate equation, thus explaining both the extensive and intensive margin of R&D within the system. Apart from the addition of another equation, the variable E_i^d shifts from the vector of exclusionary restrictions to the vector of endogenous variables, which leads to the change of the coefficient from γ to β.

[13] The current set-up worked well in the IV-tests, whereas trials to operate with further squared terms in the fourth equation (with or without interaction terms) failed to pass the Sargan test for over-identification.

Full system:

$$\begin{pmatrix} E_i^d \\ E_i \\ I_i \\ C_i \\ C_i^2 \\ P_i \end{pmatrix} = \begin{pmatrix} \alpha_1+ & \beta_1 C_i + \beta_2 C_i^2 & & +\gamma_1 S_i & +\gamma_2 O_j & & +\delta_1^k X_i^k + \upsilon_1 \\ \alpha_2+ & \beta_3 C_i + \beta_4 C_i^2 & +\beta_5 E_i^d & & +\gamma_3 O_j & & +\delta_2^k X_i^k + \upsilon_2 \\ \alpha_3+ & \beta_6 E_i & & +\gamma_4 S_i & +\gamma_5 M_j & +\gamma_6 D_l & +\delta_3^k X_i^k + \upsilon_3 \\ \alpha_4+ & \beta_7 I_i & +\gamma_7 N_i + \gamma_8 S_i & +\gamma_9 A_j & +\gamma_{10} D_l & +\gamma_{11} R_l & +\delta_4^k X_i^k + \upsilon_4 \\ \alpha_5+ & \beta_8 I_i & +\gamma_{12} N_i + \gamma_{13} S_i & +\gamma_{14} A_j^2 & +\gamma_{15} D_l & +\gamma_{16} R_l & +\delta_5^k X_i^k + \upsilon_5 \\ \alpha_6+ & \beta_9 C_i + \beta_{10} I_i & +\gamma_{17} S_i & +\gamma_{18} A_j & +\gamma_{19} D_l & +\gamma_{20} G_l & +\delta_6^k X_i^k + \upsilon_6 \end{pmatrix} \tag{10}$$

To introduce an additional exclusionary restriction, firm size is no longer covered by the vector of general explanatory variables X_i^k, but now depicted separately as S_i in all but the second equation. This reflects two stylized empirical facts from the literature (e.g., Crepon et al, 1998; Cohen, 2010): First, that large firms generally have a higher probability to engage in R&D. And second, that R&D expenditures tend to grow proportional to size among R&D performing firms. While the first stylized fact leads us to expect that firm size correlates with the extensive margin of R&D, the second fact explains, why it should not correlate with the error term in the estimation of its intensive margin.

The choice of the exclusion restrictions is motivated by theoretical considerations and supported by tests of the validity and power of the instrumental variables. We conduct the *Anderson* canonical correlation test for under-identification, the *Sargan* Test for over-identification, the *Cragg - Donald Wald F* test to detect potentially weak identification, and *Moreira's* (2003) conditional likelihood ratio test for structural models, because of its better power when identification is weak.[14] Dropping a few industry dummies to avoid problems of multicollinearity, the functions pass all the tests for under- and overidentification. Some concerns about potential weak identification remain, but Moreira's conditional likelihood ratio test confirms the sign and significance of the coefficients, even when instruments are weak. Table 6 summarizes the detailed test statistics for the different system equations.

5 Empirical findings

5.1 Sample characteristics

In the estimations we apply country dummies to control for different levels of economic development and the heterogenous ethnic, demographic, cultural and political environments within the CEE and CAC regions. Since the dummies are not displayed in the regression tables, we present the distribution of firms for the three endogenous variables in the Annex, applying the sample weights from the EBRD-WB survey.

Less than 3 percent of the firms report to be a monopoly (Table A1). About 27 percent consider the number of competitors in their main product group to be less than six. An

[14]To perform the IV tests we conduct 2SLS estimations for each equation separately and use Moreira's conditional likelihood test as implemented for Stata by Mikusheva and Poi (2006).

equal percentage of firms believe that these range from 6 to 25, whereas more than 43 percent regard them to be even higher. A minority of less than eight percent of firms reports R&D expenditures, only 1.7 percent report an R&D to sales ratio of more than five percent (Table A2). As more firms report own innovations than R&D, about 18 percent have introduced innovations that were new to the market and mostly achieved by the firm itself (Table A3). Seven percent have introduced innovations new to the market with major contributions from external partners, whereas nine percent only adopted new technologies.

A first casual inspection of the data indicates a kind of inverted-U shaped relationship between competition and innovation. For example, looking at the cross-tabulation of competition with R&D expenditures in Table 1, the share of firms with no R&D expenditures as compared to the other groups is highest among those firms considering themselves either a monopoly or reporting more than 25 competitors. Conversely, the share of firms with an R&D to sales ratio between 1.5 and 5 percent is highest among those, who reportedly have at least one but less than six competitors in their principal product. Consistent with the inverted-U hypothesis, the share of firms with an R&D to sales ratio of more than five percent is largest among those with 6 to 25 competitors.

[Insert Table 1 about here]

Turning to the cross-tabulation of competition with innovation outcome in Table 2, the relative shares of reported monopolies peaks for the group labeled 'Creative I'. In contrast, the relative share of firms with more than 25 competitors in their principal product market peaks for those firms reporting no innovations and then declines monotonically for better values on the innovation variable. Conversely, among the firms with 6 to 25 competitors the most innovative firms also capture the highest share, whereas the firms labeled 'Creative I' have the highest share among those with at least one and less than six competitors.

[Insert Table 2 about here]

5.2 The endogenous equations

Main findings

Starting with the basic system of three endogenous equations, the first column in Table 3 reports the coefficients of the *innovation opportunity* function. They are significant and positive for the linear term and negative for the quadratic term, confirming a nonlinear, inverted-U shaped impact of competition C_i on research expenditures E_i. Growing competition raises the intensive margin of R&D,[15] but at a diminishing rate. After R&D expenditures reach a maximum at intermediate levels of competition, they decrease with a further increase in the number of competitors.

[15] That is, the firms' R&D to sales ratio conditional on the R&D dummy included in the regression.

Among the control variables, a higher share of university graduates, own exports or being a supplier to other exporting firms have a positive impact on R&D expenditures. As regards firm size, the probability of a high R&D to sales ratio is largest for small firms. The firm's age and being foreign or state owned have no significant effect on R&D expenditures. Not surprisingly, being in a sector or technological regime with generally high opportunities for R&D also raises the probability to be a high R&D performer in our sample.

The second column displays the coefficients of the *innovation production* function. Research expenditures E_i have the expected positive impact on the innovation outcome I_i. Higher R&D expenditures buy a higher probability of successfull innovations. Among the exogenous variables and conditional on the given R&D expenditures, a higher share of university graduates, own exports, being a supplier to exporting firms, firm size, foreign ownership, the cumulativeness of knowledge and a high population density raise the probability to innovate successfully. In contrast, state ownership has a negative effect.

The final column reports the coefficients of the *innovation impact* function. The impact of innovation I_i is negative and thus consistent with a fundamental assumption of Schumpeterian growth models. Innovative firms face fewer competitors than technology adopters, and firms without technological innovations must compete with the largest number of rivals. Among the exogenous variables, state ownership typically goes together with a lower number of competitors, whereas competition increases with the share of university graduates, firm size, being an exporter or a supplier to exporting firms, a high population density and better regulatory quality. The number of competitors seems to rise with better appropriability conditions, which may indicate that a better protection of small innovative firms allows more competitors to stay in the market.

[Insert Table 3 about here]

The R-squared ranges from 0.621 to 0.965. Apart from the model specification and many industry and country dummies, this is also owed to the ordinal nature of the variables, which provides for a relatively robust metric. The reason is that the individual categories focus only on pronounced differences within broad boundaries and thereby tend to reduce the inevitable noise of enterprise survey data.

Sensitivity analysis

We perform several robustness checks ensuring the structural validity of the model. For example, we consecutively leave out individual countries such as Russia and Ukraine, which contribute particularly large samples, from the estimation. In each case, the significance levels and magnitude of the coefficients remain qualitatively the same. Next, we split the initial sample into the two major geographic areas to control for possible effects due to cross-country differences in the implementation of reforms (Böheim and Friesenbichler, 2016). This leaves us with 5,585 observations from the CEE and only 1,330 observations from the CAC. Reflecting their large share in the overall sample, the results for the CEE are generally very similar to the previous findings. For the CAC several exogenous variables become insignificant due to the small number of observations. But the basic relationships among the endogenous variables remain intact and surprisingly strong (Table A4).

Furthermore, we applied different thresholds for the variables on state and foreign ownership and estimated the system with a blocking minority (i.e. 25% plus) and a majority share (51% plus). The endogenous system is not sensitive to that change and the coefficients of the endogenous variables have the same sign and significance and are of very similar size. However, the signs of the coefficients on the ownership variables themselves become insignificant, which may be due to the small subsample of foreign and state-owned firms.

The taxonomies group industries according to innovation properties, and thereby should not capture the level effects of single industries, which have to be considered separately. To test for multicollinearity, we implemented single equation regressions of the basic system. The results show that multicollinearity is not a problem and the variance inflation factors (*vif*) are well below the critical ("rule-of-thumb") benchmark of 10.[16]

Finally, we also re-estimate the model using the stratified sample weights, which are available at the firm level from the database ('wstrict'). The inverted-U relationship and most exogenous variables hardly change, neither in size nor significance. However, opportunity conditions in the first equation and research expenditures in the second equation become insignificant. This may relate e.g. to larger errors for strata with low return rates (and accordingly higher weights per firm in the sample). The inverted-U relationship is also robust, if we use median weights ('wmedian') instead, but more exogenous variables become insignificant.

Extensions

Consistent with the discussion in Section 4, we first extend the basic specification with a separate equation for the quadratic term of the competition variable. The inverted-U relationship proves to be robust as do the coefficients for the endogenous variable in equation 2 and most exogenous variables overall (Table 4). Splitting the innovation impact function into separate equations for the linear and the quadratic term renders the endogenous innovation variable insignificant. However, this is no longer the case, when we estimate the full system reported below. Maybe more important is that the coefficient of appropriability conditions turns negative, which is now consistent with the general idea that intellectual property rights help firms to reduce competition.

[Insert Table 4 about here]

[16]The variance inflation factor (*vif*) is 3.76 for the opportunity conditions in equation 1, 2.95 for the cumulativeness of knowledge in equation 2, and 5.48 for the appropriability conditions in equation 3. (Even in the case of multicollinearity, results from Monte-Carlo analyses suggest that the presently used 3SLS estimator is to be preferred over other estimators; see Agunbiade, 2011).

Finally, we add a separate equation to explain the R&D dummy as an endogenous dependent variable (Table 5). This considerably improves the performance of the model, as the expected endogenous relationships are all strongly confirmed. The inverted-U shaped impact of competition on both the extensive and intensive margins of R&D, the positive effect of the intensive margin of R&D on innovation outcome and the negative impact of innovation on the linear and the squared number of competitors are all significant at the one percent level.

The technological regimes lose in significance as exclusionary restrictions. While sectoral opportunity conditions are significant only in the first but not in the second equation. The coefficient on the cumulativeness of knowledge changes its sign in the third equation, while appropriability conditions become insignificant in the fourth and fifth equation. Among the exclusionary restrictions, firm size, political instability, population density and regulatory quality appear to play a stronger role for identification.

[Insert Table 5 about here]

The R-squared hardly changes for the innovation impact functions with C_i and C_i^2 as dependent variables or the intensive margin of R&D. It drops considerably in the innovation production function and is negative in the function for the extensive margin of R&D.[17]

While the separate estimation for the CAC region could still carry the basic system with three equations, the five equations of the full system weigh too heavily on its small sample and the inverted-U becomes insignificant for the extensive margin of R&D, innovation no longer has a significant impact on competition, and many exogenous variables turn insignificant as well. In contrast, if we estimate the full system for the CEE region only, the results are again very robust (Table A5).

[Insert Table 6 about here]

In short, the inverted-U effect of competition on R&D proves to be surprisingly robust in each of the three model specifications, when we use the total firm sample, as well as for the reduced samples of the CEE and CAC, if we stick to the basic system of endogenous equations. Not surprisingly, the added complexity of the extended and the full system comes at a cost in terms of its poorer small sample performance.

[17] With 3SLS (and 2SLS alike) R-squared can be negative, because actual values are used for its determination instead of the instruments. As a consequence, the residuals are computed from a different set of regressors than those used to fit the model and its sum of squares is no longer constrained to be smaller than the total sum of squares.

5.3 The productivity equation

Productivity is a strictly dependent variable and adding the productivity equation therefore does not alter the outcome for the endogenous system. Tables 7 and 8 summarize the productivity estimates for the different system specifications in comparison to a single equation regression using the total sample. Tables A6 and A7 in the Annex report the productivity estimates for the CEE and CAC regions.[18]

To begin with the common findings that are little affected by estimating the simultaneous system: innovation, education, exports, foreign ownership and population density all contribute positively to the sales per employe. In contrast, the sectoral appropriability conditions consistently show a negative impact on productivity. The reason is that higher appropriability hinders the rapid diffusion of new technologies and allows protected firms to last, even with less efficient operations.

But for other variables the simultaneous system makes a striking difference. Most importantly, competition is insignificant in the single equation, but a significant positive driver in each of the three system estimations. In contrast, the single equation attributes much of the variation in productivity to firm size, which is however mostly insignificant in the simultaneous system. In other words, the importance often attributed to firm size is unwarranted according to the more elaborate structural model. Conversely, it is only by controlling for the endogenous co-determination of competition and innovation in the simultaneous system that we can identify both as significant drivers of firm-level productivity. The results are then surprisingly similar for the different system specifications.[19]

[Insert Table 7 about here]

Changing the productivity measure from total sales to *value added per employee* reduces the sample substantially from 7,865 to 1,913 observations for the estimation of the single equation and from 6,915 to 1,655 observations for the estimation of the basic system (Table 8). Among the exogenous variables, education, foreign ownership and appropriability stay significant in both estimations, whereas exports are significant only in the single equation and population density only in the basic system. For the single equation, competition remains insignificant and innovation is significant only at ten percent. In contrast, the basic

[18]If we restrict the sample to firms from the CEE, the results for the different dependent variables in the basic system are again very similar. For the CAC region, however, the samples become too small for any reasonable performance of the model.

[19]Because of the different methodologies a direct comparison of the magnitude of coefficients with other results reported in the literature is difficult to establish. The simultaneously estimated coefficients of innovation on productivity are in the range of other estimates using similar data. For instance, the EBRD (2014) reports results from a CDM model in which innovation is measured by dichotomous variables. The coefficients to the mean depend on the type of innovation, and range between 0.18 and 0.36. Similar magnitudes of coefficients for innovations in CDM estimations have been reported in a recent literature review by Mohnen and Hall (2013), even though the presently provided findings would be on the upper bound.

system again confirms a positive and significant simultaneous impact of both competition and innovation on firm level productivity.

Finally, using *multi factor productivity* (MFP) as the dependent variable, the sample size further reduces to 1,091 observations for the single equation and to 900 firms for the basic system. Not surprisingly, several control variables become insignificant. Only foreign ownership and firm size are significant in both specifications.[20] The basic system maintains the significant positive coefficient on innovation, whereas the competition variable becomes insignificant. The same is true, if we use either sales or value added per employee as the dependent veriable, but maintain the reduced sample of firms with data on MFP. This suggests that the small sample size is a major cause of the weaker findings on MFP. Another reason may be well known difficulties in the appropriate measurement of MFP.[21]

[Insert Table 8 about here]

Overall, by comparison with the estimation of single equations, the results demonstrate the benefits of accounting for the endogeneity between competition and innovation in a simultaneous system. In contrast, for a sufficiently large number of observations the various extensions of the basic model make little difference for the productivity equation. They unequivocally show both competition and innovation to be simultaneous drivers of firm-level productivity.

6 Summary and conclusions

Does more competition lead to more innovation and higher productivity? Though one would easily think so according to most popular notions, the scientific debate has shown the relationship to be more complex and eschew simplistic generalisations. Benefitting from the increasing availability of firm level data, it is in particular the *inverted-U* hypothesis of competition and innovation, which in recent years has triggered a rapidly growing body of empirical analyses. Due to the high demands on statistical offices, such data and according empirical studies were so far only available for a few developed economies. This left doubts about the generalization of such findings to firms in less developed countries with substantially different institutional environments.

Using a comprehensive sample of firms from the EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS) for Central Eastern Europe (CEE) as well as Central Asia and the Caucasus (CAC), our findings strongly confirm the *inverted-U* hypothesis and a simultaneous, yet independent, positive impact of competition and innovation on productivity.

From these findings we draw the following conclusions:

[20] See also Dachs and Peters (2014) for a similar finding with EU-CIS data.

[21] For example, if the residual captures a firm's profitability rather than its pure technical efficiency, one should even expect a negative relationship with the intensity of competition.

In *theoretical* perspective, the simultaneous system directs our attention towards the joint determination of separate functions and away from the often misleading interpretation of single equations. This is of particular importance for transformation processes in catching-up economies. For example, without a system approach, the typical conclusion drawn from an inverted-U relationship is that an intermediate degree of competition is most conducive to *maximize* innovation. But only under very specific circumstances will the system ever settle for a maximum of innovation. Since innovation breeds cost, a maximum of innovation is neither what individual firms aim for, nor a desirable policy objective as such. Instead, the system perspective shows that firms and markets consider trade-offs and typically settle at intermediate levels of competition *and* innovation.

With regard to *empirical* analysis, the major caveat to our findings is the limited availability of data on capital stocks and intermediary goods. This forced us to use total sales per employee as our main dependent variable. Alternative estimates for value added per employee and MFP are consistent with many of the initial results, but suffer from the much smaller number of observations. To draw firmer conclusions, further micro data on the actual use of intermediary and capital goods are warranted. Despite the enormous effort already undertaken, collecting additional data on larger samples will be key to future progress.

There is no 'one-size-fits-all' solution to *economic policy* in general, and even more so for economies in transition (EBRD, 2014, 2015). The system perspective on mutual interdependencies calls for caution with regard to mechanistic expectations about the impact of public interventions. If we understand development as the combination of growth and qualitative transformations (Schumpeter, 1947), policy must seek the dialogue with the stakeholders in place, who have the expertise about the specific industrial structures and institutions that shape the formation of competitive advantages.[22] Because of the asymmetric distribution of information, however, such a dialogue also comes at the risk of capture by incumbent interests. Ambiguous predictions about the impact of competition on innovation may then be easily exploited.

Interestingly, it is the nonlinear nature of the inverted-U hypothesis, that can resolve much of the mystery also for policy making. The idea of competition being detrimental to innovation and hence development is often used as a stalking-horse to oppose the strict application of antitrust and merger rules as well as the opening of markets to trade and entry, or other regulatory reforms. While in practice each case needs to be looked at in detail, some general conclusions emerge, if we deliberately condition our considerations on the initial intensity of competition in the market:

– Cases in *antitrust and merger control* typically are subject to closer investigation only when there already is reason to suspect a low level of initial competition. As a general rule, we should hence find most cases in the upward sloping part of the inverted-U, and expect that innovation also benefits from a strict application to maintain a high

[22] See, e.g., Rodrik (2008), Farla (2015), or Stoellinger and Holzner (2016).

number of competitors. From this follows an unambiguously positive impact on productivity.

- Similar reasoning applies to *regulatory reforms* that open markets with territorial protection, e.g. (local) monopolies, to the entry of new competitors. If we disregard other potential causes (e.g., a “natural monopoly"), starting from low initial rivalry, more competition will generally raise innovation and productivity.

- The implications of the inverted-U relationship are potentially ambiguous for the case of *trade liberalisation*, since it may also affect markets with high initial domestic competition. Again disregarding all other factors,[23] it is however consistent with the principle of asymmetric liberalization. According to it, firms in less developed economies enjoy some protection, temporarily and on a diminishing scale, from what otherwise would be an overwhelming foreign competition that can inhibit the build-up of own innovation and production capabilities. At the same time it points at the need for effective domestic competition to make the scheme work.[24]

Over and above these considerations, the productivity estimates invariably support an independent positive impact of competition, which comes in addition to the effects from innovation, education, exports or foreign ownership. This corroborates a reform agenda which combines investments in productive capabilities and innovation with openness and effective competition to further structural change and the productivity gains that come with it.

Acknowledgements

Our analysis has become possible by way of the open access generously offered by the EBRD and World Bank to their joint *Business Environment and Enterprise Performance Survey* (BEEPS). We are deeply indebted to both organisations. Furthermore, financial support from the Oesterreichische Nationalbank (OeNB, Anniversary Fund, project number: 15280) is gratefully acknowledged. We are especially grateful to Jesus Crespo-Cuaresmo, Harald Oberhofer, Philipp Schmidt-Dengler, Michael Pfaffermayr, two anonymous referees and the editor for their critical comments and constructive advise. Finally, we want to thank Anna Strauss for her technical support in handling the data and editing the text. The usual disclaimer applies.

[23] Such as higher consumer welfare, reciprocal access to larger markets, or faster technology diffusion, which all add to the arguments in favor of trade liberalization.

[24] Historically, many examples failed because of the lack of an effective domestic competition. The related infant industry argument nevertheless has a strong historical record, starting with the early industrialization in Europe and the U.S. up to the recent East Asian success stories of South Korea or China (see, e.g., Chang, 2002, 2006).

References

Acemoglu, D., Aghion, P. and Zilibotti, F. (2006). 'Distance to frontier, selection, and economic growth', Journal of the European Economic Association, 4, pp. 37-74.

Aghion, P., Bloom, N., Blundell, R., Griffith, R. and Howitt, P. (2005). 'Competition and innovation: an inverted-U relationship', Quarterly Journal of Economics, 120, pp. 701-728.

Aghion, P., Howitt, P. (1992). 'A model of growth through creative destruction', Econometrica, 60, pp. 323-351.

Aghion, P., Howitt, P. (2009). The economics of growth, Cambridge MA: MIT Press.

Angrist, J.D., Pischke, J.-S. (2009). *Mostly Harmless Econometrics. An Empiricist's Companion*, Princeton: Princeton University Press.

Agunbiade, D.A. (2011), Effect of multicollinearity and the sensitivity of the estimation methods in Simultaneous Equation Models, Journal of Modern Mathematics and Statistics, 5 (1), pp. 9-12.

Böheim, M., Friesenbichler, K. (2016), 'Exporting the competition policy regime of the European Union: success or failure? Evidence for acceding countries', Journal of Common Market Studies, Online First. DOI: 10.1111/jcms.12321.

Carlin, W., Schaffer, M. and Seabright, P. (2004). 'A Minimum of Rivalry: Evidence from transition economies on the importance of competition for innovation and growth', Contributions to Economic Analysis & Policy, 3 (1), pp. 1-43.

Castellacci, F. (2010). 'How does competition affect the relationship between innovation and productivity? Estimation of a CDM model for Norway', Economics of Innovation and New Technology, 19, pp. 1-22.

Chang, H.-J. (2002). *Kicking Away the Ladder – Development strategy in historical perspective*, Anthem Press.

Chang, H.-J. (2006). *The East Asian development experience*, Zed Books, London.

Crepon, B., Duguet E. and Mairesse J. (1998), 'Research and development, innovation and productivity: an econometric analysis at the firm level', Economics of Innovation and New Technology, 7 (2), pp. 115-158.

Dachs, B., Peters, B. (2014), 'Innovation, employment growth, and foreign ownership of firms. A European perspective', Research Policy 43 (1), 214-232.

De Bondt, R., Vandekerckhove, J. (2012). 'Reflections on the relation between competition and innovation', Journal of Industry, Competition and Trade, 12 (1), pp. 7-19.

Dutz, M.A., Vagliasindi, M. (2000). 'Competition Policy Implementation in Transition Economies: An Empirical Assessment', European Economic Review, 44 (4-6), pp. 762-72. doi:10.1016/S0014-2921(99)00060-4.

EBRD (2013). EBRD Transition Report 2013: Stuck in transition, London: EBRD Publications (available at http://2013.tr-ebrd.com/).

EBRD (2014). EBRD Transition Report 2014: Innovation in transition, London: EBRD (available at http://2014.tr-ebrd.com/).

EBRD (2015). EBRD Transition Report 2015-2016: Rebalancing finance, London: EBRD (available at http://2015.tr-ebrd.com/).

Farla, K. (2015), 'Industrial Policy for growth', Journal of Industry, Competition and Trade, 15

(3), pp. 257-282.

Friesenbichler, K., Boeheim, M. and Laster, D.-C. (2014). ‘Market competition in transition economies: A literature review’, WIFO-Working Paper No. 477, Vienna: WIFO.

Gilbert, R. (2006). ‘Looking for Mr. Schumpeter: Where are we in the competition-innovation debate?’, in: Jaffe, A.B., Lerner, J. and Stern, S. (eds), Innovation policy and the economy. Volume 6, pp. 159-215.

Gorodnichenko, Y., Svejnar, J. and Terrell, K. (2010). ‘Globalization and innovation in emerging markets’, American Economic Journal: Macroeconomics, 2 (2), pp. 194-226. doi:10.1257/mac.2.2.194.

Hare, P., Turley, G. (eds.) (2013). Handbook of the economics and political economy of transition, London: Routledge.

Kamien, M.I., Schwartz, N.L. (1976). ‘On the degree of rivalry for maximum innovative activity’, Quarterly Journal of Economics, 90, pp. 245-260.

Kaufmann, D., Kraay, A. and Mastruzzi, M., (2010). ‘The worldwide governance indicators. Methodology and analytical Issues’, Policy Research Working Paper 5430 (September), pp. 1-31.

Kornai, J. (1992). ‘The postsocialist transition and the state: reflections in the light of Hungarian fiscal problems’, American Economic Review, 82 (2), pp. 1-21.

Martin, S. (2012). ‘Market structure and market performance’, Review of Industrial Organization, 40, pp. 87-108.

Menard, S.W. (2001). *Applied Logistic Regression Analysis (Quantitative Applications in the Social Sciences)*, Sage Publications.

Mikusheva, A., Poi, B.P. (2006), ‘Tests and confidence cets with correct size when instruments are potentially weak’, Stata Journal 6 (3), pp. 335-347.

Mohnen, P., Hall, B.H (2013), ‘Innovation and Productivity: An Update’, Eurasian Business Review, 3(1), 47-65.

Moreira, M.J. (2003), ‘A conditional likelihood test for structural models’, Econometrica 71 (4), 1027-1048.

OECD (2012). Measuring R&D in developing countries – Annex to the Frascati Manual, Paris: OECD Publications.

Peneder, M. (2010). ‘Technological regimes and the variety of innovation behavior: creating integrated taxonomies of firms and sectors’, Research Policy, 39, pp. 323-334.

Peneder, M. (2012) ‘Competition and innovation: revisiting the inverted-U relationship’, Journal of Industry, Competition and Trade, 12 (1-2), pp. 1-6.

Peneder, M., Wörter, M. (2014). ‘Competition, R&D and innovation: testing the inverted-U in a simultaneous system’, Journal of Evolutionary Economics, 24 (3), pp. 653-687.

Rodrik, D. (2008). Normalizing industrial policy, Commission on Growth and Development, Working Paper No. 3, The World Bank, Washington.

Saliola, F., Seker, M. (2011). ‘Total factor productivity across the developing world’, Enterprise Surveys - Enterprise Note Series No. 23., The World Bank, IFC, Washington DC.

Schumpeter, J.A. (1947). The creative response in economic history, *Journal of Economic History* 7 (2), 149-159.

Stoellinger, R., Holzner, M. (2016), ‘State aid and export competitiveness in the EU’, Journal of

Industry, Competition and Trade (forthcoming).

Sutton, J. (1998). Technology and market structure. Theory and history, Cambridge MA: MIT Press.

Sutton, J. (1991). Sunk costs and market structure. Price competition, advertising and the evolution of concentration, Cambridge MA: MIT Press.

Syverson, C. (2011). ‘What determines productivity?’. Journal of Economic Literature, 49, pp. 326-365.

UNESCO (2010). Measuring R&D Challenges faced by developing countries. Montreal: UNESCO Publications.

Tables

Table 1: Competition and research expenditures – share of firms in percent

R&D / sales in %	Number of competitors: Monopoly	... < 6	6 - 25	... > 25	Total
None	2.8	26.8	27.1	43.3	100.0
... < 1.5	1.4	34.3	34.5	29.8	100.0
1.5 - 5	2.0	41.7	33.2	23.0	100.0
... > 5	2.5	36.0	48.0	13.5	100.0
Total	2.7	27.6	27.9	41.8	100.0

NB: Sample weights used ("wstrict").
Source: EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS).

Table 2: Innovation outcome and competition – share of firms in percent

Firm type	Number of competitors: Monopoly	... < 6	6 - 25	... > 25	Total
Adaptive I	2.5	24.9	22.6	50.0	100.0
Adaptive II	2.4	29.2	32.2	36.2	100.0
Creative I	4.4	31.0	31.0	33.6	100.0
Creative II	2.8	30.4	38.6	28.2	100.0
Total	2.7	26.7	27.0	43.5	100.0

NB: Sample weights used ("wstrict").
Source: EBRD-WB Business Environment and Enterprise Performance Survey (BEEPS).

Table 3: Basic endogenous system

VARIABLES	Research	Innovation	Competition
Competition	0.749***		
	(0.0200)		
Compet. squ	-0.132***		
	(0.00357)		
Research		0.0717**	
		(0.0322)	
Innovation			-1.613***
			(0.0501)
Education (univ)	0.000233**	0.00234***	0.00417***
	(0.000103)	(0.000489)	(0.000594)
Exports indirect	0.0319**	0.412***	0.688***
	(0.0146)	(0.0693)	(0.0849)
Exports direct	0.0224**	0.298***	0.508***
	(0.00976)	(0.0464)	(0.0572)
Size med	-0.0185***	0.159***	0.244***
	(0.00617)	(0.0293)	(0.0358)
Size large	-0.0464***	0.358***	0.459***
	(0.00992)	(0.0470)	(0.0587)
Age (ln)	-0.00486	0.0323	0.0106
	(0.00443)	(0.0209)	(0.0249)
Foreign owned	-0.0157	0.113**	-0.0184
	(0.0120)	(0.0571)	(0.0682)
State owned	0.00466	-0.193*	-0.379***
	(0.0234)	(0.111)	(0.133)
Pop. density (ln)		0.164***	0.633***
		(0.0271)	(0.0385)
Regulatory quality			0.694***
			(0.0547)
Opportunity	0.0236**		
	(0.00981)		
Cumulativeness		0.311***	
		(0.0467)	
Appropriability			0.171***
			(0.0288)
Observations	6,915	6,915	6,915
R-squared	0.965	0.727	0.621

NB: R&D dummy included in first equation.

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Source: EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS).

Table 4: Extended endogenous system

VARIABLES	Research	Innovation	Competition	
			Linear	Squared
Competition	0.534***			
	(0.138)			
Compet. squ	-0.0725*			
	(0.0371)			
Research		0.746***		
		(0.0336)		
Innovation			-0.0308	-0.257
			(0.0384)	(0.221)
Education (univ)	0.000184	0.00186***	0.000338	0.00245
	(0.000113)	(0.000494)	(0.000423)	(0.00244)
Exports indirect	0.0365**	0.263***	0.00955	0.00377
	(0.0154)	(0.0701)	(0.0604)	(0.348)
Exports direct	0.0230**	0.198***	0.0241	0.114
	(0.0101)	(0.0468)	(0.0406)	(0.234)
Size med	-0.0176***	0.133***	-0.0157	-0.0742
	(0.00648)	(0.0296)	(0.0254)	(0.146)
Size large	-0.0307**	0.299***	-0.118***	-0.616**
	(0.0130)	(0.0475)	(0.0419)	(0.242)
Age (ln)	0.00279	0.0319	-0.0484***	-0.300***
	(0.00647)	(0.0211)	(0.0177)	(0.102)
Foreign owned	0.00855	0.109*	-0.205***	-1.143***
	(0.0195)	(0.0574)	(0.0483)	(0.278)
State owned	0.0101	-0.147	-0.0909	-0.341
	(0.0246)	(0.113)	(0.0951)	(0.548)
Pop. density (ln)		0.0542**	0.655***	2.659***
		(0.0274)	(0.0275)	(0.157)
Regulatory quality			0.225***	0.327
			(0.0388)	(0.224)
Political instability			0.0416***	0.219***
			(0.00856)	(0.0494)
Opportunity	0.0452***			
	(0.0166)			
Cumulativeness		0.202***		
		(0.0471)		
Appropriability			-0.0784***	
			(0.0205)	
Approp. squ				-0.0746***
				(0.0197)
Observations	6,797	6,797	6,797	6,797
R-squared	0.961	0.740	0.919	0.787

NB: R&D dummy included in first equation.

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS).

Table 5: Full endogenous system

VARIABLES	Research		Innovation	Competition	
	Extensive	Intensive		Linear	Squared
Competition	1.209***	2.031***			
	(0.422)	(0.483)			
Compet. squ	-0.213***	-0.371***			
	(0.0757)	(0.0866)			
Research extens.		1.166***			
		(0.124)			
Research intens.			4.128***		
			(0.461)		
Innovation				-0.494***	-3.164***
				(0.154)	(0.766)
Education (univ)	0.000472***	0.000524***	-0.000968	0.00145**	0.00953***
	(0.000131)	(0.000148)	(0.000766)	(0.000617)	(0.00347)
Exports indirect	0.0826***	0.0562**	-0.416***	0.195**	1.179**
	(0.0192)	(0.0243)	(0.132)	(0.0883)	(0.490)
Exports direct	0.0620***	0.0440***	-0.274***	0.163***	0.996***
	(0.0126)	(0.0167)	(0.0909)	(0.0633)	(0.349)
Size med	0.0388***		-0.00539	0.0551	0.380*
	(0.00701)		(0.0410)	(0.0360)	(0.202)
Size large	0.0915***		0.00394	0.0441	0.427
	(0.0128)		(0.0704)	(0.0696)	(0.377)
Age (ln)	-0.00353	-0.0146**	0.0203	-0.0351*	-0.217*
	(0.00629)	(0.00690)	(0.0285)	(0.0196)	(0.117)
Foreign owned	0.0207	-0.0226	0.0764	-0.149***	-0.791**
	(0.0166)	(0.0185)	(0.0775)	(0.0553)	(0.328)
State owned	0.0251	0.0408	-0.0207	-0.172	-0.857
	(0.0328)	(0.0359)	(0.154)	(0.106)	(0.635)
Pop. density (ln)			-0.524***	1.271***	2.718***
			(0.0868)	(0.0877)	(0.240)
Regulatory quality				-0.750***	1.170***
				(0.0910)	(0.342)
Political instability				0.0737***	0.413***
				(0.0138)	(0.0752)
Opportunity	0.0392***	0.00451			
	(0.00806)	(0.0106)			
Cumulativeness			-0.239***		
			(0.0723)		
Appropriability				-0.00729	
				(0.0308)	
Approp. squ					-0.00688
					(0.0254)
Observations	6,797	6,797	6,797	6,797	6,797
R-squared	-0.149	0.913	0.208	0.896	0.712

NB: Country & industry dummies in all equations.

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS).

Table 6: Summary of IV Test Statistics

System	Under-identification		Weak identification		Over-identification	
Equation	Anderson LM	P-value	Cragg-Donald Wald F	Stock-Yogo (critical value)	Sargan	P-value
BASIC						
Equation 1:	9.420	0.024	3.115	3.32 (25%)	0.204	0.903
Equation 2:	6683.310	0.000	818.972	8.68 (10%)	0.030	0.863
Equation 3:	6.379	0.012	6.341	5.53 (25%)	exactly identified	
EXTENDED						
Equation 1:	14.540	0.0007	8.811	3.81 (15%)	0.038	0.845
Equation 4:	22.210	0.000	22.114	16.38 (10%)	exactly identified	
FULL						
Equation 1:	25.544	0.000	8.479	5.44 (10%)	0.159	0.690
Equation 2:	5.481	0.0645	1.359	not available	0.109	0.742

NB: Tests from 2SLS estimations. The *Anderson* canonical correlation statistic should reject the null-hypothesis of under-identification, the *Sargan* test fail to detect over-identification. *Cragg-Donald Wald* statistic compares to the *Stock-Yogo* critical value (closest benchmark of maximal bias in brackets). *Moreira's* conditional likelihood ratio test for structural models with weak identification (not displayed) confirms correct coverage probability with p-values ranging from 0.0000 to 0.0023 (except equation 2 of the full system, since it is not available in the presence of three endogenous variables.)

Table 7: The productivity estimates – sales per employee

Variables	Single equation	Simultaneous system Basic	Extended	Full
Competition	0.00364	0.0431**	0.701**	0.337***
	(0.0157)	(0.0177)	(0.319)	(0.12)
Innovation	0.0432***	0.451***	0.261***	0.369*
	(0.0125)	(0.0539)	(0.0552)	(0.219)
Education (univ)	0.00530***	0.00407***	0.00426***	0.00403***
	(0.000542)	(0.000595)	(0.000638)	(0.000796)
Exports (indirect)	0.138*	-0.0411	0.045	-0.00234
	(0.0736)	(0.0852)	(0.0901)	(0.126)
Exports (direct)	0.296***	0.134**	0.192***	0.161*
	(0.0484)	(0.0574)	(0.0605)	(0.0894)
Size med	0.0958***	0.00495	0.0426	0.0154
	(0.0319)	(0.0358)	(0.0385)	(0.0516)
Size large	0.216***	0.0663	0.219***	0.126
	(0.0508)	(0.0589)	(0.0758)	(0.107)
Age (ln)	-0.0203	-0.0423*	-0.00358	-0.0238
	(0.0227)	(0.0249)	(0.0305)	(0.0287)
Foreign owned	0.417***	0.359***	0.533***	0.444***
	(0.0617)	(0.0681)	(0.0997)	(0.0869)
State owned	-0.185	-0.167	-0.15	-0.16
	(0.123)	(0.133)	(0.144)	(0.146)
GDP pc (ln)	1.106***	1.009***	0.796***	0.913***
	(0.0184)	(0.0175)	(0.113)	(0.0342)
Pop. Density	0.215***	0.390***	0.274***	0.334***
	(0.0454)	(0.0328)	(0.0748)	(0.0379)
Appropriability	-0.0676***	-0.137***	-0.0522	-0.0844*
	(0.0254)	(0.0288)	(0.0413)	(0.0498)
Observations	7,865	6,915	6,797	6,797
R-squared	0.986	0.984	0.982	0.984

NB: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.
Source: EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS).

Table 8: The productivity estimates – value added per employee and MFP

Variables	VA p.e. Single equation	Basic system	MFP Single equation	Basic system
Competition	-0.00561	0.0716**	-0.000622	0.00225
	(0.0301)	(0.0346)	(0.00264)	(0.00298)
Innovation	0.0370*	0.569***	0.000692	0.0220***
	(0.0215)	(0.0904)	(0.00176)	(0.00743)
Education (univ)	0.00459***	0.00344**	0.000200**	0.000161
	(0.00118)	(0.00134)	(9.97E-05)	(0.000113)
Exports (indirect)	0.178	-0.121	0.00161	-0.00965
	(0.119)	(0.141)	(0.00939)	(0.0111)
Exports (direct)	0.304***	0.127	0.0140**	0.00874
	(0.0772)	(0.0941)	(0.00615)	(0.00753)
Size med	0.0729	0.055	-0.0128**	-0.0127**
	(0.0603)	(0.0682)	(0.00513)	(0.00581)
Size large	0.0858	-0.0912	-0.0275***	-0.0356***
	(0.0911)	(0.108)	(0.00772)	(0.00914)
Age (ln)	0.0441	0.0159	0.00268	0.00259
	(0.0428)	(0.0485)	(0.00364)	(0.0042)
Foreign owned	0.546***	0.572***	0.0209**	0.0262**
	(0.113)	(0.127)	(0.00942)	(0.0104)
State owned	-0.26	-0.255	-7.14E-05	0.00761
	(0.206)	(0.231)	(0.0202)	(0.0223)
GDP pc (ln)	1.181***	0.973***	0.00969*	0.00287
	(0.0485)	(0.0502)	(0.00559)	(0.0067)
Pop. Density	0.136	0.402***	-0.00234	-0.000981
	(0.0827)	(0.0788)	(0.00731)	(0.00912)
Appropriability	-0.281***	-0.266***	-0.000178	0.000419
	(0.0662)	(0.0739)	(0.00852)	(0.0114)
Observations	1,933	1,655	1,091	900
R-squared	0.986	0.981	0.541	0.466

NB: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.
Source: EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS).

Annex

Table A1: Intensity of competition, 2013

Country	Number of competitors (main product) None	< 6	6 - 25	> 25	Total
	Share of firms in percent				
Central and Eastern Europe (CEE)					
Albania	6.6	26.7	18.0	48.8	100.0
Belarus	6.1	34.7	25.8	33.3	100.0
Bosnia-Herzegovina	6.0	34.1	17.4	42.5	100.0
Bulgaria	1.0	24.5	30.3	44.3	100.0
Croatia	2.9	33.3	44.3	19.5	100.0
Estonia	4.1	27.8	30.7	37.3	100.0
Hungary	3.0	16.1	22.4	58.5	100.0
Kosovo	3.1	27.7	29.7	39.5	100.0
Latvia	4.7	27.6	21.8	45.9	100.0
Lithuania	4.7	20.2	28.4	46.8	100.0
Macedonia	1.8	33.1	31.1	34.0	100.0
Moldova	1.6	21.5	17.7	59.1	100.0
Montenegro	5.0	18.1	47.1	29.8	100.0
Poland	2.7	25.5	29.6	42.2	100.0
Romania	1.0	30.5	27.3	41.2	100.0
Russia	1.6	24.5	28.1	45.9	100.0
Serbia	5.9	35.0	33.3	25.9	100.0
Slovakia	0.9	47.0	38.7	13.4	100.0
Slovenia	5.7	45.0	35.9	13.4	100.0
Ukraine	0.6	21.9	26.9	50.7	100.0
Central Asia and Caucasus (CAC)					
Armenia	8.0	70.3	17.4	4.3	100.0
Azerbaijan	13.9	16.4	5.1	64.7	100.0
Georgia	5.8	29.9	16.8	47.5	100.0
Kazakhstan	0.8	23.3	22.8	53.1	100.0
Kyrgyzstan	3.6	31.7	20.2	44.6	100.0
Tajikistan	6.9	35.5	21.1	36.5	100.0
Total	2.7	26.7	27.1	43.6	100.0

NB: Sample weights used ("wstrict"). Survey data for Russia refer to 2012.
Source: EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS).

Table A2: Research expenditures, 2013

Country	None	< 1.5	1,5 - 5	> 5	Total
	R&D to sales ratio in %				
			Share of firms in percent		
Central and Eastern Europe (CEE)					
Albania	6.6	26.7	18.0	48.8	100.0
Belarus	6.1	34.7	25.8	33.3	100.0
Bosnia-Herzegovina	6.0	34.1	17.4	42.5	100.0
Bulgaria	1.0	24.5	30.3	44.3	100.0
Croatia	2.9	33.3	44.3	19.5	100.0
Estonia	4.1	27.8	30.7	37.3	100.0
Hungary	3.0	16.1	22.4	58.5	100.0
Kosovo	3.1	27.7	29.7	39.5	100.0
Latvia	4.7	27.6	21.8	45.9	100.0
Lithuania	4.7	20.2	28.4	46.8	100.0
Macedonia	1.8	33.1	31.1	34.0	100.0
Moldova	1.6	21.5	17.7	59.1	100.0
Montenegro	5.0	18.1	47.1	29.8	100.0
Poland	2.7	25.5	29.6	42.2	100.0
Romania	1.0	30.5	27.3	41.2	100.0
Russia	1.6	24.5	28.1	45.9	100.0
Serbia	5.9	35.0	33.3	25.9	100.0
Slovakia	0.9	47.0	38.7	13.4	100.0
Slovenia	5.7	45.0	35.9	13.4	100.0
Ukraine	0.6	21.9	26.9	50.7	100.0
Central Asia and Caucasus (CAC)					
Armenia	8.0	70.3	17.4	4.3	100.0
Azerbaijan	13.9	16.4	5.1	64.7	100.0
Georgia	5.8	29.9	16.8	47.5	100.0
Kazakhstan	0.8	23.3	22.8	53.1	100.0
Kyrgyzstan	3.6	31.7	20.2	44.6	100.0
Tajikistan	6.9	35.5	21.1	36.5	100.0
Total	92.4	3.8	2.1	1.7	100.0

NB: Sample weights used ("wstrict"). Survey data for Russia refer to 2012. R&D activities counted as none, when information about expenditure levels was missing.
Source: EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS).

Table A3: Innovation outcome, 2013

Country	Firm type: Adaptive I	Adaptive II	Creative I	Creative II	Total
			Share of firms in percent		
Central and Eastern Europe (CEE)					
Albania	91.1	2.8	1.5	4.7	100.0
Belarus	52.6	12.1	14.9	20.4	100.0
Bosnia-Herzegovina	50.8	9.1	7.0	33.0	100.0
Bulgaria	68.4	9.7	5.0	16.9	100.0
Croatia	51.8	13.3	5.8	29.1	100.0
Estonia	69.5	11.7	6.7	12.1	100.0
Hungary	69.2	8.8	4.5	17.6	100.0
Kosovo	38.4	8.1	9.7	43.8	100.0
Latvia	77.7	9.3	2.6	10.5	100.0
Lithuania	66.8	9.2	7.2	16.8	100.0
Macedonia	67.9	6.2	4.8	21.1	100.0
Moldova	61.0	6.4	11.6	21.0	100.0
Montenegro	81.3	13.0	2.3	3.5	100.0
Poland	61.7	11.4	4.1	22.9	100.0
Romania	47.3	15.2	13.8	23.7	100.0
Russia	67.6	6.5	6.8	19.1	100.0
Serbia	57.3	14.0	5.4	23.4	100.0
Slovakia	75.6	7.8	2.8	13.8	100.0
Slovenia	63.7	9.9	12.4	14.0	100.0
Ukraine	83.8	9.6	0.0	6.6	100.0
Central Asia and Caucasus (CAC)					
Armenia	84.9	1.6	3.8	9.7	100.0
Azerbaijan	96.7	0.3	2.0	1.1	100.0
Georgia	88.3	1.9	2.1	7.7	100.0
Kazakhstan	78.4	4.5	4.4	12.8	100.0
Kyrgyzstan	51.5	17.2	3.5	27.8	100.0
Tajikistan	76.3	6.2	3.1	14.4	100.0
Total	65.6	9.1	7.0	18.3	100.0

NB: Sample weights used ("wstrict"). Survey data for Russia refer to 2012. Adaptive I are firms that introduced no new products or processes. Adaptive II are firms that introduced either innovation, if these were new to the firm but not new to the market. Creative I are firms that introduced innovations new to the market, but with a major contribution from external partners. Creative II are firms that launched an own innovation new to the market.

Source: EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS).

Table A4: Basic endogenous system – CAC

VARIABLES	Research	Innovation	Competition
Competition	0.800***		
	(0.0237)		
Compet. squ	-0.140***		
	(0.00438)		
Research		0.310***	
		(0.0964)	
Innovation			-1.432***
			(0.118)
Education (univ)	0.000231	0.00202**	0.00568***
	(0.000161)	(0.000954)	(0.00122)
Exports indirect	0.0134	0.302*	0.550**
	(0.0299)	(0.177)	(0.224)
Exports direct	0.0196	0.294**	0.366**
	(0.0234)	(0.139)	(0.177)
Size med	-0.00861	0.0905	0.148*
	(0.0102)	(0.0605)	(0.0764)
Size large	-0.0302*	0.227**	0.207*
	(0.0165)	(0.0973)	(0.125)
Age (ln)	-0.00970	-0.0316	-0.00237
	(0.00729)	(0.0430)	(0.0539)
Foreign owned	-0.0155	0.149	-0.0796
	(0.0196)	(0.116)	(0.146)
State owned	0.146***	0.627*	0.251
	(0.0546)	(0.322)	(0.408)
Pop. density (ln)		0.145***	1.392***
		(0.0411)	(0.0802)
Regulatory quality			-0.473***
			(0.0867)
Opportunity	0.00343		
	(0.0167)		
Cumulativeness		0.0339	
		(0.100)	
Appropriability			-0.0837
			(0.0628)
Observations	1,330	1,330	1,330
R-squared	0.977	0.723	0.711

NB: Country & industry dummies in all equations; R&D dummy in first equation.
Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1
Source: EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS).

Table A5: Full endogenous system – CEE

VARIABLES	Research		Innovation	Competition	
	Extensive	Intensive		Linear	Squared
Competition	1.573***	2.323***			
	(0.436)	(0.511)			
Competit. squ	-0.290***	-0.438***			
	(0.0796)	(0.0935)			
Research extens		1.068***			
		(0.135)			
Research intens			3.438***		
			(0.411)		
Innovation				-0.285**	-2.126***
				(0.131)	(0.634)
Education (univ)	0.000496***	0.000491***	-0.000626	0.000140	0.00240
	(0.000166)	(0.000177)	(0.000800)	(0.000597)	(0.00338)
Exports indirect	0.0848***	0.0618**	-0.330**	0.0963	0.679
	(0.0229)	(0.0277)	(0.132)	(0.0833)	(0.462)
Exports direct	0.0664***	0.0550***	-0.192**	0.107*	0.729**
	(0.0144)	(0.0185)	(0.0873)	(0.0577)	(0.319)
Size med	0.0421***		0.0349	0.0216	0.227
	(0.00850)		(0.0430)	(0.0362)	(0.201)
Size large	0.0946***		0.0432	-0.0152	0.154
	(0.0158)		(0.0743)	(0.0677)	(0.362)
Age (ln)	-0.0120	-0.0249***	0.0299	-0.0570***	-0.334***
	(0.00863)	(0.00923)	(0.0304)	(0.0208)	(0.123)
Foreign owned	0.00381	-0.0396*	0.0562	-0.146***	-0.783**
	(0.0211)	(0.0225)	(0.0831)	(0.0565)	(0.335)
State owned	0.0207	0.0187	-0.176	-0.101	-0.616
	(0.0359)	(0.0380)	(0.155)	(0.108)	(0.635)
Pop. density (ln)			-0.401***	0.716***	2.711***
			(0.0754)	(0.0378)	(0.224)
Regulatory quality				0.245***	0.889***
				(0.0550)	(0.311)
Political instability				0.0458***	0.259***
				(0.0125)	(0.0687)
Opportunity	0.0324***	0.000388			
	(0.0109)	(0.0128)			
Cumulativeness			-0.179**		
			(0.0744)		
Appropriability				-0.0291	
				(0.0312)	
Approp. squ					-0.0215
					(0.0245)
Observations	5,505	5,505	5,505	5,505	5,505
R-squared	-0.402	0.878	0.374	0.915	0.760

NB: Country & industry dummies in all equations.

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS).

Table A6: The productivity estimates – CEE

	Sales p.e.		VA p.e.		MFP	
Variables	Single equation	Basic system	Single equation	Basic system	Single equation	Basic system
Competition	-0.00627	0.0457**	-0.0172	0.105***	-0.00132	0.00177
	(0.0176)	(0.0202)	(0.0342)	(0.0406)	(0.00287)	(0.00328)
Innovation	0.0344**	0.505***	0.0349	0.678***	0.000523	0.0181**
	(0.0137)	(0.0605)	(0.024)	(0.0995)	(0.00192)	(0.00788)
Education (univ)	0.00557***	0.00423***	0.00517***	0.00344**	0.000301***	0.000265**
	(0.000617)	(0.000686)	(0.00136)	(0.00159)	(0.00011)	(0.000127)
Exports (indirect)	0.121	-0.0893	0.185	-0.149	0.00227	-0.00684
	(0.0791)	(0.0931)	(0.129)	(0.154)	(0.00982)	(0.0116)
Exports (direct)	0.297***	0.122**	0.301***	0.0666	0.0138**	0.0101
	(0.0512)	(0.0616)	(0.0827)	(0.104)	(0.00645)	(0.00786)
Size med	0.108***	-0.0038	0.0727	0.0628	-0.00944*	-0.009
	(0.0357)	(0.0409)	(0.0683)	(0.0792)	(0.00561)	(0.0064)
Size large	0.234***	0.0503	0.0265	-0.261**	-0.0280***	-0.0355***
	(0.0567)	(0.0675)	(0.104)	(0.129)	(0.00842)	(0.0103)
Age (ln)	-0.00945	-0.0402	0.0516	0.0435	0.00234	0.00294
	(0.0254)	(0.0284)	(0.0483)	(0.0564)	(0.0039)	(0.00451)
Foreign owned	0.431***	0.372***	0.611***	0.703***	0.0237**	0.0290**
	(0.0691)	(0.0777)	(0.13)	(0.151)	(0.0105)	(0.0116)
State owned	-0.205	-0.125	-0.336	-0.225	0.000637	0.00424
	(0.132)	(0.143)	(0.225)	(0.253)	(0.0213)	(0.0233)
GDP pc (ln)	1.137***	0.993***	1.114***	0.824***	0.00615	-0.000734
	(0.0502)	(0.0157)	(0.0485)	(0.0529)	(0.00497)	(0.00631)
Pop. Density	0.146	0.393***	0.188**	0.409***	0.00516	0.00714
	(0.114)	(0.032)	(0.0811)	(0.0833)	(0.00647)	(0.00925)
Appropriability	-0.0466*	-0.143***	-0.213***	-0.173*	-0.000452	-0.00102
	(0.028)	(0.0331)	(0.0787)	(0.0911)	(0.00869)	(0.0115)
Observations	6,500	5,585	1,572	1,309	954	769
R-squared	0.987	0.984	0.986	0.979	0.526	0.469

NB: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.
Source: EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS).

Table A7: The productivity estimates – CAC

	Sales p.e.		VA p.e.		MFP	
Variables	Single equation	Basic system	Single equation	Basic system	Single equation	Basic system
Competition	0.0436	0.0246	0.0627	-0.0162	0.00329	0.00294
	(0.0357)	(0.0371)	(0.0634)	(0.0625)	(0.00725)	(0.00671)
Innovation	0.0845***	0.0973	0.0397	-0.0198	0.000394	0.0204
	(0.0316)	(0.12)	(0.0511)	(0.213)	(0.00501)	(0.0225)
Education (univ)	0.00383***	0.00391***	0.00134	0.00218	-0.000411	-0.000302
	(0.00116)	(0.00119)	(0.00241)	(0.00233)	(0.000253)	(0.000258)
Exports (indirect)	0.253	0.252	-0.233	-0.0873	0.01	0.00384
	(0.21)	(0.217)	(0.349)	(0.393)	(0.0512)	(0.0493)
Exports (direct)	0.173	0.124	0.199	0.206	-0.00495	-0.012
	(0.165)	(0.171)	(0.257)	(0.247)	(0.0296)	(0.0311)
Size med	0.0482	0.0404	0.0481	0.0606	-0.0372***	-0.0363***
	(0.0731)	(0.0739)	(0.132)	(0.132)	(0.0136)	(0.0133)
Size large	0.144	0.144	0.252	0.244	-0.00933	-0.0143
	(0.117)	(0.12)	(0.197)	(0.195)	(0.0227)	(0.0231)
Age (ln)	-0.0638	-0.0682	0.0558	0.0426	0.00474	0.00264
	(0.0524)	(0.052)	(0.0954)	(0.0916)	(0.0115)	(0.0112)
Foreign owned	0.324**	0.282**	0.116	0.126	0.00435	0.00167
	(0.14)	(0.141)	(0.226)	(0.224)	(0.0249)	(0.0244)
State owned	-0.128	-0.265	0.221	-0.676	0.0261	0.0473
	(0.376)	(0.394)	(0.575)	(0.652)	(0.0727)	(0.0737)
GDP pc (ln)	1.132***	1.134***	1.269***	1.319***	0.00718	0.00666
	(0.0268)	(0.0372)	(0.0773)	(0.0897)	(0.0114)	(0.011)
Pop. Density	0.170***	0.179***	0.143**	0.142*	-0.00299	0.00434
	(0.0368)	(0.0386)	(0.0653)	(0.0778)	(0.0099)	(0.0123)
Appropriability	-0.155**	-0.151**	-0.456***	-0.471***	0.00124	-0.00613
	(0.0616)	(0.0607)	(0.123)	(0.117)	(0.0204)	(0.0212)
Observations	1,365	1,330	361	346	137	131
R-squared	0.986	0.986	0.988	0.988	0.701	0.642

NB: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Source: EBRD-WB *Business Environment and Enterprise Performance Survey* (BEEPS).