LUND UNIVERSITY
School of Economics and Management

Labor productivity impact of Internet infrastructure: evidence from a panel data

by

Nikita Napalkov

August 2017

Department of Economics

Supervisor: Klas Fregert

Abstract

The study aims to shed more light on the relationship between Internet penetration and labor productivity by analyzing the aggregate cross-country panel data available for a wide range of countries in the period of 2001-2015. An overview of hypothetical mechanisms behind the communication technologies' impact on productivity is provided, while the Augmented Solow model is used as a theoretic framework which motivates the choice of variables for empirical analysis. The estimates are obtained by using a selection of econometric estimators: fixed-effects OLS, mean-group common correlated effects and "Difference GMM" for additional robustness. Extensive empirical analysis is performed in order to account for certain well-known factors which can cause a bias in the estimates. When using a "penetration index" comprising a few dimensions of a country's Internet development as a main variable of interest, the paper finds significant positive effect only in developed countries' sample. Difference GMM estimates, however, are not significant. Additional analysis suggests that there may be a more pronounced connection between mobile subscriptions and labor productivity (and more than just in the developed countries). The independent variables' estimates all have a theoretically expected signs. The results are fairly robust and allow a cautiously optimistic view on the relationship between Internet development and labor productivity growth. Nevertheless the extent of effect may sensitive to some unobserved country characteristic or industry.

Keywords: Internet penetration, productivity growth, Augmented Solow model, Common correlated effects, dynamic panel data.

Acknowledgements

I am very grateful to my supervisor Klas Fregert, Ph.D. for his valuable guidance during the whole process of work on the paper. I would also like to extend my thanks to Prof. Joakim Westerlund for his advice on the econometric matters which are essential part of this work.

Table of Contents

I. Introduction

It is hard to argue that the advances in the development of communication technologies since the end of the 20th century have been very significant. One of the most striking examples is the Internet penetration growth rates: in 2014 in the OECD on average 81% of the adult population had an Internet access and 75% used it every day while in 2002 only around 34% had the access (OECD, 2015). It is argued that communication technologies throughout the years transformed the way the economy is structured and have created new industries and products. As pointed out by Melody (2009) existence of "the new knowledge economy" hinges on the effective ways to create, interpret, store and exchange information. Therefore communication technologies are presumed to be of central importance for economic growth in the context of emerging knowledge economies.

Fostering economic growth and productivity efficiency is a cornerstone of many countries' national economic policies. As such it is necessary to answer the question regarding the role of the Internet, as a part of telecommunication technologies, in the recent economic development, in order to establish, if the countries need to more actively promote development of the infrastructure behind the Internet services in their countries. In fact in 2010 the European Commission initialized a special programme, called the Digital Agenda for Europe, aimed at increasing the Internet's quality and coverage in the Union. Set of goals defined by the Agenda implies spending up to 221 billion Euro on its implementation, since only a fraction of the population would be covered if only market mechanisms and incentives are used (Gruber et al., 2014).

Clearly the policymakers expect positive economic outcomes from investment in ICT (Information and Communication Technologies). It is thus interesting to observe that a question regarding the aggregate effect of ICT on the productivity is not one with a straightforward empirical answer. In 1987 Robert Solow stated famously that "you can see the computer age everywhere but in the productivity statistics" (Solow, 1987, p. 36). The quote became known as the "Solow Paradox" because empirical findings at the time indeed could not establish a significantly positive relationship between these two variables. Since then studies employing the growth accounting framework have argued that there is a positive impact of ICT on labor productivity in the US (van Ark et al., 2008). More recently researchers claimed finding a positive relationship between ICT and economic growth on the aggregate level. However the studies which use a more sophisticated empirical methodology and wider samples tend to come up with an effect on economic growth of smaller magnitude than initially suggested (see the Literature review section). Moreover an important question regarding the size of the impact is

whether it continues to be a relevant factor for labor productivity and not just economic growth in general, especially when the developed countries have already accumulated high stocks of ICT capital (in form of infrastructure). Is it possible that the effect of infrastructure per se is small and insignificant (see Evangelista et al., 2014) or even negative (Thompson & Garbacz, 2007) in this group of countries? Is there heterogeneity in the impact depending on country characteristics? Does the effect hold under scrutiny of different empirical methods? These are the questions which motivate an empirical study based on the panel data from a broad range of countries.

This research will argue that the effect of Internet development (penetration) on aggregate labor productivity is not necessarily positive, but rather ambiguous. It could dependent on the country sample, methodology used and possibly the industry concerned. The aim of the study is to measure the impact of broadband and mobile penetrations' growth on total labor productivity growth by means of empirical analysis of the cross-country panel data over the period from 2001 to 2015. The paper adopts the Augmented Solow model of growth, which motivates the choice of control variables, and analyzes the effect of interest using panel fixed-effects OLS, Common Correlated Effect Estimator and Difference GMM for robustness. These estimators are used to account for various measurement issues associated with unobserved heterogeneity and endogeneity.

The remainder of this paper is structured as follows: chapter 2 describes the possible mechanisms of ICT's impact on GDP and labor productivity while chapter 3 provides a brief summary of the previous literature findings regarding the economic effect of ICT in general and Internet in particular. Chapter 4 introduces the theoretical framework used in this research. Chapter 5 describes data at hand as well as difficulties associated with empirical analysis of the impact of ICT and motivates the design of econometric specification. Chapter 6 presents and discusses the main findings from the selected models which are followed by additional robustness tests. Chapter 7 summarizes the findings of this paper and provides a conclusion.

II. ICT and the economy: mechanisms of impact

In order to come up with an adequate econometric strategy for measuring the impact of the Internet penetration growth on labor productivity it is important to first describe the hypothetical theoretical mechanisms behind this relationship.

As suggested by Melody (2009) ICT in general is mostly used as an intermediate good, such that the value comes not from the technology or product itself but from the application of said product. The positive productivity impact is thus expected to be achieved by transformation of the way the business is conducted, since it becomes cheaper to generate and share information. Indeed, according to Leff (1984), decreasing communication costs resulting from the expansion of the appropriate infrastructure should lower the cost of acquiring new information and thus increase quantity of information readily available, which "can be expected to promote increased arbitrage and enhance market efficiency" (p. 261). Enhanced market efficiency implies that market prices shift to a more competitive level, which should reduce resource misallocation and quasi rents. Litan and Rivlin (2001) also argue that potential gains to productivity come as a result of reduced transaction costs, increased management efficiency and increasing competition, which pressures market agents to adapt cost-saving technologies. They highlight the fact that cost-savings may be especially high in information-intensive sectors of the economy (health-care, financial services, retail, public sector and so on). At the time the authors predicted "annual contribution [of the Internet] to productivity growth of 0.2-0.4 percent" over five years (Litan & Rivlin, 2001, p. 316).

In neoclassical growth model widely used in growth-accounting frameworks ICT contributes to aggregate productivity by either being one of the inputs in the production function and increasing productivity of industries using it, or by increasing total factor productivity (TFP) of ICT producing sectors (Draca et al., 2009). TFP is a portion of output growth which cannot be attributed to contributions of the inputs in the production function and it is also known as multifactor productivity. This variable is usually defined as the residual in the output growth equation and it accounts for gains in efficiency of use of other inputs (van Ark et al., 2008).

This formal measure (TFP/MFP) reflects the idea that telecommunication infrastructure may be regarded as a GPT – general purpose technology, which changes how economic activity is organized (Bloom et al., 2011). Most famous historic examples of GPT are steam engine and electricity. Aghion et al. (2014) provides 3 features defining a GPT: wide use in many sectors of the economy; tendency to underperform when first being invented / delivering productivity gains during the life cycle; ability to make it easier to create subsequent technologies. As such the channels of impact are both direct (factor productivity growth) and indirect. It is also important

to mention that if the Internet is a true GPT it could render older technologies obsolete in a Schumpeterian fashion, reducing employment in traditional industries thus initial impact of this technology may very well be negative in terms of employment and productivity. In the initial step restructuring of the businesses to adopt new technology induces costly initial R&D investment; these resources are naturally taken out of production (Aghion et al., 2014). In fact employment effects from such creative destruction in certain industries may adversely affect labor productivity since the measure used in this paper is GDP per persons employed. As such productivity growth technically depends on both changes in GDP and employment. As pointed out by Evangelista and Savona (2003) a direct labor-saving effect may be more severe in certain industries such that number of jobs created by ICT proliferation is smaller than those destroyed (e.g. labor intensive jobs in services).

Table 1 below summarizes a number of possible channels through which Internet infrastructure stock and labor productivity may be connected. Adopted in part from Czernich et al. (2012).

Table 1. Causal channels of Internet's impact on labor productivity (LP)

Directly	**As GPT**					**Higher chance of technological spillovers**
Demand for Internet infrastructure	Increased market efficiency of using industries (cost savings)	Lower entry barriers for startups in services	New business models (e.g. e-commerce) and products	Better job matching	Initial introduction of technology	Faster adoption of new technologies devised by others
↓	↓	↓	↓	↓	↓	↓
Increased economic output in construction sector	Lower prices, higher demand	Increased employment + increased GDP	Employment effect ambiguous (depends on industry)	↓	Capital taken out of production to invest in R&D	Higher pace of spread of any GPT
↓	↓	↓	↓	↓	↓	↓
GDP growth	GDP growth	LP growth	LP effect is ambiguous	LP growth	Reduction in LP in the period just after introduction	GDP and LP growth

Evangelista et al. (2014) is one of the first to also suggest productivity impact from the side of the individuals' increasing competency resulting from more active Internet usage. Their arguments for this are summarized in the table below.

Table 1 (continued). Causal channels of Internet's impact on labor productivity (LP)

Individual side effects (depend on intensity of usage of telecommunication stock)			
Improvement in tech-competence of individuals using Internet (gaining advanced skills) ↓ Skilled labor input is complementarity necessary for productivity growth through improved processing of information in IT-using firms --- Firm's investment in IT + skilled labor ↓ Increasing firm's productivity, ec. growth	Prerequisite for working and studying from home (new work practices) ↓ Increased employment and competence especially in disadvantaged groups ↓ LP ambiguous	Easier to search for jobs ↓ Better job matching ↓ LP growth	Easier to find information and start new business ↓ Combined with lower entry barriers for services industry ↓ GDP growth

The causality however is not one-directional as the literature almost universally warns (and also claims to prove in some cases). Indeed, it is possible that ICT infrastructure/usage growth leads to increases in labor productivity (and GDP per capita as well). But it is also possible that individuals and firms in countries with higher GDP per capita also have higher ability to pay for ICT / invest in ICT, which would lead to a more rapid ICT growth. Moreover it may be that countries with higher GDP per capita tend to have stricter regulatory environments (higher chance of state intervention) which would also affect speed of growth of ICT (Czernich et al., 2012). In such cases endogeneity becomes an issue and more basic empirical approaches could not be used to claim causality even if significant correlation is robustly established. The only solutions to the endogeneity issue are instrumental variables techniques or natural experiments. More detailed discussion regarding the choice of estimation method is given in chapter V.

Overall this chapter highlights the fact that there are multiple links between ICT in general (and therefore the Internet as well) and labor productivity growth, most of which are expected to be positive. However certain aspects of ICT as a general purpose technology may partly offset the positive effect on the aggregate level. As the next chapter shows these suspicions have been indeed substantiated by some of the researchers.

III. Literature review

ICT infrastructure is a part of public infrastructure, and as such it is worth mentioning that the early time-series study by Aschauer (1989) suggests a large and significant effect of public infrastructure investment on total factor productivity growth. However the major impact was shown to decrease dramatically once more robust econometric specifications are employed (Roller & Waverman, 2001). Subsequent studies generally highlight a positive link between telecommunication/Internet infrastructure and GDP growth, but some recent studies suggest a much more modest or even negative impact on economic growth and labor productivity (especially in the developed countries). Differences in intensity of adoption of new technologies and adoption lag are shown by Comin and Mestieri (2014) to account for an absence of convergence in GDP per capita between developed countries and the rest of the world, thus indicating that technology adoption is arguably the main reason for the systematic income differences.

Table 2 below summarizes main findings in the empirical literature concerning the effect of telecommunication and Internet growth on the economic development. These studies employ different theoretic frameworks and empirical methods and have been selected on the basis of using mostly time-series cross-country aggregate data for the analysis which is close to the setup of this paper.

Table 2 Summary of the main results found in the literature

Sample	*Period*	*Theoretic / Empirical model*	*Results*
Leff (1984) - Externalities, information costs and social benefit-cost analysis for economic development: An example from telecommunications			
LDCs		Descriptive social benefit-cost analysis	Communication investment projects in LDCs can influence economic development through numerous channels: lower transaction costs, reduced uncertainty improve factor allocation and resource mobilization. Telecommunication projects can provide vast external economies.
Cronin et al. (1991) – Telecommunications infrastructure and economic growth. An analysis of causality			
USA	1958-1988	Causality tests: Granger test, Sims and Modified Sims tests.	There is a bi-directional causality link between GNP and telecommunications investment. The result is significant at 10 percent level.
Madden & Savage (1998) – CEE telecommunications investment and economic growth			
27 transitional economies in CEE	1990-1995	Cross-country economic growth model OLS for 11 economies.	Telecommunication investment is positively associated with GDP growth at 5% significance level. Changes in growth are two-way casually connected to changes in telecommunications investment.

Röller & Waverman (2001) – Telecommunications infrastructure and economic development: A simultaneous approach			
21 OECD countries	1970-1990	Structural model that endogenizes telecomm. investment and economic growth. Micro-model of supply and demand is specified and jointly estimated with the macroeconomic production equation. Non-linear GMM is used.	Significant and positive relationship between telecommunications stock and GDP growth is established. Point estimates are reduced significantly when allowing country fixed-effects. In that case 1 percent increase in telecommunications penetration increases economic growth by circa 0.045 percent. One third of GDP growth in OECD in 20 years may be due to telecom development!
Datta & Agarwal (2004) – Telecommunications and economic growth: a panel data approach			
22 OECD countries	1980-1992	Cross-country growth framework of Barro (1991). Dynamic fixed-effects panel model is applied.	Stock of telecommunications access lines is significantly positively correlated with GDP growth. Size of the effect appears to be weakly inversely related to its prior level (negative squared term).
Thompson & Garbacz (2007)- Mobile, fixed line and internet effects on global productive efficiency			
93 countries	1995-2003	Stochastic-frontier production function approach	Higher mobile, telephone or Internet penetration decreases productive inefficiency. However developed countries as opposed to other country groupings show significant response to changes only in Internet penetration (the effect is negative)! Authors argue that this is due to the countries already operating near their productive frontier.
van Ark et al. (2008) - The productivity gap between Europe and the United states: trends and causes			
USA and the EU	1950-2006	Neoclassical growth accounting framework	Multifactor productivity growth slowdown is evident in the EU in 1995-2004 compared to the US. MPG can be a result of changes induced by ICT development. Contribution of this factor to the economic growth in the EU together with investment in ICT and changes in labor composition declined by 0.5 percent whereas in the US it increased. Thus productivity divergence can be attributed to slower emergence of the knowledge economy in the European countries.
Koutroumpis (2009) – The economic impact of broadband on growth: a simultaneous approach			
22 OECD countries	2002-2007	Structural econometric model in the spirit of Roller & Waverman. Limited information estimation (IV) and non-linear 3SLS GMM.	Increase in broadband penetration and use is significantly positively associated with GDP growth (1 percent increase = 0.023 percent growth). 0.40 percent of annual ec. growth in OECD countries can be attributed to growth of BB penetration.

Czernich et al. (2012) – Broadband infrastructure and economic growth			
25 OECD countries	1996-2007	Endogenous growth theory framework. IV regression: broadband penetration is instrumented. The 1st stage is based on the diffusion of the existing telephony and cable networks that predict BB penetration.	Significant and positive causal effect of broadband penetration on GDP per capita is found. 10 percent increase in instrumented penetration accounts for 0.9-1.5 percent GDP per capita growth. Statistically significant positive effect is only apparent once 10% threshold level of BB penetration is passed, while reaching threshold beyond that has no additional effect.
Evangelista et al. (2014) – The economic impact of digital technologies in Europe			
27 EU countries	2004-2008	3 sets of equations measuring the impact of "access, usage, empowerment" on labor productivity, GDPpc and employment. Pooled GLS and Areallo-Bond Difference GMM estimators.	3 separate dimensions of digitalization affect macro-economic variables differently. Access dimension (index formed by weighted sum of infrastructure and price components) has no significant effect on labor productivity, GDPpc or employment. Internet usage index is positively associated with labor productivity only when lagged for one period.
Gruber et al. (2014) – Broadband access in the EU: An assessment of future economic benefits			
27 EU countries	2005-2011	Endogenous growth theory. Structural econometric model in the spirit of Roller & Waverman.	The use of broadband connection is estimated to have contributed 1.36 percent to GDP annually. Some evidence of growth impact from a speed of the Internet connection. The hypothesis of threshold value for BB coverage (15%) is confirmed – in countries with values above this level the effect of broadband is higher.

The summarized studies suggest that on the aggregate level studies employing cross-country panel data tend to come up with positive effect of telecommunication or Internet variables on GDP / GDP per capita and mixed results for labor productivity. However at the industry-level there are fewer significant results, which, as argued by Stiroh (2002), may be either due to possibility of no effect of ICT, too much aggregation or due to model misspecification.

IV. Theoretical Framework

Modern growth theory is based largely on the work of Solow (1956) in which he uses the neoclassical production function and assumes exogenous technological change to come up with a model in which rate of saving and growth of labor force are the main determinants of GDP per capita steady-state growth paths. However the crucial property of the model is decreasing returns to capital investment which means that without technological change the growth would dwindle (Aghion & Howitt, 1998). Neoclassical growth theory also suggests that there should be a catching up effect, such that countries with initially low levels of GDP per capita should grow faster, so convergence in GDP growth paths is expected (Mankiw et al., 1992). However the lack of empirical support to these predictions caused a number of researchers to come up with alternative growth models, the most popular of which is endogenous growth model by Romer (1990). This branch of growth literature incorporates knowledge into the production process arguing that technological change is not exogenous but happens because of people's response to market incentives (endogenous). This model allows for intentional investment in R&D, whereas larger markets create more incentives for research, but population size is not the right measure of market size. Romer (1990) argues that it is human capital, proxied by some measure of formal training, which drives investment in research. This ultimately means that "growth rate is increasing in the stock of human capital" (p. 73). Mankiw et al. (1992) on the other hand suggest that the original Solow model is consistent with empirical evidence when growth equation is also augmented by human capital accumulation. The "augmented Solow model" is adopted in this paper to derive how ICT is connected with productivity growth.

In neoclassical production function inputs of capital, labor and technology provide the sources of growth. Due to diminishing returns to accumulation of capital, exogenous technological advances lie at the heart of economic growth. Thus there is a reason to expect that such technology as the Internet being a part of technological process has certain effect on productivity. The starting point for the theoretical model used in this paper is the Cobb-Douglas production function of the following form:

$$Y_t = (K_t)^a (A_t L_t)^{1-a} \qquad 0 < a < 1$$

Where a is capital investment's share in income, Y is output, K – capital, L – labor, A is the level of technology. In the standard Solow model the last two are assumed to grow exogenously at rates n and g respectively. $A_t L_t$ grows at rate $n+g$.

$$L_t = L_0 e^{nt};\ A_t = A_0 e^{gt}$$

The constant fraction of output *(s)* is invested, which ultimately leads to the steady-state GDP per capita equation derived from the production function above:

$$\ln\frac{Yt}{Lt} = \ln A(0) + gt + \frac{a}{1-a}\ln(s) - \frac{a}{1-a}\ln(n+g+\delta)$$

Where δ is rate of depreciation. Mankiw et al. (1992) mentions that since a has a fixed value of about one third, in the standard Solow model elasticity of GDP per capita w.r.t. investment in capital *(s)* should be around 0.5 and elasticity w.r.t. $n + g + \delta$ around -0.5. They then develop the so-called augmented Solow model which expands the original by including human capital *(H)* in the production function:

$$Y_t = (K_t)^a (H_t)^\beta (A_t L_t)^{1-a-\beta} \qquad a + \beta < 1$$

Which eventually leads to the following steady-state GDP per capita function:

$$\ln\frac{Yt}{Lt} = \ln A(0) + gt + \frac{a}{1-a-\beta}\ln(s_k) + \frac{\beta}{1-a-\beta}\ln(s_h) - \frac{a+\beta}{1-a-\beta}\ln(n+g+\delta)$$

Where s_k is a fraction of income invested in physical capital and s_h – fraction invested in human capital. GDP per capita now depends on growth of population, accumulation of both physical and human capital. What is left is to accommodate for ICT growth in the model. Internet infrastructure as mentioned earlier facilitates distribution of information and ideas, thus in the formal model it would affect the technology growth parameter g. Originally Mankiw et al. (1992) assume that both g and δ are constant across countries, this implies that advancement of knowledge has the same pace across countries and only the initial resource endowments (a_0) differ:

$$\ln A(0) = a_o + \epsilon$$

However for the analysis of impact of ICT on productivity growth it is important to allow for g to differ across countries thus g gains a subscript i. Czernich et al. (2012) suggest that diffusion of Internet is connected with technological growth parameter in the following way:

$$g_i = g_c + a_1 B_i$$

Where B_i is the broadband Internet penetration rate and g_c is a constant. In this paper in addition to broadband penetration I am interested in mobile-cellular penetration rates. Both would form an index that in my specification also varies over time and is reflected in variable $B_{it.}$ With this the final empirical specification for a single country looks like:

$$\ln\frac{Y_{it}}{L_{it}} = a_i + g_c t + \sum_{t=1}^{t} a_1 B_{it} + \frac{a}{1-a-\beta}\ln(s_{kit}) + \frac{\beta}{1-a-\beta}\ln(s_{hit}) - \frac{a+\beta}{1-a-\beta}\ln(n_{it} + [g_c + a_1 B_{it}] + \delta) + \epsilon_{it}$$

Since the technology growth parameter is now present twice in the equation, exact predictions about the effect of change in the *Internet penetration index* B_{it} is impossible, since its accumulation would directly increase GDP per capita value, but at the same time would "pull" some investment from physical and human capital in the same way that growth in population does.

Moreover due to the fact that $n + g + \delta$ enter the equation under one coefficient, it becomes non-trivial to disentangle influence of population growth from an effect of technological growth analytically, unless assumptions about some of the values are made. Originally Mankiw et al. (1992) assume that $g + \delta = 0.05$ since "In U.S. data the capital consumption allowance is about 10 percent of GNP, and the capital-output ratio is about three, which implies that δ is about 0.03…" (p. 413).

Since the interest lies in finding the effect on productivity growth I take first differences, which results in:

$$\Delta\ln\frac{Y_{it}}{L_{it}} = g_c + a_1 B_{it} + \frac{a}{1-a-\beta}\Delta\ln(s_{kit}) + \frac{\beta}{1-a-\beta}\Delta\ln(s_{hit}) - \frac{a+\beta}{1-a-\beta}\Delta\ln(n_{it} + [g_c + a_1 B_{it}] + \delta) + \epsilon_{it}$$

In the standard Solow model elasticity of GDP per capita w.r.t. investment in capital *(s)* should be around 0.5 and elasticity w.r.t. $n + g + \delta$ around -0.5. The augmented model in turn implies that if shares of physical and human capital in income are around 1/3 then coefficient on $\ln(s_k)$ on average should be 1 while coefficient on $\ln(n + g + \delta)$ should be -2. So the presence of human capital increases the effect of accumulation of physical capital (Mankiw et al., 1992). In short, in difference terms we should expect positive coefficient on s_k and negative on n, while the coefficient for of s_h and B are theoretically ambiguous.

The augmented Solow model demonstrates how GDP per capita of working-age population is theoretically connected with capital investment, investment in human capital, population and technological growth rates. Moreover it motivates the choice of variables for empirical specification and gives a set of predictions regarding the expected slopes of variables. The next chapter presents such specification and provides an overview of the data and empirical method used.

V. Model specification

Measuring the impact of Internet and communication technologies on economic growth is a non-trivial pursuit due to the number of methodological and empirical difficulties. As evident from the previous chapters these difficulties arise due to the high number of channels through which Internet development can affect GDP/employment ratio and because of the reverse causality and spurious relationship issues.

V.1 Empirical methods and explanatory variables

The research is based on the panel data available for a large sample of countries and thus it becomes possible to employ three estimators to reduce the aforementioned biases. The first one is the standard fixed-effects OLS regression. Fixed-effects (γ_{it}) essentially mean estimating only within-country changes in variables which by construction would eliminate the unobservable country-specific effect (the same way differencing does in case of two time periods). If one additionally includes year dummies (λ_t) it then becomes possible to account for both country-specific heterogeneity and year-specific effects (common for all countries in the sample for a given time-period).

$$\Delta\ln(LPROD)_{it} = C_i + \beta_1 \Delta\ln(PEN)_{it} + \beta_2 \Delta\ln(I/GDP)_{it} + \beta_3 \Delta\ln(HC)_{it} + \beta_4 \Delta\ln(N)_{it} + \gamma_{it} + \lambda_t + \varepsilon_{it}$$

The drawback of FE OLS is that it becomes impossible to estimate time-invariant variables or account for time-varying country specific heterogeneity. Accounting for time-varying heterogeneity is possible in the second estimator used in this work. It is a more recently developed estimator – Common Correlated Effects Mean-Group estimator (CCEMG) proposed by Pesaran (2006) and it is aimed at accounting for unobservable common factors which are allowed to have differential impact on each country. The author proves that by augmenting the country-specific equation by cross-sectional averages of all variables it becomes possible to have a consistent estimator with the desired property as $T, N \rightarrow \infty$. Since it is a more general model than FE there are some additional assumptions: common effects are distributed independently of individual errors; error terms are distributed independently for *i, j, t*; slope coefficients follow random coefficient model and some more available in the paper (Pesaran, 2006). In practice this means that instead of year-dummies the regression equation includes cross-sectional means of both dependent and independent variables. Mean-group variation means that the coefficients are calculated separately for each country and then averaged across them. The Stata code for CCEMG including more options as well as the cross sectional dependence test was developed by Ditzen (2016).

$$\Delta\ln(\mathrm{LPROD})_{it} = C_i + \beta_1\Delta\ln(\mathrm{PEN})_{it} + \beta_2\Delta\ln(\mathrm{I/GDP})_{it} + \beta_3\Delta\ln(\mathrm{HC})_{it} + \beta_4\Delta\ln(\mathrm{N})_{it} + \delta_0\overline{\Delta\ln(\mathrm{LPROD})_t} + \delta_1\overline{\Delta\ln(\mathrm{PEN})_t} + \delta_2\overline{\Delta\ln(\mathrm{I/GDP})_t} + \delta_3\overline{\Delta\ln(\mathrm{HC})_t} + \delta_4\overline{\Delta\ln(\mathrm{N})_t} + \varepsilon_{it}$$

Still arguably the most often described issue with measuring the impact of technological change on growth is the issue of endogeneity. As previously discussed this can be caused if there is reverse causality (and in our case bidirectional relationship is indeed suspected). As such this issue would hinder arguing about the causality direction when using some type of OLS model (for OLS to work there should be no correlation between independent variables and the error term[1]). The best solution to this particular issue is a natural experiment or good instrumental variables, but these are not readily available due to data limitations. The next best technique is a group of estimators known as dynamic panel data estimators, specifically the Difference generalized method of moments (GMM) popularized by Areallo and Bond (1991). The idea behind the estimator is to instrument changes in variables by previous levels of variables (two and more periods back). This way the endogenous variables are instrumented by their lagged values which are not correlated with the current error term.

While sounding like a solid solution on paper this estimator as any other has its own weaknesses. As Roodman (2009) puts it – Difference GMM is more suitable for panels with large N and small T; linear functional relationship and "heteroskedasticity and autocorrelation within individuals but not across them" (p. 1). To satisfy the last condition year dummies are included. However if series displays random walk properties past levels of dependent variable may convey little information about future changes. The issue called weak instrumentation set is inherent to GMM estimators. The special Hansen test is then used to check for joint validity of instruments. The instrument count however is quadratic in number of time periods which is problematic: Bowsher (2002) shows that as instrument count rises the Hansen test is weakened. Finite sample may not have enough information to estimate large matrix of instruments (this paper's case). Of importance is the fact that consistency is still not compromised, but standard errors are not efficient anymore (Roodman, 2009). This paper adopt Difference GMM in addition to OLS estimators, the idea being that it would give a more conservative estimate of the relationship between Internet penetration and labor productivity growth.

[1] Such correlation will arise in case of simultaneity since if *Yit* determines *Xit* and *Xit* determines *Yit* simultaneously then *Xit* would be correlated with the error term (*Eit*). This is easy to see by writing out structural equations. In such case the independence assumption required for unbiasedness of OLS would be violated.

Turning to the choice of independent variables it should be first of all noted that in the empirical growth literature "as many growth determinants have been proposed as there are countries for which data are available" (Durlauf et al., 2004). For example, the famous empirical analysis by Barro (1991) suggests that economic growth empirically is positively connected with the initial human capital/GDP per capita relationship, measures of political stability and physical investment. It is also inversely related to the share of government consumption in GDP and initial level of real GDP per capita. In this work I restricted the amount of explanatory variables to these directly emerging from the Augmented Solow model to increase the number of observations available, avoid larger gaps in data (important for FE and CCE estimators) and due to the instrument proliferation problem (relevant in case of GMM models). The list of all variables is given below in the Table 3. My variable of interest in the main specifications is defined as Internet penetration index (PEN), which is weighted sum of three normalized variables which act as proxies for the development of Internet infrastructure in a country. Broadband subscribers' share of population comprises 60% of the index, mobile-cellular subscribers' share – 30% and number of secure servers per million of individuals – 10%. The weighting would mean that coefficient for PEN would reflect the influence of broadband Internet infrastructure the most. As such it is supposed to reflect the higher importance of fastest type of Internet connection but still not discard mobile Internet use which becomes more and more popular. Such weighting order is also used by the newest DESI (Digital Economy and Society Index) developed by the EC, the difference being the larger set of variables which Eurostat gathers (however only available for European countries and for much fewer years than variables used in this paper). Human capital variable is also a problematic one to capture since proxies used in the literature often have limited coverage and large gaps. For this reason there are three proxies alternatively used in this paper.

Table 3 Variables used in the estimations

Name	**Description**
LPROD	GDP per person employed (in 2015 US dollars)
I/GDP	Gross-fixed capital formation (share of GDP) – *proxy for capital input*
N	Labor force (Employed + seeking employment) growth rate
SCHOOL	Share of working-age population (aged 15-19) in secondary school
TERT_Enrl	Tertiary school enrollment rate
TERT_Ed	Share of population aged 26 to 64 with tertiary education
PEN:	Composite index of Internet penetration – *proxy for infrastructure*
└ 0.6**BB_sub*	Broadband subscribers (of total population), normalized
└ 0.3**MB_sub*	Mobile-cellular network subscribers (of total population), normalized
└ 0.1**SERV*	Number of secure servers per million of individuals, normalized

Additionally of interest are the slow adjustments of the dependent variable, which reflect transition between steady-state levels of labor productivity. This means that there is a reason to include lagged dependent variable (LDV) as one of the regressors. Unfortunately including LDV in fixed-effects estimation would bias the coefficients of the independent variables downwards as shown by Nickell (1981), however this is exactly the type of equation (dynamic panel) GMM would be most useful for.

V.2 Data

The panel data for which Internet penetration variables could be obtained comprises observations from 121 countries ranging from the least developed countries to most developed ones according to the UN classification. However if there was only one year where all the variables were available, the country had to be excluded due to differencing, which also limits a number of total available observations a bit. Individual variable series were obtained from a wide range of sources, these are provided in the Appendix A. The sample time dimension was restricted by data availability of the main variables of interest – broadband/mobile subscribers and secure servers count. The final database covers the years 2001 to 2015 and has 1815 available observations on the dependent variable. The summary statistics is presented in Table 4.

Table 4 Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
LPROD	1815	47605.12	39820.04	1564.81	198533.7
I/GDP	1740	.2229694	.0625937	.0172893	.5770911
N	1694	.0176792	.0229032	-.1104785	.2062265
SCH	1545	.1022894	.0298984	.0143583	.1842923
TER_Enr	1336	.4164845	.2643936	.0037384	1.138718
TER_Ed	523	.2839062	.1047817	.0639777	.5517366
PEN:	1594	.1725248	.1521619	.0000534	.6600128
└ BB_sub	1629	.0947177	.114117	0	.4580223
└ MB_sub	1800	.7727293	.4900812	0	3.104378
└ SERV	1739	209.5308	473.1159	.007215	3406.738

There were different proxies for the chosen variables potentially usable for analysis (e.g. share of households with Internet access at home; share of population aged 25+ with a completed bachelor's degree etc.), but they had to be excluded since most of them were restricted to particular groups of countries, limited number of years or both. Natural logarithms of all the variables except growth rate of labor force are used for better linear fit and easier interpretation. Correlations table is presented in the Appendix A. For robustness checks lagged differences of

independent variables would also be used, however this could mean losing further time periods of data. Additionally later I consider replacing the cumulative PEN index by broadband and mobile subscriber's shares separately as well as use share of population "using" the Internet as a replacement.

It is important to note that if residual series are non-stationary, it is possible to find significant relationship even between two unrelated non-stationary variables – this is called "spurious regression" (Enders, 2015). The Augmented Dickey-Fuller test fails to reject H0 of non-stationarity for all variables in levels (see Appendix B). Differencing of the macroeconomic series is expected to turn series into stationary. In my case ADF test after first-differencing rejects non-stationarity for labor productivity (but only weakly) and HC variables. In fact labor productivity and capital investment are first-difference stationary over long run, but my sample period is only 15 years, so the cycles are less evident. Graph of the labor productivity series in Appendix B implies that growth rates of productivity did not return to the pre-2008 rates in the recent years, which is why the stationarity is undermined. Labor force growth and PEN index display downward-sloping trends in first-differences. Second differences completely eliminate non-stationarity in every variable. This is desirable since Difference GMM estimation would use differences of my first-differenced variables. By using first-differenced variables I expect to have R-squared and t-statistics which are not (significantly) biased. Addition of year-effects should help eliminating additional effects having to do with common external time-specific shocks influencing all countries in the sample.

VI. Main results and discussion

Results for the main specifications using FE OLS and CCEMG estimators are presented in Table 5 and Table 6 accordingly, for GMM – in Table 7. Each table reports a number of sub-specifications that differ in variables included and sample of countries analyzed. Consistency and efficiency increases with number of observations, but it is still important to have country groupings analyzed separately due to the possibility of heterogeneity in the effect depending on some common property of countries in a sample (heterogeneous slope coefficients). Columns (1)–(4) report results only for the group of developed countries according to the EU classification (acronym: DVLPD), first specification omits HC completely and the later include different proxies – one at a time. Specification (4) has fewer countries due to the fact that share of population with tertiary education was only available for OECD countries plus a few more developed ones. Columns (5)–(7) report the coefficients for a group of developing countries including the least developed ones which had too few observations to be analyzed separately (DEV). Finally columns (8) and (9) present results for the whole sample for variables available (ALL).

VI.1 FE OLS results

By looking at the results obtained from FE OLS it is apparent that all of the significant estimates for the variables have the expected signs. From the theory we know that capital investment and human capital investment should have positive effect on economic growth, while growth rate of labor force should have the opposite sign. Our main variable of interest – change in PEN index has a significant and positive correlation with labor productivity growth but only in the developed countries' sample. The effect appears to be larger in specification (1) where human capital is not controlled for. Since natural logarithms are on both sides of the equation, coefficients in FE OLS model can be seen as elasticities, so for example column (2) suggests that 1% increase in the Internet penetration *growth rate* is associated with 0.021% increase in GDP per employed *growth rate.* Human capital proxies with exception for positive effect of tertiary school enrollment rate in developed countries all have no statistically significant estimates.

Point estimates for Internet penetration growth, while positive, become insignificant in the developing countries' case, but it seems that the fit becomes quite poor as indicated by lower R^2. The reason may be the limited data available for developing countries, specifically the human capital data for the developing countries has many gaps leaving many observations to be lost in columns (6)-(7).

Table 5 Results for Fixed-effects OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Δln(LPROD)	DVLPD	DVLPD	DVLPD	OECD+	DEV	DEV	DEV	ALL	ALL
Δln(PEN)	0.0389***	0.0212**	0.0234**	0.0170	0.00884	0.0143	-0.0116	0.0157	0.0163
	(0.0115)	(0.00891)	(0.0110)	(0.0131)	(0.0153)	(0.0164)	(0.0146)	(0.0124)	(0.0131)
Δln(I/GDP)	0.0482*	0.0690***	0.0678***	0.0443	0.0553**	0.0572***	0.0529*	0.0560**	0.0612***
	(0.0254)	(0.0215)	(0.0230)	(0.0313)	(0.0268)	(0.0103)	(0.0307)	(0.0229)	(0.00919)
N [LF growth]	-0.545***	-0.538***	-0.526***	-0.496***	-0.509***	-0.570***	0.558***	-0.524***	-0.566***
	(0.0896)	(0.0863)	(0.0844)	(0.127)	(0.102)	(0.134)	(0.132)	(0.0772)	(0.0909)
Δln(SCH)		0.0218				0.0130			0.00870
		(0.0197)				(0.0286)			(0.0187)
Δln(Ter_Enr)			0.0503**				0.0129		
			(0.0243)				(0.0150)		
Δln(Ter_ed)				-0.0380					
				(0.0299)					
Constant	0.0150***	0.00237	-0.00118	0.0188***	0.0319***	0.0238***	0.0404**	0.0291***	0.0147***
	(0.00550)	(0.00476)	(0.00714)	(0.00435)	(0.0115)	(0.00384)	(0.0159)	(0.00738)	(0.00276)
Year effects significant	Y	Y	Y	Y	N	Y	N	N	Y
Obs	556	519	469	389	843	697	521	1,399	1,216
R-squared	0.463	0.483	0.502	0.500	0.166	0.180	0.156	0.227	0.252
Countries	42	41	40	31	76	68	64	118	109

Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1)

Another possibility is that there are some common effects that affect all countries in the sample in the given time period, but affect them differently (this required a different model to capture). Human capital growth rates appear not to be significantly correlated with labor productivity growth rate, nevertheless capital investment and LF growth variables still have the expected coefficients. The same applies to the full sample in columns (8)-(9). Overall the coefficients for capital investment growth and growth rate of labor force are robustly significant across FE specifications, but evidence of positive connection between Internet and labor productivity is only observed in the sample of developed countries.

VI.2 CCEMG estimator results

I now turn to Common correlated effects mean group estimator developed by Pesaran (2006) results for which are reported in Table 6. In general there is no reason to believe that regressors are identically and independently distributed across countries in the samples, there may be unobserved common factors that are correlated with regressors. The main difference from FE OLS with year dummies is that now we control for common effects having "differential impacts on individual units, while at the same time allowing them to exhibit an arbitrary degree of correlation among themselves and with the individual-specific

regressors" (Pesaran, 2006, p. 969). Country-specific errors can be serially correlated and heteroscedastic and country-specific variables do not need to be strictly i.i.d. Regressors however should be stationary and exogenous. Stationarity condition as mentioned earlier is assumed to be satisfied due to differencing; however endogeneity due to simultaneity could still be an issue (for now we have to assume exogeneity of the Internet penetration index).

Ideally it would have been possible to consistently estimate a dynamic panel (with lagged dependent variable) thanks to the extension of CCE by Chudik and Pesaran (2015), but unfortunately the data requirements for such estimation method are quite high. Since the estimation procedure requires adding lags of cross-sectional means, in this paper I end up having more variables than observations preventing me from such estimation by means of CCEMG.

Results for CCE estimator show that estimates have the expected signs, with exception for the case where all countries are in the sample. Moreover R^2 values have increased and suggest that around 70% of variation in labor productivity is explained by our variables and their cross-sectional averages. Point estimates for capital input growth effect are now higher than in case of FE OLS, while labor force growth rate coefficient is in the same frame albeit not significant for the developed countries. Human capital growth is still not significant determinant for labor productivity growth, moreover it becomes impossible to estimate specification (7) due to too many variables.

Table 6 Results for CCEMG

	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
Δln(LPROD)	DVLPD	DVLPD	DVLPD	OECD+	DEV	DEV	ALL	ALL
Δln(PEN)	0.0258*	0.0220	0.0507**	0.0516***	0.0423**	0.103***	0.131	-0.0352
	(0.0142)	(0.0207)	(0.0218)	(0.0177)	(0.0196)	(0.0324)	(0.112)	(0.0378)
Δln(I/GDP)	0.105***	0.159***	0.207***	0.0744**	0.0266	0.0669	0.0280	0.0447
	(0.0263)	(0.0349)	(0.0616)	(0.0288)	(0.0287)	(0.0571)	(0.0232)	(0.0402)
N [LF growth]	-0.333	-0.197	-0.217	-0.486**	-1.846	-1.119*	0.498	-0.813**
	(0.231)	(0.222)	(0.399)	(0.199)	(1.277)	(0.532)	(0.982)	(0.411)
Δln(SCH)		0.105				0.0382		0.0975
		(0.0918)				(0.151)		(0.0961)
Δln(Ter_enr)			0.0417					
			(0.0822)					
Δln(Ter_ed)				0.00227				
				(0.0754)				
Constant	0.00559	0.00348	0.00759	0.00887**	0.0335	0.0304	-0.179	-0.0632
	(0.00349)	(0.00659)	(0.00385)	(0.00413)	(0.0608)	(0.0929)	(0.150)	(0.0565)
Observations	556	519	469	389	843	697	1,399	1,216
R-squared	0.719	0.767	0.791	0.712	0.668	0.720	0.655	0.796
Countries	42	41	40	31	76	68	118	109
F-value	1.678	0.877	0.640	0.631	0.777	0.0644	0.917	0.451
CD-test stat.	2.496	0.940	2.252	1.061	-0.104	3.150	0.216	3.008
p>CD stat.	0.0126	0.347	0.0246	0.289	0.918	0.00164	0.829	0.00263

Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1)

The PEN variable growth is now significantly associated with GDP per employed growth in both developed and developing sample, suggesting the effect close to the one reported in FE OLS table (0.02-0.05% faster growth rate of labor productivity resulting from 1% higher growth rate of Internet penetration). However there are certain concerns regarding the consistency of the estimator in some cases. If the cross-sectional means (of all variables, partialed out during estimation) do not take care of all dependence between countries, the error terms will contain cross-sectional dependence and will not be identically independently distributed any more, this would make OLS not consistent (Ditzen, 2016). Chudik and Pesaran (2015) developed a way to test for cross-sectional dependence called CD test. Under H0 errors are weakly dependent; the test statistic is asymptotically normally distributed. CD test postestimation suggests that cross-sectional dependence is not entirely eliminated in specifications (2), (4), (5) and (8), so the respective coefficients should be taken with skepticism. In addition, it is evident from the value of F-test that for the specifications (4), (6) and (9) the hypothesis of no joint significance for variables cannot be rejected (due to insignificant cross-sectional averages). This leaves us with two specifications that have the better diagnostic results – (1) and (3) of which the first one is preferred. These are results for the case of developed countries that confirm the previous findings but also suggest a larger effect of capital input: 1% increase in growth of gross fixed capital formation to GDP ratio leads to 0.1-0.2% rise in labor productivity growth rate (as opposed to 0.05-0.07% rise).

VI.3 Difference GMM and additional robustness tests

So far I have used OLS approach to estimate the panel data at hand and it is well know that for OLS to be unbiased and consistent in time-series setting the independent variable should not be correlated with the error terms past or present. This assumption can be violated if there are measurement errors, but this particular issue is outside of control of the researcher. Secondly omitted variables can lead to correlation, we tried to control for this by means of fixed-effects, year dummies and common correlated effects on top of the theoretically motivated regressors choice. But there are still more factors that could potentially lead to violation of strict exogeneity assumption, the main being the simultaneity between ICT variable and GDP suggested in the literature on ICT's effect on growth. This would be an issue even if the critical assumption that capital investment and labor force growth rate are independent of the error term holds. The solution used in absence of good instrumental variables is the generalized method of moments which also by design is most suitable for dynamic panels. The main idea and issues with the

estimator of choice – Difference GMM were discussed in the previous chapter and now we turn to the results obtained by this estimator reported in Table 7.

The solution to the endogeneity problem in Difference GMM lies in using lagged variables as instruments for their current values. The researcher has a choice of differenced variables that are deemed endogenous, predetermined or exogenous. The first group is instrumented by the second and following lags of its levels (to maximize the sample), predetermined variables (usually LDVs) are instrumented by the first and following lags. For the method to work it is of course critical that all instruments are orthogonal to the contemporaneous error term. This is where the so-called Hansen test comes into play – it is used to test joint validity of the instruments. Under H0 the instruments are jointly valid (exogenous), so we do not want to reject the null hypothesis here. But as Roodman (2009) reports there is an issue – in case of too many instruments endogenous variables can be "overfited"; on top of that the Hansen test statistic never rejects the H0 in such case, and a "telltale sign is a perfect Hansen statistic of 1.000" (p. 43). This is the reason for limiting lag lengths (to 6 and 3 for each specification) in my regressions: although fewer lags can mean lower efficiency, longer lags lead to too many instruments. Another reason for two lag lengths for each specification is to test whether Hansen test statistic varies a lot with lag lengths – if instruments are valid it should not.

I start with specification including all countries in the sample (1)-(6), since it becomes more important to have as many observations as possible to increase degrees of freedom. As mentioned earlier I assume one predetermined variable – lagged dependent variable, one endogenous variable – Internet PEN index, and the rest are treated as exogenous variables. For the estimation I use xtabond2 Stata procedure maintained by D. Roodman. Year dummies are also included in order for errors to be possibly correlated only within countries, but nor across them. This is important since Windmeijer (2005) finite-sample correction to standard errors (reducing downward bias) is used and the procedure assumes no such correlation in errors. Lastly AR(2) stands for Arellano-Bond test for serial correlation in the first-differenced residuals, which could also signify that lagged instruments are invalid (H0 – no serial correlation). Serial correlation in AR(1) is expected, but there should be no serial correlation in AR(2), meaning that for instruments to be valid H0 should *not* be rejected.

From the results of Difference GMM estimation it is evident that there are some dynamic effects in the data – coefficient for LDV is weakly significant in most of the specifications. Weak significance in differenced equation signals that the series is not very mean reverting but rather follows cycles (with random walk tendencies). The autoregressive nature of the labor productivity growth series is confirmed.

Table 7 Results for Difference GMM

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δln(LPROD)	ALL	ALL	ALL	ALL	ALL	ALL	DVLPD	DVLPD	DEV	DEV
L1. -//-	0.183**	0.182**	0.115*	0.129*	0.117	0.154	0.0996	0.106*	0.203*	0.191*
	(0.0872)	(0.0869)	(0.0625)	(0.0667)	(0.100)	(0.0953)	(0.0604)	(0.0614)	(0.106)	(0.103)
Δln(PEN)	0.0318	0.0136	0.0198	0.0228	-0.0101	-0.00276	0.0571*	0.0340	0.0104	-0.0152
	(0.0194)	(0.0216)	(0.0183)	(0.0221)	(0.0221)	(0.0319)	(0.0307)	(0.0356)	(0.0206)	(0.0239)
Δln(I/GDP)	0.0505**	0.0483**	0.0595***	0.0577***	0.0691***	0.0704***	0.0399	0.0465	0.0510**	0.0508*
	(0.0235)	(0.0233)	(0.0124)	(0.0128)	(0.0246)	(0.0240)	(0.0344)	(0.0383)	(0.0271)	(0.0267)
N [LF growth]	-0.643***	-0.609***	-0.573***	-0.544***	-0.564***	-0.576***	-0.560***	-0.572***	-0.680***	-0.654**
	(0.101)	(0.0934)	(0.0937)	(0.0936)	(0.108)	(0.116)	(0.125)	(0.130)	(0.134)	(0.125)
Δln(SCH)			-0.00985	-0.0172						
			(0.0206)	(0.0210)						
Δln(Ter_Enr)					0.0208	0.0191				
					(0.0149)	(0.0155)				
Obs	1,212	1,212	1,047	1,047	817	817	481	481	731	731
Countries	117	117	107	107	98	98	42	42	75	75
Instrum. (Z)	121	70	122	71	122	71	121	70	121	70
Lag lim. on Z	6	3	6	3	6	3	6	3	6	3
AR(2) p-value	0.334	0.324	0.349	0.288	0.912	0.701	0.434	0.535	0.230	0.246
Hansen test p-value	0.374	0.362	0.800	0.479	0.813	0.160	1	0.976	1	0.351
F-stat	10.95	11	10.28	9.653	9.721	9.270	24.06	18.21	6.28	6.67
F-stat: p-val.	0	0	0	0	0	0	0	0	0	0
Year-eff signif	N	Y	N	N	Y	N	Y	Y	N	N

Robust standard errors with small-sample correction in parentheses (*** p<0.01, ** p<0.05, * p<0.1)

Apendix C additionally reports results for regression using variables in levels, where AR process is much more evident (however Hansen and AR tests statistics has warning signs of weak instruments). From the table above however we see that adding human capital proxies weakens the instrument set and does not bring additional explanatory power, falling into the trend of HC variables not being significant. When only subsamples are analyzed the Hansen test is meaningless since there are too many instruments per observations, but AR(2) test still suggests that they are jointly valid. Therefore I regard specifications (1)-(2) and (7)-(8) as the most likely to have good instrumentation sets. These specifications, while confirming the previous findings for capital inputs and growth of labor force coefficients, suggest no significant effect of PEN variable (except for the case of developed countries but only at 10 percent significance level). Considering that the variable of interest is instrumented by its own lagged levels, this may indicate that OLS coefficients are biased upwards due to endogeneity, but insignificance in coefficients makes it impossible to say that with confidence (95% confidence interval suggests that the effect may be either higher or lower than that found using OLS). To understand the effect of permanent shift in independent variable in case the LDV is included calculation of the long-run multiplier is needed. The long-run multiplier is calculated by

assuming that *Y* variable is in its steady-state on both sides of the regression equation. If α is a coefficient for LDV then it follows that the LR multiplier for *Xit* is calculated as follows:

$$Y_i^{ss} = \frac{\beta}{1-\alpha} X_{it}$$

In case of specification (7), if we accept the significance level, this implies that LR multiplier for PEN variable is equal to 0.063. In other words 1% faster PEN growth leads to 0.063% increase in productivity growth.

Now I turn to various robustness checks, which would be based on CCE and GMM estimators only, for space reasons and since CCE is a more general OLS estimator. Firstly it seems important to replace the PEN index by the general statistics on the Internet users (all means) share of population.

Table 8 Results for CCE and GMM with differenced log of Internet users

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(CCE)	(CCE)	(CCE)	(CCE)	(CCE)	(GMM)	(GMM)	(GMM)	(GMM)	(GMM)
Δln(LPROD)	DVLPD	DVLPD	DVLPD	ALL	ALL	DVLPD	ALL	ALL	ALL	ALL
L1. -//-						0.0841*	0.0765	0.0458	0.00394	-0.00204
						(0.0430)	(0.122)	(0.123)	(0.113)	(0.114)
Δln(USER)	0.0404*	-0.0151	0.0889	0.00795	0.0134	0.0394	0.0129	-0.0224	0.0164	-0.0187
	(0.0227)	(0.0321)	(0.0585)	(0.0130)	(0.0169)	(0.0393)	(0.0216)	(0.0251)	(0.0212)	(0.0250)
Δln(I/GDP)	0.0923***	0.123***	0.0792*	0.0442***	0.0491**	0.0582	0.0122	0.00833	0.0108	0.00553
	(0.0224)	(0.0293)	(0.0404)	(0.0148)	(0.0240)	(0.0353)	(0.0375)	(0.0368)	(0.0418)	(0.0425)
N [LF growth]	-0.0760	-0.231	0.567	0.0882	-0.927***	-0.612***	-0.70***	-0.687***	-0.660***	-0.636**
	(0.395)	(0.144)	(0.876)	(1.260)	(0.321)	(0.128)	(0.105)	(0.106)	(0.107)	(0.110)
Δln(SCH)		0.110			0.118				0.0158	0.0119
		(0.0741)			(0.0960)				(0.0297)	(0.0274)
Δln(Tert_enr)			0.121							
			(0.0846)							
Obs	578	542	491	1,588	1,362	495	1,354	1,354	1,161	1,161
Countries	42	41	41	118	110	42	118	118	110	110
F-value	1.706	0.991	0.685	1.009	0.597	23.63	9.721	10.47	7.839	8.488
R-squared	0.703	0.755	0.776	0.597	0.715					
p>CD stat.	0.0809	0.592	0.461	0.0292	5.36e-05					
Instrum. (Z)						80	121	70	122	71
Lag lim. on Z						4	6	3	6	3
AR(2) p-value						0.118	0.173	0.250	0.114	0.140
Hansen test p-value						0.999	0.553	0.140	0.469	0.122

Robust standard errors (with small-sample correction for GMM) in parentheses (*** p<0.01, ** p<0.05,* p<0.1)

Table 8 suggests that growth of number of people having some kind of Internet access is not significantly connected with changes in labor productivity outside the developed countries (and only in case of CCEMG). Postestimation tests suggest that the preferred specifications are

(1), (5) and (8). The first column can be viewed as weak support for the previous findings. However, in case of developing countries (omitted in the table above) and all countries together the results are less conclusive than before. Lower significance is arguably due to a one-dimensional measure of Internet access being a less robust indicator of Internet infrastructure development in a country.

Next I allow the two components of PEN index: broadband subscribers share of population and mobile-cellular subscribers' share to enter regression separately. Secure servers percentage is not considered unlike in previous setups. The results are presented in Table 9.

Table 9 Results for Difference GMM with Broadband and Mobile subscribers' shares separate

	(GMM)	(GMM)	(GMM)	(GMM)	(GMM)	(GMM)	(GMM)	(GMM)	(GMM)	(GMM)
Δln(LPROD)	DVLPD	DVLPD	DVLPD	DVLPD	DEV	DEV	ALL	ALL	ALL	ALL
L1. -//-	0.0643	0.0633	0.0776	0.0665	0.196*	0.0194	0.179**	0.174**	0.0923	0.0408
	(0.0515)	(0.0513)	(0.0518)	(0.0507)	(0.105)	(0.133)	(0.0820)	(0.0857)	(0.110)	(0.108)
Δln(BB)	0.0258**	0.0323			-0.0065		0.0162	0.0149		
	(0.0126)	(0.0200)			(0.0096)		(0.0107)	(0.0112)		
Δln(MB)			0.0564***	0.0550***		0.0376**			0.0242**	0.0345**
			(0.0119)	(0.0137)		(0.0156)			(0.00984)	(0.0152)
Δln(I/GDP)	0.0477	0.0447	0.0492	0.0518	0.0498**	0.0029	0.0526**	0.0511**	0.0109	0.00810
	(0.0350)	(0.0384)	(0.0363)	(0.0386)	(0.0244)	(0.0374)	(0.0211)	(0.0211)	(0.0351)	(0.0346)
N [LF growth]	-0.552***	-0.583***	-0.604***	-0.616***	-0.654***	-0.785***	-0.602***	-0.607***	-0.708***	-0.701***
	(0.127)	(0.143)	(0.130)	(0.138)	(0.124)	(0.128)	(0.0917)	(0.0912)	(0.103)	(0.102)
Obs	484	484	502	502	758	872	1,242	1,242	1,374	1,374
Countries	42	42	42	42	77	76	119	119	118	118
F-value	23.29	19.38	38.49	29.78	7.15	6.06	12.17	11.16	10.08	9.982
Instrum. (Z)	121	70	121	70	70	70	121	70	121	70
Lag lim. on Z	6	3	6	3	3	3	6	3	6	3
AR(2) p-value	0.451	0.481	0.275	0.235	0.188	0.267	0.357	0.370	0.161	0.360
Hansen test p-value	1	0.999	1	0.984	0.379	0.061	0.265	0.195	0.310	0.147

Robust standard errors (with small-sample correction for GMM) in parentheses (*** p<0.01, ** p<0.05,* p<0.1)

The results are interesting because they seem to suggest a somewhat higher positive correlation between the growth of mobile subscribers' number and labor productivity than between broadband subscribers and the dependent variable. It is difficult to come with an intuitive explanation as to why this could happen. One possibility is indeed a higher impact of mobile phones usage on labor productivity in case higher changes represent some innovation allowing mobile phones to be more extensively used in the work environment (while broadband subscription became a standard office utility earlier, thus majority of changes in share of

subscribers may come from home users' dynamics) [2]. However since regressions with mobile share all have insignificant capital share coefficient, the possibility of some estimation issue (having to do with correlations and omitted effects) exists. It should be noted that differenced "mobile" series exhibits less much less fluctuations compared to "broadband" series, so the fit of isolated series has a higher chance of being incidental.

In the developed countries it seems that both variables have positive and significant coefficient, but significance of broadband subscribers share of population' coefficient disappears in the other samples. Difference GMM thus supports the hypothesis of heterogeneity in slope coefficients depending on the sample. It is possible that in the developing countries (including LDCs) growth of broadband penetration is not significantly correlated with productivity growth because of possible lack of human capital to implement certain cost-saving business practices related to Internet usage. On the other hand the instrumentation set is once again weak in case of developing countries as suggested by Hansen test and AR(2) statistic. Absence of significance of the LDV in most of the specification (except the case of the full sample) implies that there is no need to calculate long-run multipliers, as the there are no apparent dynamic effects in labor productivity.

There is a reason to believe that the productivity effect of changes in penetration of ICT may be not contemporaneous, due to people having to adapt to new technologies / enter labor market. So the reason for checking the effect of lagged independent variables is to test for delayed effects of these variables. The results of this exercise are presented in Table 10. For space reason the same set of test for separate measures of Internet penetration was omitted. Evidently the fit of the model falls as we increase lag lengths of PEN index, moreover the Hansen statistic indicate that the instrument set becomes invalid once they are included. Therefore there seems to be no apparent connection between today labor productivity growth and growth of Internet penetration index one or more years back. Consistent with previous coefficients obtained by Difference GMM there is no significance in the effect of Internet penetration (as measured by PEN index) in the full sample.

[2] According to OECD most firms in member countries have a broadband connection – 95% of all enterprises (OECD, 2015)

Table 10 Results for Difference GMM with lags of PEN index

	(1)	(2)	(3)	(4)	(5)	(6)
Δln(LPROD)	ALL	ALL	ALL	ALL	ALL	ALL
L1. -//-	0.131*	0.0959*	0.119	0.0804	0.106	0.0602
	(0.0772)	(0.0539)	(0.0721)	(0.0529)	(0.0701)	(0.0683)
L1.Δln(PEN)	0.0199			0.0210		0.0622**
	(0.0292)			(0.0304)		(0.0249)
L2. -//-		-0.00151		-0.0173	0.0276	-0.0186
		(0.0206)		(0.0168)	(0.0310)	(0.0178)
L3 -//-			-0.00632		-0.00653	-0.00140
			(0.0267)		(0.0128)	(0.0103)
Δln(I/GDP)	0.0506**	0.0597**	0.0745***	0.0592**	0.0732***	0.0713***
	(0.0248)	(0.0261)	(0.0263)	(0.0267)	(0.0262)	(0.0266)
N [LF growth]	-0.551***	-0.656***	-0.762***	-0.629***	-0.687***	-0.687***
	(0.0871)	(0.107)	(0.122)	(0.0988)	(0.122)	(0.104)
Obs	1,158	1,045	935	1,036	926	919
Countries	116	115	114	114	113	112
F-value	12.65	12.93	10.04	12.74	9.194	10.30
Instrum. (Z)	68	64	59	75	69	79
Lag lim. on Z	3	3	3	3	3	3
AR(2) p-value	0.798	0.419	0.497	0.586	0.332	0.539
Hansen test p-value	0.194	0.0531	0.0857	0.126	0.0791	0.0374

Robust standard errors with small-sample correction in parentheses (*** p<0.01, ** p<0.05, * p<0.1)

VII. Conclusion

The goal of the paper was to measure the impact of Internet penetration's growth on total labor productivity using the various econometric methods, aimed at reducing the biases identified by the previous empirical literature. The study is particularly relevant due to the importance of the Internet as a major innovation in communication technologies sector, highlighted by a number of theoretical and empirical studies, as well as official governmental programmes (e.g. the Digital Agenda for Europe). Increased availability and quality of the Internet connection is expected to facilitate distribution of information, allow further cost-savings and development of new business practices. It could also benefit the society in general since the development of Internet infrastructure benefits service providers only partially, while spillovers into the economy at large are highly probable if the Internet is the general purpose technology.

This study's empirical findings highlight the following. Firstly as expected from the Solow model there is a significant positive link between investment in capital stock and labor productivity; negative link between labor force growth rate and labor productivity. The effect of changes in human capital investment on productivity growth has a positive point estimate but is insignificant across the specifications (with exception for tertiary enrollment rate changes in developed countries, which has a weakly significant coefficient). This may be due to endogeneity of education variable or due to discrepancies in data reporting between countries and gaps in the data. Since human capital proxies weakened instrumentation sets, in later GMM tests the standard Solow model was tested.

As far as results of OLS estimators go, the significant positive connection (in differences) between Internet penetration index and productivity is established only for the developed countries' sample. The coefficient suggests that 1% increase in the Internet penetration growth rate is associated with about 0.02-0.03% increase in GDP per employed growth rate.

When using Difference GMM, that is supposed to account for reverse causality bias, the significance disappears. However robustness tests indicate that when the index is replaced by either broadband subscribers' or mobile subscribers' shares the later has a positive and significant coefficient across specifications, while the former is significant only in case of developed countries. As such heterogeneity in impact of Internet penetration variables due to the country sample is established using multiple estimators. 1% faster growth of mobile subscribers' share of population is associated with about 0.03-0.05% faster labor productivity growth universally. The positive role of mobile usage is in line with the latest empirical firm-level findings (see Bertscheck & Niebel, 2016).

The findings allow arguing that a positive productivity effect from growth of Internet providing infrastructure (which allows growth of Internet usage) is expected. The regressions were performed on first differences, so the positive effect would come from acceleration of growth, which is possible only in case of some innovation in the field of ICT, since the rates of growth have been declining in recent years. On the firm-level such innovations could be represented by something which can allow more rapid implementation of electronic orders and cloud computing. It is however possible that the further productivity effect may come from improvements in quality of Internet connection not reflected by subscriptions per se (the OECD has recently started to gather statistics on speeds of Internet access, which could be increasing with investment in infrastructure despite lower growth of subscribers). However the cost-benefit analysis was beyond the scope of this paper, so it is not possible to tell exactly how justified would be some kind of governmental programmes promoting Internet infrastructure growth.

Certain data limitations prevented a more in-depth analysis of ICT factors behind changes in productivity. For example low number of observations on Internet usage (both by individuals and firms) hinders some types of analysis (especially when using CCEMG, System GMM estimators). It should be noted that the choice of subscribers' share plus number of secure servers may only partially reflect the (potential) productivity influence of such complex phenomenon as the Internet.

These considerations suggest that future research conducted on more fine-grained variables, reflecting Internet infrastructure development (which have started to be gathered recently in some countries), could paint a more precise picture of the connection between the Internet and productivity. Of course such research should also incorporate advance econometric methods and/or less aggregate data, in order to distinguish between productivity effects in certain industries. Lastly, the future research could aim at exploring the productivity effects of increased computer literacy and competency in conjunction with increasing Internet penetration.

References

Aghion, P., Akcigit, U., Howitt, P. (2014). What Do We Learn From Schumpeterian Growth Theory? in Aghion, P., Durlauf, S. (eds) Handbook of Economic Growth Volume 2, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 5 May 2017]

Aghion, P., Howitt, P. (1998). Endogenous Growth Theory, Third Edition., London: The MIT Press

Arellano, M., Bond, S. (1991). Some test of specifications for panel data: Monte Carlo evidence and an application to employment equation, Review of Economic Studies, [e-journal] vol. 58, iss. 2, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 10 May 2017]

Aschauer, D. (1989). Is public expenditure productive?, Journal of Monetary Economics, [e-journal] vol. 23, iss. 2, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 11 May 2017]

Barro, R. (1991). Economic Growth in Cross Section of Countries, The Quarterly Journal of Economics, [e-journal] vol. 106, iss. 2, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 May 2017]

Bertschek, I., Niebel, T. (2016). Mobile and more productive? Firm-level evidence on the productivity effects of mobile internet use, Telecommunications Policy, [e-journal] vol. 40, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 10 June 2017]

Bowsher, S. (2002). On testing overidentifying restrictions in dynamic panel models, Economics Letters, [e-journal] vol. 77, iss. 2, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 20 June 2017]

Chudik, A., Pesaran, H. (2015). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors, Journal of Econometrics, [e-journal] vol. 188, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 26 May 2017]

Comin, D., Mestieri, M. (2014). Technology Diffusion: Measurement, Causes, and Consequences in Aghion, P., Durlauf, S. (eds) Handbook of Economic Growth Volume 2, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 5 May 2017]

Cronin, F., Parker, E., Colleran, E., Gold, M. (1991). Telecommunications infrastructure and economic growth, An analysis of causality, Telecommunications Policy, [e-journal] vol. 15, iss. 6, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 May 2017]

Czernich, N., Falck, O., Kretschmer, T., Woessman, L. (2011). Broadband Infrastructure and Economic Growth, The Economic Journal, [e-journal] vol. 121, iss. 552, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 19 April 2017]

Datta, A., Agarwal, S. (2004). Telecommunications and economic growth: a panel data approach, Applied Economics, [e-journal] vol. 36, iss. 15, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 17 May 2017]

Ditzen, J. (2016). xtdcce: Estimating Dynamic Common Correlated Effects in Stata, SEEC Discussion Paper 1601, available at: http://seec.hw.ac.uk/images/discussionpapers/SEEC_DiscussionPaper_No8.pdf [Accessed 20 May 2017]

Draca, M., Sadun, R., Van Reenen, J. (2009). Productivity and ICTs: A review of the evidence in Avgerou, C., Mansell, R., Quah, D., Silverstone, R. (eds) The Oxford Handbook of Information and Communication Technologies, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 28 May 2017]

Durlauf, S., Johnson, P., Jonathan, T. (2004). Growth Econometrics in Aghion, P., Durlauf, S. (eds) Handbook of Economic Growth, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 29 May 2017]

Enders, W. (2015). Applied econometric time series, Fourth edition, USA: John Wiley & Sons, Inc.

Evangelista, R., Guerrieri, P., Meliciani, V. (2014). The economic impact of digital technologies in Europe, Economics of Innovation and New Technology, [e-journal] vol. 23, iss. 8, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 11 April 2017]

Evangelista, R., Savona, M. (2003). Innovation, employment and skills in services. Firm and sectoral evidence, Structural Change and Economic Dynamics, [e-journal] vol. 14, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 29 April 2017]

Gruber, H., Hatönen, J., Koutroumpis, P. (2014). Broadband access in the EU: An assessment of future economic benefits, Telecommunications Policy, [e-journal] vol. 38, iss. 3, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 2 June 2017]

Koutroumpis, P. (2009). The economic impact of broadband on growth: A simultaneous approach, Telecommunications Policy, [e-journal] vol. 33, iss. 9, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 18 April 2017]

Leff, N. (1984). Externalities, Information Costs, and Social Benefit-Cost Analysis for Economic Development: An Example from Telecommunications, Economic Development and Cultural Change, [e-journal] vol. 32, iss. 2, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 15 May 2017]

Litan, R., Rivlin, A. (2001). Projecting the Economic Impact of the Internet, Journal of Financial Transformation, [e-journal] vol. 2, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 May 2017]

Madden, G., Savage, S. (1997). CEE telecommunications investment and economic growth, Information Economics and Policy, [e-journal] vol. 10, iss. 2, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 May 2017]

Mankiw, G., Romer, D., Weil, D. (1992). A Contribution to the Empirics of Economic Growth, The Quarterly Journal of Economics, [e-journal] vol. 107, iss. 2, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 15 May 2017]

Melody, W. (2009). Markets and policies in new knowledge economies in Avgerou, C., Mansell, R., Quah, D., Silverstone, R. (eds) The Oxford Handbook of Information and Communication Technologies, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 28 May 2017]

Nickel, S. (1981). Biases in dynamic models with fixed-effects, Econometrica , [e-journal] vol. 49, iss. 6 , Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 1 June 2017]

OECD (2015). OECD Digital Economy Outlook 2015 [pdf] Available at: http://www.oecd.org/internet/oecd-digital-economy-outlook-2015-9789264232440-en.html [Accessed 18 April 2017]

Pesaran, H., (2006). Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure, Econometrica, [e-journal] vol. 74, iss. 4, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 27 May 2017]

Romer, P. (1990). Endogenous Technological Change, Journal of Political Economy, [e-journal] vol. 98, iss. 5, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 15 May 2017]

Roodman, D. (2006). How to do xtabond2: An Introduction to “Difference” and “System” GMM in Stata, Center for Global Development Working Paper 103, available at: http://www.cgdev.org./files/11619_file_HowtoDoxtabond6_12_1_06.pdf [Accessed 9 May 2017]

Roodman, D. (2009). A Note on the Theme of Too Many Instruments, Oxford Bulletin of Economics & Statistics, [e-journal] vol. 71, iss. 1, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 15 May 2017]

Röller, L-H., Waverman, L. (2001). Telecommunications Infrastructure and Economic Development: A Simultaneous Approach, American Economic Review, [e-journal] vol. 91, iss. 4, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 May 2017]

Solow, R. (1957). Technical Change and the Aggregate Production Function, The Review of Economics and Statistics, [e-journal] vol. 39, iss. 3, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 5 May 2017]

Solow, R. (1987). We’d better watch out, New York Times Book Review, July 12, 1987, p.36

Stiroh, K. (2002). Information Technology and the U.S. Productivity Revival: What Do the Industry Data Say?, The American Economic Review , [e-journal] vol. 92, iss. 5 , Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 10 May 2017]

Thompson, H., Garbacz, C. (2007). Mobile, fixed line and Internet service effects on global productive efficiency, Information Economics and Policy, [e-journal] vol. 19, iss. 2, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 17 May 2017]

van Ark, B., O’Mahony, M., Timmer, M. (2008). The Productivity Gap between Europe and the United States: Trends and Causes, Journal of Economic Perspectives, [e-journal] vol. 22, iss. 1, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 20 April 2017]

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators, Journal of Econometrics, [e-journal] vol. 126, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 19 May 2017]

Appendix A. Data sources / correlations table

Table 1 Data sources

Name	Source
LPROD	Total Economy Database™ - Output, Labor and Labor, Productivity 1950-2016
I/GDP	World Bank - World Development Indicators
N	Total Economy Database™ - Output, Labor and Labor, Productivity 1950-2016
SCHOOL	UNECSO Institute for Statistics (UIS); United Nations Population Division: World Population Prospects 2017
TERT_Enrl	World Bank – World Development Indicators
TERT_Ed	OECD Education Statistics: http://www.oecd-ilibrary.org/education/data/oecd-education-statistics_edu-data-en
PEN:	
└ 0.6**BB_sub*	ITU World Telecommunication/ICT Indicators Database: http://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2017/Fixed_broadband_2000-2016.xls
└ 0.3**MB_sub*	ITU World Telecommunication/ICT Indicators Database: http://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2017/Mobile_cellular_2000-2016.xls
└ 0.1**SERV*	Netcraft (netcraft.com) and World Bank population estimates: http://data.worldbank.org/indicator/IT.NET.SECR.P6

Table 2 Correlations (levels and differences)

```
. corr lnlprod lns_intinf lns_inv glf lns_sc lns_enroll lns_ter
(obs=419)
```

	lnlprod	lns_in~f	lns_inv	glf	lns_sc	lns_en~l	lns_ter
lnlprod	1.0000						
lns_intinf	0.5274	1.0000					
lns_inv	-0.1683	-0.1251	1.0000				
glf	0.0282	-0.1605	0.2106	1.0000			
lns_sc	-0.1286	-0.2308	-0.0450	0.2911	1.0000		
lns_enroll	0.3507	0.4848	0.0024	-0.3081	-0.1338	1.0000	
lns_ter	0.5912	0.6555	-0.0717	-0.0798	0.0613	0.5238	1.0000

```
. corr dlnlprod dlns_intinf dlns_inv glf dlns_sc dlns_enrolls dlns_ter
(obs=382)
```

	dlnlprod	dlns_i~f	dlns_inv	glf	dlns_sc	dlns_e~s	dlns_ter
dlnlprod	1.0000						
dlns_intinf	0.3772	1.0000					
dlns_inv	0.3168	0.2247	1.0000				
glf	-0.1799	0.1811	0.2139	1.0000			
dlns_sc	-0.0651	-0.1512	-0.0002	-0.0061	1.0000		
dlns_enrolls	0.2114	0.2424	0.0263	0.1151	-0.0259	1.0000	
dlns_ter	0.0761	0.1103	-0.0312	-0.0227	-0.1541	0.1327	1.0000

Appendix B

Table 1 Augmented Dickey-Fuller test (MacKinnon approximate p-value)

	Levels	**1st difference**	**2nd difference**
LPROD	0.2438	0.1056	0.0002
I/GDP	0.4435	0.1306	0.0000
N	0.8514	0.1180	0.0000
SCHOOL	0.0207	0.0146	0.0005
TERT_Enrl	0.8189	0.0504	0.0112
TERT_Ed	0.7928	0.0105	0.0005
PEN	0.4731	0.1200	0.0000

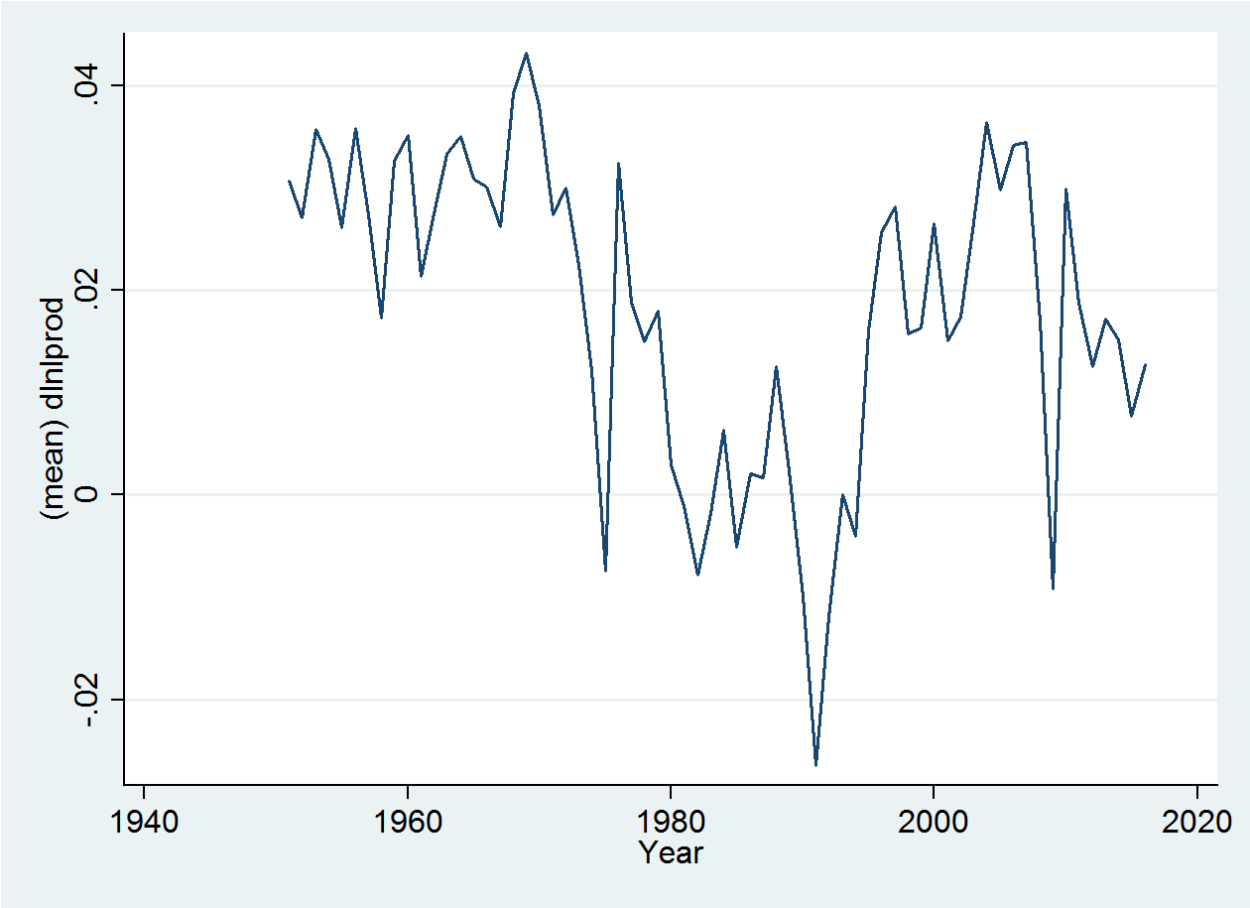

Fig.1 Labor productivity (first differenced) series

Appendix C. Difference GMM on levels of variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ln(LPROD)	ALL	ALL	ALL	ALL	ALL	ALL	DVLPD	DVLPD	DEV	DEV
L1. -//-	0.796***	0.780***	0.761***	0.733***	0.736***	0.728***	0.831***	0.814***	0.795***	0.783***
	(0.0426)	(0.0350)	(0.0378)	(0.0430)	(0.0579)	(0.0554)	(0.0556)	(0.0722)	(0.0525)	(0.0443)
ln(PEN)	0.0186*	0.0250***	0.0238***	0.0292***	0.0166**	0.0229**	0.0122	0.0154	0.0219*	0.0317***
	(0.00947)	(0.00838)	(0.00685)	(0.00920)	(0.00725)	(0.00991)	(0.0150)	(0.0207)	(0.0112)	(0.0110)
ln(I/GDP)	0.0678**	0.0620**	0.0581***	0.0557***	0.0668***	0.0632***	0.0371**	0.0332*	0.0725**	0.0676**
	(0.0279)	(0.0264)	(0.0119)	(0.0127)	(0.0195)	(0.0204)	(0.0183)	(0.0185)	(0.0323)	(0.0301)
N [LF growth]	-0.553***	-0.542***	-0.518***	-0.521***	-0.445***	-0.475***	-0.413***	-0.442***	-0.604***	-0.560***
	(0.0791)	(0.0719)	(0.0765)	(0.0740)	(0.0795)	(0.0715)	(0.0897)	(0.0878)	(0.121)	(0.111)
ln(SCH)			0.00138	0.00305						
			(0.0151)	(0.0183)						
ln(Ter_Enr)					0.0230	0.0271*				
					(0.0157)	(0.0155)				
Obs	1,340	1,340	1,166	1,166	943	943	525	525	815	815
Countries	118	118	109	109	104	104	42	42	76	76
Instrum. (Z)	133	76	134	77	134	77	133	76	133	76
Lag lim. on Z	6	3	6	3	6	3	6	3	6	3
AR(2) p-value	0.650	0.713	0.924	0.837	0.862	0.894	0.0257	0.0237	0.911	0.834
Hansen test p-value	0.662	0.0398	0.809	0.0923	0.872	0.158	1	0.996	1	0.157
F-stat	172.5	168	113	103.9	86.26	87.51	956.3	631.6	87.91	85.77
F-stat: p-val.	0	0	0	0	0	0	0	0	0	0
Year-eff signif	Y	N	N	N	N	N	N	Y	N	N

Robust standard errors with small-sample correction in parentheses (*** p<0.01, ** p<0.05, * p<0.1)