A panel Granger-causality test of endogenous vs. exogenous growth

Jochen Hartwig

Swiss Economic Institute at ETH Zurich (KOF ETH)

Weinbergstrasse 35, 8092 Zurich

Switzerland

Telephone: +41-44-632-7331

Fax: +41-44-632-1218

E-mail: hartwig@kof.ethz.ch

Keywords: Exogenous growth, endogenous growth, panel Granger-causality tests

JEL classification: C12; C23; O41

Abstract

The paper proposes a new test of endogenous vs. exogenous growth theories based on the Granger-causality methodology and applies it to a panel of 20 OECD countries. The test yields divergent evidence with respect to physical and human capital. For physical capital, the test results favor Solow-type exogenous growth theory over AK-type endogenous growth models. On the other hand, the test results lend support to human capital oriented endogenous growth models – like the Uzawa-Lucas model – rather than to the human capital augmented Solow model.

1. Introduction

In the standard neoclassical Solow (1956, 1957) model, economic growth is exogenous.[1] Once the economy reaches its steady state, technological progress – which is not explained by the model and is hence 'exogenous' – is the only driver of per-capita growth. In particular, investment in physical capital cannot increase growth any longer. The same is true for investment in human capital (with which the original Solow model has been 'augmented' later on, see Mankiw, Romer and Weil, 1992).

Widespread dissatisfaction with the fact that the Solow model cannot 'explain' long-term economic growth has sparked off an alternative *endogenous* growth theory which emerged in the 1980s. At the risk of over-simplifying, we can distinguish three classes of endogenous growth models. In the first class, the AK models (instigated by Romer, 1986),[2] technology is assumed to grow in line with capital accumulation. Therefore, growth remains dependent on capital accumulation even in the long run, and the 'endogenous' technical progress offsets the growth-cushioning effects of diminishing returns to capital accumulation that characterize the Solow model.

The second class of endogenous growth models highlights human capital accumulation. Uzawa (1965) and Lucas (1988) present models in which both physical capital and human capital enter the production function. Workers use only a fraction of their human capital in the production process and allocate the rest to an increase in human capital through education. The growth of human capital in the economy depends on the share of total time spent on education and a productivity parameter which measures the efficiency of education in increasing human capital. Per-capita gross domestic product (GDP) growth can continue as long as time is devoted to education. In the Uzawa-Lucas model, the growth rate of human capital acquires the role that technical progress plays in the Solow model, yet unlike technical progress, (optimal) human capital accumulation can be 'explained' (from an inter-temporal utility maximization calculus).

Following Romer's (1990) lead, a third class of endogenous growth models aims at endogenizing technological progress by focusing on research and development (R&D) and innovation. The works of Grossman and Helpman (1991, 1994) are prime examples for this line of research. In their model, productivity growth is the result of specialization and the

[1] I use the term 'Solow model' for short. Swan (1956) is widely regarded as an equally important early contribution so that the neoclassical growth model is also known as the Solow-Swan model.

[2] Aghion and Howitt (1998) note that Frankel (1962) presented an early AK model that went largely unnoticed by the profession.

division of labor – an idea that sits comfortably with the legacy of Adam Smith. Specialization can be increased by devoting more labor to research and development. As a consequence, new innovative products will be forthcoming, and the variety of intermediate products that serve as inputs into the production of final products will rise.[3] In other words, the more intermediate goods there are, the greater is the division of labor and the more productive is the labor that produces the final products. The structure of the argument is not too different from the Uzawa-Lucas model. In both cases, the society can secure long-term growth by devoting some of its time-resources to increasing the stock of an input factor – human capital and research labor, respectively – on which output hinges. However, there are differences between the models, too. For instance, unlike the Uzawa-Lucas model, the Romer (1990) and Grossman-Helpman models imply a ‘scale effect’. This means that the growth rate of GDP will be the higher, the bigger the economy (the labor force) is. Also, the ultimate reasons for the possibility of long-term non-zero per-capita growth are different across the models. In the AK and Uzawa-Lucas models, the assumption of constant returns to physical and human capital, respectively, allows for long-term growth. In the Romer (1990) and Grossman and Helpman models on the other hand, long-term growth either results from spillovers in research activities whose benefits cannot be privatized or from imperfect competition. Innovative activities that increase the number of intermediate products will only occur if the innovator has the chance to avert immediate imitation by competitors. Therefore, endogenous growth models of the ‘third class’ usually assume monopolistically competitive instead of perfectly competitive markets. This implies that the factors of production (capital and labor) generally receive less than their marginal products, and that the innovator can earn monopoly rents.

It is natural to ask which growth theory is ‘right’ because the answer to this question is of obvious relevance for economic policy. If endogenous growth theory is the ‘right’ theory, then economic policy can have an impact on the long-term growth rate. If, on the other hand, growth is exogenous, then policy can only raise the growth rate temporarily above its steady state value to which it must ultimately return. Not everybody shares the agnosticism of Temple (1999: 152) who writes: “the debate on whether policy affects the long-run growth rate or just the steady state level of income is almost impossible to resolve, and not much of practical importance will turn on it”. On the contrary, a considerable literature has emerged that seeks to test endogenous vs. exogenous growth theories empirically. The next section will

[3] In Aghion and Howitt’s (1992, 1998) approach, innovations displace rather than complement the old product variants – a process that Aghion and Howitt link to Schumpeter’s theory of creative destruction.

review this literature (for the first time as far as I see). The key finding of this review is that no consensus exits, neither on which methods are the most adequate for performing the empirical tests, nor on whether growth is actually endogenous or exogenous. Against this backdrop, section 3 proposes a new test of endogenous vs. exogenous growth which is based on the panel Granger-causality methodology. Section 4 introduces the data to be used in the estimations. Section 5 presents the results of the empirical analysis, and section 6 concludes.

2. Tests of endogenous vs. exogenous growth – A review of the literature

A useful starting point for a review of the empirical literature that aims at testing endogenous vs. exogenous growth theories is the seminal contribution by Mankiw, Romer and Weil (1992). Regressing the log GDP per working-age person in 1985 on measures of physical capital (the share of real investment – including government investment – in real GDP in 1985)[4] and human capital in the form of education (secondary school enrollment rates) in a cross-sectional framework of up to 98 countries, Mankiw, Romer and Weil (1992: 407) conclude that "an augmented Solow model that includes accumulation of human as well as physical capital provides an excellent description of the cross-country data". Mankiw, Romer and Weil's analysis, which has been extended to panel data by Islam (1995) and others, thus seemingly bolsters the theory of *exogenous* growth.

However, Bernanke and Gürkaynak (2001: 12) point out that Mankiw, Romer and Weil's "basic estimation framework is broadly consistent with *any* growth model that admits a balanced growth path"; hence this framework is arguably unable to discriminate between theories of endogenous and exogenous growth. The same point has been made more recently by Arnold, Bassanini and Scarpetta (2007) who derive a growth equation that nests both the Solow model and the Uzawa-Lucas model of endogenous growth. Arnold, Bassanini and Scarpetta argue that this equation offers the possibility to discriminate between the two models on the basis of the parameter estimate for the speed of convergence. The Solow model predicts a slower convergence than the Uzawa-Lucas model; and Arnold, Bassanini and Scarpetta claim that their empirical estimate for the speed of convergence parameter (based on OECD data) is too high to be consistent with the Solow model.[5] Hence they interpret their results – although these admittedly rest on a number of calibrations, the validity of which might be disputed – to yield evidence in favor of the theory of *endogenous* growth. Drawing

---

[4] This is problematic since 'real shares' do not have any conceptual meaning. Shares of demand components in GDP should therefore always be calculated in nominal terms.

[5] This point had already been made (in passing) by Bassanini and Scarpetta (2002).

on the same data sources as Mankiw, Romer and Weil (1992), Bernanke and Gürkaynak (2001) also reject several restrictions imposed by the Solow model, namely that the long-run growth rates of GDP and total factor productivity (TFP) are uncorrelated with investment rates in physical and human capital.[6] Bernanke and Gürkaynak do not find much evidence in favor of endogenous growth either, although they feel that their data are more in line with the Uzawa-Lucas and AK models than with the Solow model.

While Mankiw, Romer and Weil (1992) and Bernanke and Gürkaynak (2001) use cross-sectional data for their respective tests, Arnold, Bassanini and Scarpetta (2007) use pooled cross-sectional and time-series data. Tests of endogenous vs. exogenous growth based on time-series data have been instigated by Jones (1995). Jones concentrates on the first and third classes of endogenous growth models described above and argues that the large increases in the rate of physical investment and R&D expenditure that took place in the US (and elsewhere) over the past decades should have boosted GDP growth, given the predictions of endogenous growth theory. Since no such growth acceleration occurred, Jones concludes that the theories of endogenous growth are flawed. Aghion and Howitt (1998, ch. 12) defend endogenous growth theory against this influential critique, showing that modified endogenous growth models could reconcile the observed patterns of R&D and physical capital expenditure on the one hand, and economic growth on the other.

A fundamental problem for all time-series based tests of endogenous vs. exogenous growth is how to filter out the short-term – or business cycle – effects in output in order to concentrate on the long-term interrelations that are relevant for the theory of economic growth. The most widely-accepted technique – used by Grier and Tullock (1989), Cashin (1995), Islam (1995), Caselli, Esquivel and Lefort (1996), Devarajan, Swaroop and Zou (1996), Mendoza, Milesi-Ferretti and Asea (1997), Kneller, Bleaney and Gemmell (1999) and others – is to take averages of five years (or more) of the data. Jones (1995) uses a different approach, though. He specifies an error-correction model for *annual* data. In an error-correction framework, the cointegrating relationship between the dependent and independent variables is usually interpreted as the long-run equilibrium relationship between these variables. Jones – as well as Arnold, Bassanini and Scarpetta (2007) – tests endogenous vs. exogenous growth by drawing inferences from the estimated coefficients for the contemporaneous relationship between the variables, which are interpreted as ‘long-run

[6] It has to be noted, though, that Inklaar, Timmer and van Ark (2008) were not able determine any influence of investment in information and communication technology or human capital on TFP growth in a new dataset for advanced economies, the EU KLEMS database (available at www.euklems.net).

parameters'. Jones finds no evidence that the coefficient of the investment rate is significantly positive. Therefore, a permanent positive shock to investment will not permanently raise the GDP growth rate, which speaks against AK model-style endogenous growth. Furthermore, the 'scale effect' predicted by the 'class three' endogenous growth models is in sharp contrast to the empirical evidence, according to Jones.[7] It has to be noted, though, that neither Jones (1995) nor Arnold, Bassanini and Scarpetta (2007) report results of cointegration tests so that it remains unclear whether their interpretation of the coefficients is indeed warranted.

Another argument leveled against endogenous growth theories is that countries – at least the advanced economies – seem to converge toward a common long-run growth rate (cf. Evans, 1996, 1998). This empirical finding is contrary to what endogenous growth models predict because, in these models, growth depends on factors that are country-specific and hence different across countries. Kocherlakota and Yi (1995) show, however, that convergence regressions cannot discriminate between endogenous and exogenous growth theories. More specifically, depending on the persistence of technology shocks, the coefficient of initial income may be positive even if growth is exogenous, and it may be negative even if growth is endogenous. In two other contributions, Kocherlakota and Yi (1996, 1997) produce evidence that is favorable for *endogenous* growth models. In their 1996 contribution, Kocherlakota and Yi draw on a distinction between endogenous and exogenous growth models that they claim to be less well known, namely that in the former class of models *temporary* innovations in government policies can have a permanent effect on the *level* of GDP, while in the latter class, they cannot. To test endogenous vs. exogenous growth, Kocherlakota and Yi regress GDP levels on their own lags and on lags of policy variables. They test whether the sum of the coefficients of the lagged policy variables is significantly different from zero, in which case the null hypothesis of exogenous growth could be rejected. They assume the error terms to be independently and identically distributed (IID) and

[7] Gong, Greiner and Semmler (2004) contend that the property of the Uzawa-Lucas model that an increase in the time spent on education raises the GDP growth rate permanently was also a 'scale effect', which they claim to be unrealistic. They remove this alleged scale effect and test the modified model, which – as they emphasize – is not an endogenous growth model anymore. Gong, Greiner and Semmler find the modified – exogenous – growth model to be compatible with time series data for the US and Germany. Dinopoulos and Thompson (2000), on the other hand, remove scale effects from Romer's model of endogenous technological change in a way that preserves the endogeneity of long-run growth. They find some empirical support for this 'augmented' Romer model in a sample of high-income countries, which is not robust to the way in which human capital is measured, however.

estimate the parameters with ordinary least squares (OLS).[8] Kocherlakota and Yi (1996) cannot reject the null for most of their policy variables (mostly tax rates and shares of different 'real' government investment expenditures in 'real' GDP[9]). However, they find a positive and robustly significant effect for non-military structural capital and interpret this finding as favorable for endogenous growth models.

In their 1997 contribution, Kocherlakota and Yi use the same testing strategy (including the aforementioned assumptions) and reach basically the same conclusions as in their 1996 paper. The new twist in the 1997 paper is that Kocherlakota and Yi include both public (capital) spending and revenue variables in their growth regressions. They find both sets of lagged coefficient sums to be statistically different from zero, with opposite signs. Kocherlakota and Yi interpret this finding not only as evidence in favor of endogenous growth theory,[10] but also as evidence against Jones who had declared: "If we characterize endogenous growth theory by the prediction that permanent changes in policy variables lead to permanent changes in growth, then this lack of persistent change in growth rates imposes a strong restriction on these models: either the variables that have permanent effects on growth exhibit little persistent change, or somewhat miraculously the movements in these variables have been offsetting" (Jones 1995: 521). Kocherlakota and Yi (1996) claim that the 'miracle' is actually happening: permanent changes in public investment would raise GDP growth rates if they were not exactly offset by permanent changes in tax rates – the latter being a natural consequence of the government budget constraint. Bleaney, Gemmell and Kneller (2001) confirm this result in dynamic panel estimations with both annual and five-year averaged data.[11]

Li (2002) also challenges Jones's (1995) results. He modifies Jones's regression equation by removing the lags of the dependent variable from the right-hand side and adding leads of the differenced investment variable. He then also tests the null hypothesis that the coefficient of the contemporaneous investment rate is zero against the alternative that it is positive. Li

[8] Whether this assumption is warranted will be discussed below.

[9] Again, this is problematic (see above, fn. 4).

[10] It has to be mentioned, though, that Kocherlakota and Yi (1996: 252) admit the possibility that their regressions "are only picking up transitional dynamics from exogenous growth models".

[11] Their table 4 shows, however, that when a terms of trade index – which Mendoza, Milesi-Ferretti and Asea (1997) found to be a significant explanatory variable for economic growth – is included in the regression, the investment ratio becomes insignificant, and the coefficient of the 'productive government expenditure' variable switches from significantly positive to significantly negative. Bleaney, Gemmel and Kneller do not comment on this effect.

finds significantly positive coefficients for the majority of his 24 OECD countries and for the full panel. Like Kocherlakota and Yi (1996), Li (2002: 94f.) admits, however, "that the positive link between growth and investment found in this study may also be consistent with other types of endogenous or exogenous growth models".

The tests of endogenous vs. exogenous growth discussed so far are basically regressions of the long-term GDP growth rate on variables that the Solow model predicts to be unimportant for growth in the long run (like investment or policy variables). Feenstra et al. (1999) choose a different approach. They concentrate on the above-mentioned 'third class' of endogenous growth models – models of the Romer-Grossman-Helpman type in which increased product diversity drives productivity growth. Feenstra and colleagues test whether changes in the relation of product variety between Taiwan and South Korea are correlated with changes in the difference in total factor productivity (TFP) growth between the two countries. They find "some degree of confirmation" (Feenstra et al., 1999: 319) for this link, which they take as supportive for the endogenous growth model.

3. A new test of endogenous vs. exogenous growth – panel Granger-causality

As the previous section has demonstrated, the empirical evidence on whether long-run economic growth is exogenous or endogenous is mixed. While there seems to be a slight preponderance of the endogenous growth evidence, this evidence is generally subject to certain reservations and qualifications. The time-series based tests of endogenous vs. exogenous growth mostly rely on annual data, despite the standard practice in growth econometrics being to take averages of five years (or more). As some authors admit, tests based on annual data have difficulties in discerning permanent effects of investment or policy shocks on the GDP growth rate from temporary effects that last for several years.[12] Li (2002: 111) calls for new tests to discriminate between exogenous and endogenous growth.

The present paper intends to meet this demand by introducing the panel Granger-causality testing methodology to this field of research. In line with most of the earlier literature, I will concentrate on the relatively clear-cut early endogenous growth models of the AK and Uzawa-Lucas type and test them against the Solow model.

We know that the basic tenet of endogenous growth models of the AK and Uzawa-Lucas type is that an increase in the investment rates in either physical or human capital will raise the steady state GDP growth rate. Therefore, in order to corroborate these models empirically we would need to find a significantly positive correlation between lagged investment growth

[12] Temple (1999: 132f.) also warns against using annual data in growth studies.

and GDP growth. This insight has been at the core of many earlier attempts to test endogenous vs. exogenous growth described in the previous section. For instance, in his pioneering study, Jones (1995) notes that the AK model suggests a dynamic relationship like:

$$g_t = A(L)g_{t-1} + B(L)i_t + \varepsilon_t \qquad (1),$$

where g_t denotes GDP growth in period t, i_t denotes the rate of investment in period t, ε_t is a stochastic shock, and A(L) and B(L) are two lag polynomials with roots outside the unit circle. Jones interprets the endogenous growth model as predicting that the sum of the coefficients in the polynomial B(L) is significantly positive.

Now, what is less recognized in the literature is that the Solow model, on the other hand, implies a significantly *negative* correlation between lagged investment growth and GDP growth. This has been pointed out, however, by Vanhoudt (1998) who notes that the Solow model's convergence hypothesis can be written as (I keep Vanhoudt's notation):

$$\gamma_t \equiv \ln\left(\frac{y_t}{y_{t-1}}\right) = \beta \cdot [\ln(y^*_t) - \ln(y_{t-1})] \qquad (2).$$

The parameter β ($0 \leq \beta \leq 1$) stands for the speed of convergence of GDP (y) to its steady state level (y^*), which is determined by the investment regime (s_t) according to:

$$\ln(y^*_t) = \frac{\alpha}{1-\alpha} \cdot \ln(s_t) - \frac{\alpha}{1-\alpha} \cdot \ln(n + x + \delta) \qquad (3).$$

Of course, n, x and δ stand for population growth, technological progress, and the depreciation rate, respectively. Now, if we differentiate equation (2) and insert (3), we get:

$$\Delta\gamma_t = -\beta \cdot \gamma_{t-1} + \beta \cdot \frac{\alpha}{1-\alpha} \cdot \ln(s_t) - \beta \cdot \frac{\alpha}{1-\alpha} \cdot \ln(s_{t-1}) \qquad (4),$$

from which follows that:

$$\gamma_t = (1-\beta) \cdot \gamma_{t-1} + \beta \cdot \frac{\alpha}{1-\alpha} \cdot \ln(s_t) - \beta \cdot \frac{\alpha}{1-\alpha} \cdot \ln(s_{t-1}) \qquad (5).$$

If a positive shock to investment growth raises the investment share above its steady state value in period t, this will affect GDP growth positively in the same period, but will have a negative impact on next period's growth as the GDP growth rate falls back to its steady state value *ceteris paribus*. Therefore, the Solow model predicts a negative sign for lagged investment growth in a growth regression if lagged GDP growth and current investment growth are controlled for.[13] It is important to note, however, that this is true only for medium-run lags. In the long run, the Solow model predicts no influence whatsoever of changes in

[13] Cf. also Podrecca and Carmeci (2001). Equation (5) can easily be augmented by human capital as Vanhoudt (1998) has demonstrated. The same is true, of course, for equation (1).

accumulation rates on per-capita GDP growth as the latter is solely driven by technological progress.

The Granger-causality testing methodology seems to be an ideal tool to examine the different sets of predictions about parameter signs empirically. In fact, as Vanhoudt (1998: 80) notes, equation (5) is “very close to a simple Granger causality equation”. Granger’s (1969) testable definition of causality has spawned a vast literature in which ‘Granger-causality tests’ have established themselves as a widely-used analytical tool in applied economics.[14] According to Granger’s definition of causality, a stationary time series Y_t is said to ‘cause’ another stationary time series X_t if – under the assumption that all other information is irrelevant – the inclusion of past values of Y_t significantly reduces the predictive error variance of X_t. In econometric practice, Granger-causality tests are carried out by regressing X_t on its own lags and on lags of Y_t. If the lags of Y_t are found to be jointly statistically significant, then the null hypothesis that Y_t does not Granger-cause X_t can be rejected. For the context of this paper this means that if – after lagged GDP growth and contemporaneous investment are controlled for – Granger-causality running from lagged investment to GDP growth is found to be significantly positive, then this is evidence in favor of endogenous growth theory.[15] If, however, negative Granger-causality in the medium run *and* no Granger-causality in the long run are found, then this speaks in favor of exogenous growth theory (cf. Vanhoudt, 1998).

Most Granger-causality tests that can be found in the literature involve only two variables X_t and Y_t (as described in the previous paragraph). For the purpose of the present paper, this is unfortunate because if we want to test exogenous growth against the Uzawa-Lucas model we should include both physical and human capital as right-hand side variables. Fortunately, to conduct *multivariate* Granger-causality tests is no problem at all. On the contrary, moving

[14] Some remain skeptical, however, since the concept of Granger-causality draws on only one aspect of the multi-faceted philosophical concept of causality, namely antecedence. Granger (1980) readily conceded this. Nevertheless, Thurman and Fisher (1988), for instance, tried to ridicule the concept by applying it to the ‘chicken-egg problem’, showing that eggs ‘cause’ chicken in the Granger sense but not vice versa. Also, the notion that ‘Christmas card sales Granger-cause Christmas’ is reiterated in the literature (cf. Atukeren, 2008, who shows, however, that the concept of Granger-causality survives this criticism because it refers to stochastic events, and Christmas is not stochastic). To avoid this kind of squabble, it probably would have been better not to introduce the term ‘causality’ at all and to speak of ‘improved predictability’ instead. As it has become a technical term in the meantime, ‘Granger-causality’ is kept here.

[15] The inclusion of contemporaneous values of Y_t is unusual in tests of whether Y_t Granger-causes X_t. Yet if the method is to be used to test endogenous vs. exogenous growth theories, then contemporaneous investment has to be controlled for, as is demonstrated by equation (5).

from a bivariate to a multivariate setting can help to alleviate the problem of 'spurious causality' (cf. Hsiao, 1982). Spurious causality can arise in a bivariate setting when both variables have 'common causes' that are absent from the regression equation. In this case, even if there is no other relationship between X_t and Y_t, the test will erroneously find Granger-causality. If all 'common cause' variables are included in the regression, however, the spurious causation between X_t and Y_t will vanish.[16] On the other hand, if no causality between X_t and Y_t is found in a multivariate setting, the conclusion of no-causality is only valid if there is also no causality in a bivariate setting (and if Y_t does not cause any variable that in turn causes X_t).

Besides the multivariate framework, another distinctive feature of the present paper against most of the earlier literature is its use of panel data. The notion of Granger-causality has not found its way into panel econometrics until the beginning of the new millennium. Hartwig (2009), who studies the causal nexus between health and economic growth in a panel Granger-causality framework, lists a dozen papers that have used panel Granger-causality tests earlier. Since the testing methodology is not implemented identically in that literature, a few words are necessary on how the method will be implemented here.

First, all data will be transformed into five-year average annual growth rates. This will on the one hand introduce the standard practice of eliminating the cyclical component into the literature on testing endogenous vs. exogenous growth and, on the other hand, help to avoid the problems in discerning permanent from temporary effects that tests based on annual data face.[17] Jones (1995) finds that the effects of shocks to investment on economic growth disappear after six years. (He takes this finding as evidence against endogenous growth theory). It will be remembered that Vanhoudt (1998) argues that if growth is exogenous, the Granger test should yield significantly negative medium-term coefficients because after the initial stimulus is over, the GDP growth rate returns to its steady state value. The long-term effect of investment on economic growth is zero. Based on Jones's (1995) finding, we can identify a five-year lag of the investment variables as the medium-term lag, whose coefficient should be negative, and a ten-year lag as the long-term lag, whose coefficient should be insignificant if the Solow model was right. Positive coefficients, on the other hand, would support the AK and Uzawa-Lucas models.

---

[16] Therefore, any finding of Granger-causality is in principle *prima facie* – as Atukeren (2007: 10) makes clear – because "the missing cause problem is not necessarily solved in a multivariate framework".

[17] Bleaney, Gemmell and Kneller (2001) have used five-year average growth rates before in this field of research.

As Granger-causality tests require stationary data, all time series will be tested for the presence of unit roots, applying a battery of now standard panel unit root tests. When these tests fail to detect unit roots, the panel estimation models can be set up. As only long-run coefficients are estimated, the restriction of identical coefficients of the lagged X_{it}, Y_{it} and Z_{it} variables across countries will be imposed.[18] Thus, I will estimate a time-stationary VAR model adapted to a panel context (as in Holtz-Eakin, Newey and Rosen, 1988) of the form:

$$X_{it} = \alpha_0 + \sum_{l=1}^{m} \beta_l X_{it-l} + \sum_{l=0}^{m} \delta_l Y_{it-l} + \sum_{l=0}^{m} \varphi_l Z_{it-l} + \mu_i + u_{it} \qquad (6).$$

X_{it}, Y_{it} and Z_{it} are the five-year averages of the growth rates of per-capita GDP, per-capita physical investment and per-capita human capital investment, respectively. N countries (indexed by i) are observed over T periods (indexed by t). I allow for country-specific effects μ_i. The disturbances u_{it} are assumed to be independently distributed across countries with a zero mean. They may display heteroscedasticity across time and countries, though.

Estimating equation (6) with pooled OLS presents an endogeneity problem since if the dummy variables (country-specific effects) affect GDP growth in one period they presumably affected them in the previous period also (cf. Nickell, 1981). The first step into the direction of correcting this endogeneity problem in dynamic panels is to take the first difference of all variables and to thereby eliminate the individual effects. Still, there remains a correlation between the lagged dependent variable, which is now in differences, and the error term. As a way around this problem, Arellano and Bond (1991) have proposed to use lags of the dependent variable from at least two periods earlier (in levels) as well as lags of the right-hand side variables as instruments in a Generalized Method of Moments (GMM) estimator. Arellano and Bover (1995) and Blundell and Bond (1998) have suggested to difference the instruments instead of the regressors in order to make them exogenous to the fixed effects. This leads from the ‘difference’ GMM to the ‘system’ GMM estimator, which is a joint estimation of the equation in levels and in first differences (cf. Roodman, 2006). In the next section, I will present results using the OLS, Arellano-Bond one-step system GMM, and Arellano-Bond two-step system GMM estimators. Based on the estimation results, a conclusion on causality will be reached by running Wald tests on the coefficients of the lagged Y_t’s and Z_t’s to check whether they are jointly statistically different from zero.

[18] Arnold, Bassanini and Scarpetta (2007), in their error correction model based on annual data, also impose this restriction on their long-run coefficients, but allow the short-run coefficients to vary across countries.

4. Data

Empirical tests of endogenous vs. exogenous growth require data on GDP and investment in physical and human capital. While OECD data for GDP and gross fixed capital formation can be obtained easily,[19] data on human capital formation are more cumbersome to retrieve. Pioneering studies like Barro (1991) and Mankiw, Romer and Weil (1992) used school enrollment rates as a proxy for human capital formation. More recently, years of schooling have been found more adequate (cf. for instance Bils and Klenow, 2000). Yet, even for OECD countries, the quality of schooling data is low (cf. Krueger and Lindahl, 2001, De la Fuente and Domenech, 2006).[20] Therefore, I will follow Barro and Sala-i-Martin's (1995: 422) and Gong, Greiner and Semmler's (2004) lead in using public expenditure on education as an alternative proxy for human capital formation.[21] The data source for public expenditure on education is the World Bank's educational statistics database (EdStats).[22] This database contains time series for the share of public education expenditure in GDP which, for most OECD countries, cover the period 1970 to 2005. Multiplying these shares with nominal GDP yields nominal education expenditure. After deflation by the GDP deflator, real expenditure is divided by population numbers. Finally, real per-capita education expenditure is transformed into five-year average growth rates, taking geometric means.

It is perhaps apposite to stress the symmetry between this measure and the variable that proxies physical capital formation. In both cases, capital formation is measured by deflated expenditure on the respective capital good. One might object that if there are inefficiencies in the educational system, education expenditure need not be closely correlated with human capital formation. Although that is true, similar objections could be leveled against using

[19] These were extracted from the OECD's National Accounts website (http://www.oecd.org/std/national-accounts).

[20] Cohen and Soto (2007) claim to have reduced the measurement error somewhat by better taking into account the age structure of the population.

[21] Actually, Barro and Sala-i-Martin (1995) use the share of public expenditure on education in GDP. But if we want to explain GDP with this variable, the fact that GDP appears as the denominator of the share will introduce a negative bias in the estimated coefficient (cf. also Cohen and Soto, 2007: 70). Therefore, five-year average growth rates of the investment variables rather than investment shares will be used. This also avoids the problem mentioned above in fn. 4.

[22] Cf. http://www.worldbank.org/education/edstats.

gross fixed investment to gauge physical capital formation. Nevertheless, this is the standard measure. Here, a similar measure will be chosen for human capital formation.[23]

Data for real per-capita education expenditure with a frequency of at least five years and starting point 1970 are available for 20 OECD countries from the World Bank database. These countries are *Australia*, *Austria*, *Canada*, *Denmark*, *Finland*, *France*, *Greece*, *Iceland*, *Ireland*, *Japan*, *South Korea*, *Luxemburg*, *Netherlands*, *New Zealand*, *Norway*, *Spain*, *Sweden*, *Switzerland*, the *United Kingdom*, and the *United States*. There are no data for Canada and Luxemburg for the last five-year period (2000-2005) so that the sample consists of 138 observations altogether. For studies working with five-year average data, this is a decent number.

There remains the problem, however, that the part of human capital that is formed by *private* education expenditure is excluded from the analysis. This exclusion is dictated by data availability: time-series data for private education expenditure back to 1970 do not exist. However, the OECD (2008) has recently published the split between public and private education expenditure for its member states in 2000 and 2005. Table B3.1 of the OECD study shows that next to all continental European countries in our sample had public shares in total education expenditure above 90 percent both in 2000 and in 2005.[24] All non-European countries and the UK, on the other hand, had much lower shares – with South Korea at the bottom end (58.9 percent in 2005). However, if the split between public and private education expenditure remains stable over time, public expenditure growth can still serve as a proxy for total expenditure growth in dynamic analysis. Low but stable public expenditure shares can be observed in South Korea and in the US. In the US, the share has remained at 67.3 percent between 2000 and 2005. There are two countries in our sample, however, where a marked shift from public to private education expenditure has occurred between 2000 and 2005 (and probably already earlier). These two countries are the UK, where the public share fell by 5.2 percentage points (PP) to 80.0 percent in 2005, and Canada, with a 4.4 PP drop to 75.5 percent.[25] For these two countries at least, public education expenditure growth will probably

[23] Vanhoudt (1998: 80) also proposes to “assume that all variables are accumulated in a similar way – i.e. by investing a fraction of foregone output”.

[24] Only Spain had a slightly lower share of 88.6 percent in 2005 (up from 87.4 in 2000). For Luxemburg, data are missing. For Norway and Switzerland (both of which had shares above 90 percent in 2000), 2005 data are missing.

[25] The public education expenditure share also declined in Japan (by 2.2 PP) and in Australia (by 1.9 PP). New Zealand had a public share of 78.4 percent in 2005. The 2000 share is unknown.

understate the true human capital formation. I will control for a possible bias due to the inclusion of the UK and Canada in a jackknifing exercise to be presented in the next section.

5. Empirical results

A reasonable first step in empirical analysis is a visual inspection of the data. Figures 1 to 3 show the histograms of the five-year average growth rates of real per-capita GDP, real per-capita gross fixed investment, and per-capita public education expenditure deflated by the GDP deflator for our sample of 20 OECD countries. All three variables exhibit large outliers. Ireland's per-capita GDP has grown by an amazing 8 ½ percent per year on average over the period 1995-2000 while its gross fixed investment growth surpassed 16 percent at the same time. On the other hand, Finland's per-capita real gross fixed investment almost halved over the period 1990-1995. Real public spending on education also shows a large positive and negative outlier, namely South Korea (1975-1980) and Luxemburg (1980-85), respectively. Outliers like these strengthen the case for carrying out the jackknifing exercise proposed in the previous section as a robustness test.

< *Insert Figures 1 to 3 around here* >

Another way to look at the data is to examine bi-variate pooled scatter graphs. Figures 4 to 6 show no pronounced contemporaneous association between public education expenditure growth and either GDP growth or fixed investment growth, but a strong positive association between fixed investment growth and GDP growth. Outliers become apparent here also.

< *Insert Figures 4 to 6 around here* >

As was mentioned in section 3, Granger-causality tests require stationary time series. Unfortunately, the available panel unit root tests are mainly designed for panels where both the time dimension and the cross section dimension are relatively large.[26] In panels such as ours with a time dimension of only 7 observations, the analysis can proceed only under restrictive assumptions like, for instance, dynamic homogeneity. This has to be kept in mind when interpreting the results of panel unit root tests reported in Table 1. As the table shows, the tests reject the null hypothesis of non-stationarity for all three variables. For what they are worth, these test results at least do not speak against proceeding to the Granger-causality tests.

[26] For a detailed account of panel unit root tests, cf. Breitung and Pesaran (2008).

< Insert Table 1 around here >

Since Granger-causality test results are sensitive to the choice of lag length *m* in the time-stationary VAR model given by equation (6), it is important to specify the lag structure appropriately. I follow Miyakoshi and Tsukuda (2004) and Atukeren (2007) in estimating equation (6) with OLS and basing the choice of the optimal lag length on the Schwarz Information Criterion (SIC). Table 2 shows that – based on this criterion – the optimal lag length is 2.

< Insert Table 2 around here >

Table 3 shows the results for estimating the VAR model (6) with OLS, with the Arellano-Bond one-step system GMM estimator and with the Arellano-Bond two-step system GMM estimator, respectively.[27] The OLS specifications include country-specific fixed effects, while the GMM specifications include period-specific effects (as is recommended in the literature). Lags of the dependent variable from at least two periods earlier as well as lags of the investment variables serve as GMM-style instruments.[28] For the two-step estimator, the small sample correction proposed by Windmeijer (2005) is implemented.

< Insert Table 3 around here >

The three bottom lines of the table report specification test results for the GMM estimations. The Sargan test is a test of the null hypothesis that the instruments are uncorrelated with the error term u_{it} (which they must be in order to be valid instruments). Table 3 shows that the null hypothesis is always accepted.[29] Note, however, that it was necessary to 'collapse' the set of instruments in order to achieve that the Sargan test accepts the over-identifying restrictions in the GMM estimations. While in the standard instrument matrix each instrumenting variable generates one column for each time period and lag available to that time period, Roodman

[27] The Arellano-Bond one-step estimator uses the identity matrix as a weighting matrix. The two-step estimator weighs the instruments asymptotically efficient using one-step estimates.

[28] Roodman's 'xtabond2' command was used in Stata (v. 9) for the GMM estimations; and Roodman's (2006) examples geared my handling of the syntax.

[29] The Sargan statistic, which is the minimized value of the one-step GMM criterion function, is not robust to heteroskedasticity or autocorrelation. The Hansen statistic (which is the minimized value of the two-step GMM criterion function) is robust.

(2006) proposes to ‘collapse’ the instrument set into a single column to limit the instrument count. This option is available in Stata (v. 9) and has been used here. The Arellano-Bond test of no second-order autocorrelation in the disturbances of the first differenced equation is used to detect first-order autocorrelation in the underlying level variables, which must not be present. The test accepts the null hypothesis at the five percent level.

The upper part of Table 3 reports the estimated coefficients, which can be used to draw inferences on whether economic growth is endogenous or rather exogenous. Recall that finding significantly positive coefficients would support the idea that growth is endogenous. Exogenous growth requires negative medium-term and no long-term Granger-causality running from physical and human capital accumulation to GDP growth.

The coefficients for lagged per-capita fixed investment are positive in the OLS estimation and negative in the GMM estimations. They are always insignificant, and the Wald test never rejects the hypothesis that the coefficients are jointly equal to zero. The human capital coefficients, on the other hand, are always positive and mostly significant. The Wald test always rejects the null hypothesis that the coefficients are jointly equal to zero.

Before we can accept the null hypothesis of no Granger-causality running from physical capital accumulation to GDP growth, we have to check the bi-variate setting though (cf. Atukeren, 2007). Table 4 shows that the bi-variate tests produce no evidence of Granger-causality either. The coefficients for lagged gross fixed investment growth remain positive in the OLS estimation, negative in the GMM estimations, and insignificant.

< *Insert Table 4 around here* >

With respect to physical capital accumulation, these findings cannot discriminate between Solovian exogenous growth theory and endogenous growth models of the AK type. Admittedly, the coefficients for the medium run (five-year) lag of fixed investment growth are negative in the GMM estimations, which speaks in favor of the exogenous growth model. However, since these coefficients are not statistically significant, no empirical support for the Solow model can be derived from them.

With respect to human capital, on the other hand, the results clearly support endogenous growth theory of the Uzawa-Lucas type, which posits a positive impact of human capital accumulation on long-term economic growth. Human capital formation through public education expenditure is found to Granger-cause per-capita GDP growth with a positive sign. Against the backdrop of the vivid debate over the growth effects of human capital accumulation, which has moved from great optimism (Lucas, 1988, Romer, 1990) to outspoken pessimism (Benhabib and Spiegel, 1994, Bils and Klenow, 2000, Pritchett, 2001), this result is in line with more recent studies that have vindicated a positive impact of human capital formation on long-term economic growth (Cohen and Soto, 2007, Arnold, Bassanini and Scarpetta, 2007).

There remains the possibility, however, that the results are driven by outliers. In order to check this, I re-estimate equation (6) with the Arellano-Bond one-step [AB(1)] estimator, dropping each of the 20 countries in turn. This robustness test yields some interesting results. Table 5 shows that the negative sign of the first lag of gross fixed investment growth is robust to the exclusion of countries from the sample. Furthermore, when certain countries – namely Canada, Denmark, the Netherlands, or the US – are dropped, the first lag becomes significant. The second lag of fixed investment growth, on the other hand, never becomes significant. It will be remembered that Solow-type exogenous growth theory predicts a significantly negative medium-term and an insignificant long-term coefficient for the fixed investment growth variable. The robustness test shows that our previous finding of an insignificant medium-term coefficient was only due to the inclusion of certain countries in the sample, whereas in the bulk of the economies the relationship predicted by exogenous growth theory in fact holds true. The robustness test thus points in the direction that – contrary to the prediction of AK-style endogenous growth models – long-term economic growth is not driven by physical capital accumulation.[30]

Things look quite different with regard to human capital. For the full sample, the conclusion was that long-term economic growth is driven by human capital formation. This finding is in principle confirmed by the robustness test. There is the case of South Korea, however. When South Korea is dropped from the sample, the coefficient for the first lag of public education expenditure growth – although it is still positive – becomes insignificant. The coefficient for the second lag remains significantly positive at the ten percent level, but the Wald test now marginally fails to reject the null hypothesis that the two coefficients are jointly insignificant. Apparently, the positive Granger-causation is exceptionally strong in

[30] Podrecca and Carmeci (2001) also find negative coefficients for lagged fixed investment – they use investment shares in GDP – in bi-variate tests for Granger-causality between fixed investment and GDP growth; and they also interpret this finding as evidence in favor of the Solow model and against the AK model. Podrecca and Carmeci do not control for contemporaneous investment, however, which is not in line with what the Solow model requires (see equation 5 above). Curiously, Podrecca and Carmeci find insignificant first (five-year) lags and significantly negative second (ten-year) lags for the fixed investment variable. The Solow model suggests the contrary.

Korea, which leads to an upward bias in the estimated coefficients for the full sample. Nevertheless, there is still a solid indication of positive Granger-causation in the sample excluding South Korea – the second lag of public education expenditure growth remains significant – so that there seems to be no need to revise the conclusions drawn from the full-sample estimations in the light of the robustness test.

< *Insert Table 5 around here* >

6. Conclusion

Around twenty years ago, endogenous growth theory emerged as an alternative to neoclassical exogenous growth theory. Endogenous growth models are appealing because they imply that economic growth depends on factors that policy has an impact on, like infrastructure capital, education, or R&D. But is endogenous growth theory also better in line with reality than its neoclassical counterpart? A body of empirical literature (that has been reviewed here for the first time) has confronted this question, yet no consensus has been reached either on which methods should be used to test endogenous vs. exogenous growth theory or on whether growth is actually endogenous or exogenous. Against this backdrop, this paper proposes a new test of endogenous vs. exogenous growth which draws on the fact that endogenous and exogenous growth theories make different predictions about the relationship between variations in physical or human capital formation on the one hand and GDP growth on the other hand over time. In the AK and Uzawa-Lucas endogenous growth models, an increase in the investment rates in physical or human capital, respectively, will raise the steady state GDP growth rate. Therefore, intervals of high (low) capital formation should antecede intervals of high (low) GDP growth. Exogenous growth theory makes the exact opposite prediction. As more investment can only increase GDP growth instantaneously, high (low) capital formation in one interval should cause GDP growth rates in consecutive intervals to decline (increase) as GDP growth moves back toward its steady state value. Over a longer time horizon, there should be no significant correlation between investment and economic growth according to neoclassical growth theory because growth is held to be independent of investment in the long run.

This paper applies the method of Granger-causality testing to a panel of 20 OECD countries to investigate the correlation between variations in spending on physical and human capital in the form of education on the one hand and subsequent GDP growth on the other hand. All data is transformed into five-year average growth rates, and the estimators chosen are Arellano-Bond type system GMM estimators. Although the results are not absolutely robust to the exclusion or inclusion of certain countries from respectively in the sample, the following picture emerges. The medium-term coefficients (five-year lag) for lagged physical capital growth are significantly negative while the long-term coefficients (ten-year lag) are insignificant. This finding favors Solow-type exogenous growth theory over AK-type endogenous growth models. With regard to human capital growth, things look quite different. The coefficients for both lags are positive and jointly significant, which means that human capital formation in the form of education Granger-causes GDP growth with a positive sign. This result lends support to human capital oriented endogenous growth models – like the Uzawa-Lucas model – rather than to the human capital augmented Solow model.

Acknowledgements

I would like to thank Erdal Atukeren for stimulating discussions on the topics of causality and, more narrowly, Granger-causality which have clearly improved my understanding of these issues. A special thanks to Axel Dreher and Martin Gassebner for their support with the Stata estimations. The usual disclaimer applies.

References

Aghion, P., Howitt, P., 1992. A model of growth through creative destruction. Econometrica 60, 323–351.

Aghion, P., Howitt, P., 1998. Endogenous Growth Theory. Cambridge, Mass.: MIT Press.

Arnold, J., Bassanini, A., Scarpetta, S., 2007. Solow or Lucas? Testing growth models using panel data from OECD countries. OECD Economics Department Working Papers, No. 592, Paris.

Arellano, M., Bond, S. R., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies 58, 277–297.

Arellano, M., Bover, O., 1995. Another look at the instrumental variables estimation of error-components models. Journal of Econometrics 68, 29–51.

Atukeren, E., 2007. A causal analysis of the R&D interactions between the EU and the US. The B. E. Global Economy Journal 7(4), Article 1.

Atukeren, E., 2008. Christmas cards, Easter bunnies, and Granger-causality. Quality and Quantity 42, 835–844.

Barro, R. J., 1991. Economic growth in a cross section of countries. Quarterly Journal of Economics 106, 407–443.

Barro, R. J., Sala-i-Martin, X., 1995. Economic Growth. New York: McGraw-Hill.

Bassanini, A., Scarpetta, S., 2002. Does human capital matter for growth in OECD countries? A pooled mean-group approach. Economics Letters 74, 399–405.

Benhabib, J., Spiegel, M. M., 1994. The role of human capital in economic development: Evidence from aggregate cross-country data. Journal of Monetary Economics 34, 143–173.

Bernanke, B. S., Gürkaynak, R. S., 2001. Is growth exogenous? Taking Mankiw, Romer, and Weil seriously. NBER Macroeconomics Annual 16, 11–57.

Bils, M., Klenow, P. J., 2000, Does schooling cause growth? American Economic Review 90, 1160–1183.

Bleaney, M., Gemmell, N., Kneller, R., 2001. Testing the endogenous growth model: Public expenditure, taxation, and growth over the long run. Canadian Journal of Economics 34, 36–57.

Blundell, R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87, 115–143.

Breitung, J., Pesaran, M. H., 2008. Unit roots and cointegration in panels. In: Matyas, L., Sevestre, P. (Eds): The Econometrics of Panel Data, 3rd ed. Berlin: Springer, 279–322.

Caselli, F., Esquivel, G., Lefort, F., 1996. Reopening the convergence debate: A new look at cross-country growth empirics. Journal of Economic Growth 1, 363–389.

Cashin, P., 1995. Government spending, taxes and economic growth. IMF Staff Papers 42, 237–269.

Cohen, D., Soto, M., 2007. Growth and human capital: Good data, good results. Journal of Economic Growth 12, 51–76.

De la Fuente, A., Domenech, R., 2006. Human capital in growth regressions: How much difference does data quality make? Journal of the European Economic Association 4, 1–36.

Devarajan, S., Swaroop, V., Zou, H.-F., 1996. The composition of public expenditure and economic growth. Journal of Monetary Economics 37, 313–344.

Dinopoulos, E., Thompson, P., 2000. Endogenous growth in a cross-section of countries. Journal of International Economics 51, 335–362.

Evans, P., 1996. Using cross-country variances to evaluate growth theories. Journal of Economic Dynamics and Control 20, 1027–1049.

Evans, P., 1998. Using panel data to evaluate growth theories. International Economic Review 39, 295–306.

Feenstra, R. C., Madani, D., Yang, T.-H., Liang, C.-Y., 1999. Testing endogenous growth in South Korea and Taiwan. Journal of Development Economics 60, 317–341.

Frankel, M., 1962. The production function in allocation and growth: A synthesis. American Economic Review 52, 995–1022.

Gong, G., Greiner, A., Semmler, W., 2004. The Uzawa-Lucas model without scale effects: Theory and empirical evidence. Structural Change and Economic Dynamics 15, 401–420.

Granger, C. W. J., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424–438.

Granger, C. W. J., 1980. Testing for causality: A personal viewpoint. Journal of Economic Dynamics and Control 2, 329–352.

Grier, K., Tullock, G., 1989. An empirical analysis of cross-national economic growth, 1951–80. Journal of Monetary Economics 24, 259–276.

Grossman, G. M., Helpman, E., 1991. Innovation and Growth in the Global Economy. Cambridge, Mass: MIT Press.

Grossman, G. M., Helpman, E., 1994. Endogenous innovation in the theory of growth. Journal of Economic Perspectives 8, 23–44.

Hartwig, J., 2009. Is health capital formation good for long-term economic growth? – Panel Granger-causality evidence for OECD countries. Journal of Macroeconomics (forthcoming).

Holtz-Eakin, D., Newey, W., Rosen, H. S., 1988. Estimating vector autoregressions with panel data. Econometrica 56, 1371–1395.

Hsiao, C., 1982. Autoregressive modeling and causal ordering of economic variables. Journal of Economic Dynamics and Control 4, 243–259.

Inklaar, R., Timmer, M. P., van Ark, B., 2008. Market services productivity across Europe and the US. Economic Policy 23, 141–194.

Islam, N., 1995. Growth empirics: A panel data approach. Quarterly Journal of Economics 110, 1127–1170.

Jones, C. I., 1995. Time series tests of endogenous growth models. Quarterly Journal of Economics 110, 495–525.

Kneller, R., Bleaney, M. F., Gemmell, N., 1999. Fiscal policy and growth: Evidence from OECD countries. Journal of Public Economics 74, 171–190.

Kocherlakota, N. R., Yi, K.-M., 1995. Can convergence regressions distinguish between exogenous and endogenous growth models? Economics Letters 49, 211–215.

Kocherlakota, N. R., Yi, K.-M., 1996. A simple time series test of endogenous vs. exogenous growth models: An application to the United States. Review of Economics and Statistics 78, 126–134.

Kocherlakota, N. R., Yi, K.-M., 1997. Is there endogenous long-run growth? Evidence from the United States and the United Kingdom. Journal of Money, Credit and Banking 29, 235–262.

Krueger, A. B., Lindahl, M., 2001. Education for growth: Why and for whom? Journal of Economic Literature 34, 1101–1136.

Li, D., 2002. Is the AK model still alive? The long-run relation between growth and investment re-examined. Canadian Journal of Economics 35, 92–114.

Lucas, R. E., 1988. On the mechanics of economic development. Journal of Monetary Economics 22, 3–42.

Mankiw, N. G., Romer, D., Weil, D. N., 1992. A contribution to the empirics of economic growth. Quarterly Journal of Economics 107, 407–437.

Mendoza, E. G., Milesi-Ferretti, G. M., Asea, P., 1997. On the effectiveness of tax policy in altering long-run growth: Harberger's superneutrality conjecture. Journal of Public Economics 66, 99–126.

Miyakoshi, T., Tsukuda, Y., 2004. The causes of the long stagnation in Japan. Applied Financial Economics 14, 113–120.

Nickell, S., 1981. Biases in dynamic models with fixed effects. Econometrica 49, 1417–1426.

OECD, 2008. Education at a Glance 2008: OECD Indicators. Organisation for Economic Cooperation and Development, Paris.

Podrecca, E., Carmeci, G., 2001. Fixed investment and economic growth: New results on causality. Applied Economics 33, 177–182.

Pritchett, L., 2001. Where has all the education gone? World Bank Economic Review 15, 367–391.

Roodman, D., 2006. How to do xtabond2: An introduction to 'difference' and 'system' GMM in Stata. Center for Global Development Working Paper, No. 103, Washington, D.C.

Romer, P. M., 1986. Increasing returns and long-run growth. Journal of Political Economy 94, 1002–1037.

Romer, P. M., 1990. Endogenous technological change. Journal of Political Economy 98, S71–S102.

Solow, R. M., 1956. A contribution to the theory of economic growth. Quarterly Journal of Economics 70, 65–94.

Solow, R. M., 1957. Technical change and the aggregate production function. Review of Economics and Statistics 39, 312–320.

Swan, T. W., 1956. Economic growth and capital accumulation. Economic Record 32, 334–361.

Temple, J., 1999. The new growth evidence. Journal of Economic Literature 37, 112–156.

Thurman, W. N., Fisher, M. E., 1988. Chickens, eggs, and causality, or which came first. American Journal of Agricultural Economics 70, 237–238.

Uzawa, H., 1965. Optimum technical change in an aggregative model of economic growth. International Economic Review 6, 18–31.

Vanhoudt, P., 1998. A fallacy in causality research on growth and capital accumulation. Economics Letters 60, 77–81.

Windmeijer, F., 2005. A finite sample correction for the variance of linear two-step GMM estimators. Journal of Econometrics 126, 25–51.

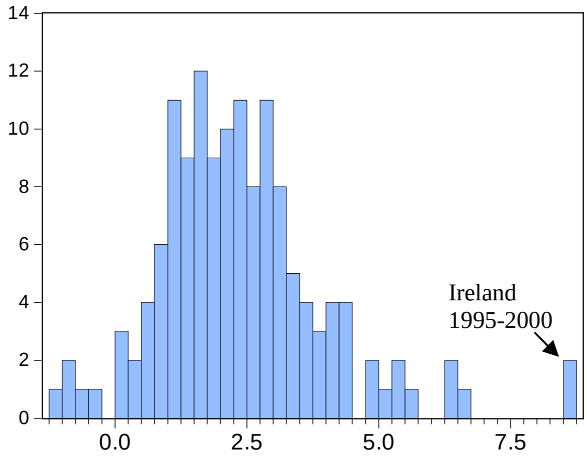

Figure 1: *Histogram of real per-capita gross domestic product growth rates for 20 OECD countries (five-year averages, 1970-2005)*

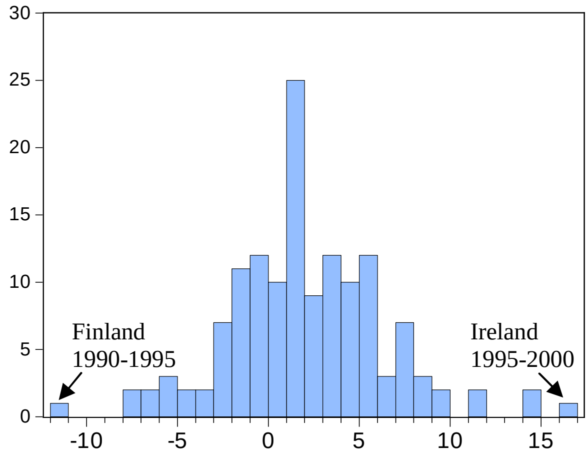

Figure 2: *Histogram of real per-capita gross fixed investment growth rates for 20 OECD countries (five-year averages, 1970-2005)*

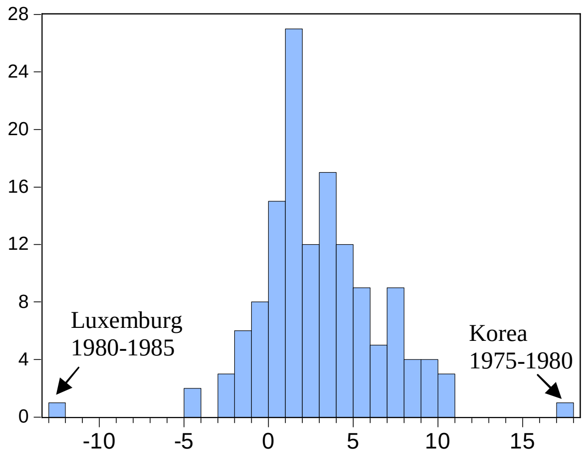

Figure 3: *Histogram of real per-capita public education expenditure growth rates for 20 OECD countries (five-year averages, 1970-2005)*

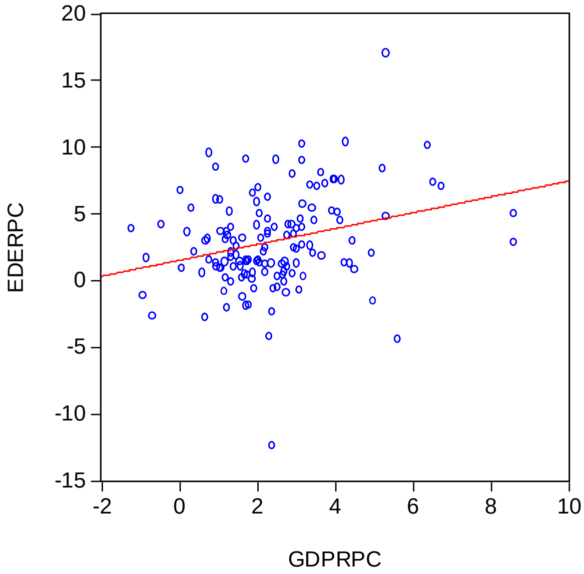

Figure 4: *Scatter graph of real per-capita gross domestic product (GDPRPC) growth rates versus real per-capita public education expenditure (EDERPC) growth rates for 20 OECD countries (five-year averages, 1970-2005)*

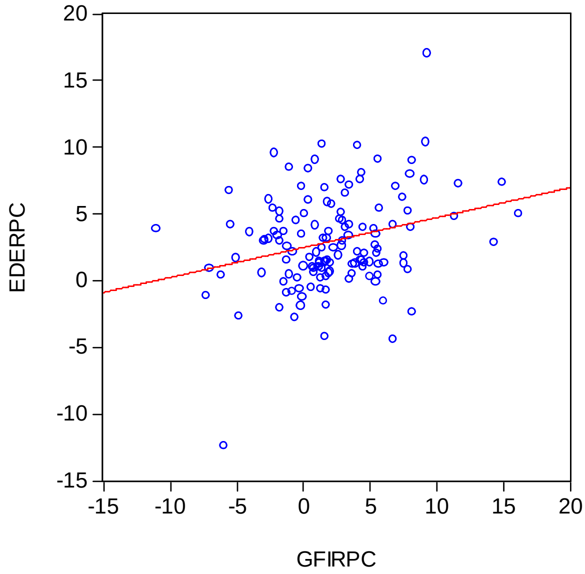

Figure 5: *Scatter graph of real per-capita gross fixed investment (GFIRPC) growth rates versus real per-capita public education expenditure (EDERPC) growth rates for 20 OECD countries (five-year averages, 1970-2005)*

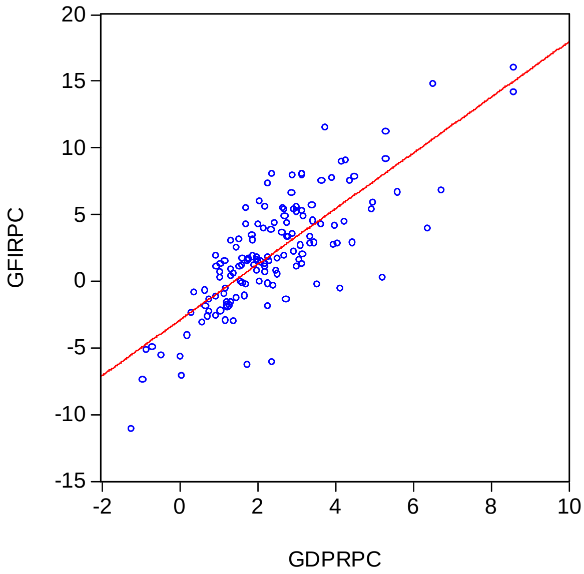

Figure 6: *Scatter graph of real per-capita gross domestic product (GDPRPC) growth rates versus real per-capita gross fixed investment (GFIRPC) growth rates for 20 OECD countries (five-year averages, 1970-2005)*

Table 1: Panel unit root test results (20 OECD countries, 1970-2005)

	GDPRPC			GFIRPC			EDERPC		
H_0: Unit root in level	Stat.	Prob.	Obs.	Stat.	Prob.	Obs.	Stat.	Prob.	Obs.
Levin, Lin & Chu t*	–11.738	0.000	120	–17.810	0.000	120	–15.358	0.000	118
Im, Pesaran & Shin W-stat	–4.299	0.000	120	–6.796	0.000	120	–4.765	0.000	118
ADF – Fisher Chi-square	96.658	0.000	120	128.900	0.000	120	101.543	0.000	118
PP – Fisher Chi-square	111.651	0.000	120	177.304	0.000	120	137.747	0.000	118

GDPRPC = real per-capita GDP growth rates (five-year averages), GFIRPC = real per-capita gross fixed investment growth rates (five-year averages), EDERPC = real per-capita public education expenditure growth rates (five-year averages)

Note: Individual intercepts are included as exogenous variables in the test equations. For the first three tests listed in the table, maximum lags are automatically selected based on the Schwarz Information Criterion. The remaining test uses the Bartlett kernel for the Newey-West bandwidth selection. The probabilities for the Fisher tests are computed using an asymptotic Chi-square distribution. The other tests assume asymptotic normality. EViews (v. 6) was used for the estimations.

Table 2: Optimal lag length for equation (6)

Lag	1	2	3
SIC	3.061	2.770	2.959

SIC = Schwarz Information Criterion

Table 3: Estimation results for equation (6)

	GDPRPC		
	(1) OLS	(2) Arellano-Bond one-step GMM	(3) Arellano-Bond two-step GMM
GDPRPC(-1)	–0.042	0.507***	0.353
	(0.098)	(0.109)	(0.597)
GDPRPC(-2)	–0.070	0.121	0.187
	(0.088)	(0.109)	(0.544)
GFIRPC	0.262***	0.236***	0.221*
	(0.017)	(0.050)	(0.108)
GFIRPC(-1)	0.004	–0.071	–0.026
	(0.036)	(0.044)	(0.212)
GFIRPC(-2)	0.021	–0.004	–0.008
	(0.030)	(0.038)	(0.061)
EDERPC	0.038	0.044	0.018
	(0.029)	(0.053)	(0.071)
EDERPC(-1)	0.066***	0.067***	0.048
	(0.024)	(0.025)	(0.039)
EDERPC(-2)	0.048**	0.056**	0.065*
	(0.022)	(0.026)	(0.034)
Number of obs.	98	98	98
Wald test (p-level) GFIRPC	0.779	0.265	0.972
Wald test (p-level) EDERPC	0.011	0.005	0.004
Sargan test (p-level)	–	0.356	0.356
Hansen test (p-level)	–	–	0.939
AB test (p-level)	–	0.058	0.559

GDPRPC = real per-capita GDP growth rates (five-year averages), GFIRPC = real per-capita gross fixed investment growth rates (five-year averages), EDERPC = real per-capita public education expenditure growth rates (five-year averages)

Standard errors are in parenthesis. *, ** and *** denote significance at the 10, 5 and 1 percent levels, respectively. Estimates for constant terms not shown. AB test = Arellano-Bond test for AR(2) in first differences.

Table 4: Bi-variate Granger causality tests

	GDPRPC		
	(4)	(5)	(5)
	OLS	Arellano-Bond one-step GMM	Arellano-Bond two-step GMM
GDPRPC(-1)	–0.069	0.445***	0.323*
	(0.100)	(0.133)	(0.158)
GDPRPC(-2)	–0.053	0.113	0.165*
	(0.088)	(0.101)	(0.081)
GFIRPC	0.267***	0.236***	0.265***
	(0.017)	(0.181)	(0.032)
GFIRPC(-1)	0.024	–0.055	–0.021
	(0.036)	(0.043)	(0.034)
GFIRPC(-2)	0.034	0.000	–0.011
	(0.030)	(0.035)	(0.025)
Number of obs.	100	100	100
Wald test (p-level)	0.505	0.404	0.770
Sargan test (p-level)	–	0.085	0.085
Hansen test (p-level)	–	–	0.642
AB test (p-level)	–	0.277	0.579

GDPRPC = real per-capita GDP growth rates (five-year averages), GFIRPC = real per-capita gross fixed investment growth rates (five-year averages)

Standard errors are in parenthesis. *, ** and *** denote significance at the 10, 5 and 1 percent levels, respectively. Estimates for constant terms not shown. AB test = Arellano-Bond test for AR(2) in first differences.

Table 5: Robustness test – Cross-national stability of parameters, AB(1) estimator: excluded countries

	GDPRPC									
	Australia	Austria	Canada	Denmark	Finland	France	Greece	Iceland	Ireland	Japan
GDPRPC(–1)	0.459*** (0.132)	0.505*** (0.138)	0.512*** (0.131)	0.534*** (0.141)	0.498*** (0.131)	0.494*** (0.136)	0.443*** (0.123)	0.494*** (0.211)	0.389*** (0.119)	0.468*** (0.122)
GDPRPC(-2)	0.136 (0.107)	0.110 (0.112)	0.124 (0.108)	0.103 (0.112)	0.122 (0.112)	0.118 (0.111)	0.080 (0.110)	0.026 (0129)	0.218* (0.113)	0.154 (0.101)
GFIRPC	0.258*** (0.048)	0.230*** (0.050)	0.236*** (0.046)	0.221*** (0.050)	0.237*** (0.048)	0.239*** (0.049)	0.225*** (0.047)	0.261*** (0.060)	0.328*** (0.044)	0.270*** (0.044)
GFIRPC(–1)	–0.057 (0.043)	–0.072 (0.045)	–0.075* (0.043)	–0.076* (0.045)	–0.060 (0.042)	–0.069 (0.044)	–0.067 (0.041)	–0.078 (0.069)	–0.007 (0.043)	–0.047 (0.041)
GFIRPC(–2)	–0.008 (0.038)	–0.003 (0.040)	–0.008 (0.039)	–0.001 (0.039)	–0.000 (0.041)	–0.005 (0.039)	0.009 (0.038)	0.013 (0.044)	–0.031 (0.040)	–0.010 (0.037)
EDERPC	0.035 (0.052)	0.048 (0.054)	0.048 (0.052)	0.046 (0.050)	0.044 (0.054)	0.044 (0.055)	0.061 (0.056)	0.043 (0.054)	–0.043 (0.046)	0.020 (0.049)
EDERPC(-1)	0.068*** (0.024)	0.066*** (0.025)	0.070*** (0.025)	0.069*** (0.025)	0.066*** (0.025)	0.068*** (0.025)	0.066*** (0.024)	0.079*** (0.024)	0.073*** (0.023)	0.061** (0.025)
EDERPC(-2)	0.056** (0.026)	0.060** (0.027)	0.057** (0.026)	0.052* (0.027)	0.059** (0.028)	0.056** (0.027)	0.053** (0.025)	0.044* (0.026)	0.039 (0.024)	0.051** (0.025)
Number of obs.	93	93	94	93	93	93	93	93	93	93
Wald test GFIRPC (p-level)	0.418	0.267	0.219	0.228	0.339	0.292	0.213	0.401	0.739	0.401
Wald test EDERPC (p-level)	0.004	0.005	0.004	0.007	0.006	0.006	0.005	0.003	0.004	0.013

Table 5 – cont'd

	GDPRPC									
	Korea	Luxemburg	Netherlands	New Zealand	Norway	Spain	Sweden	Switzerland	UK	US
GDPRPC(–1)	0.470*** (0.117)	0.407*** (0.121)	0.513*** (0.147)	0.434*** (0.111)	0.547*** (0.157)	0.456*** (0.133)	0.497*** (0.137)	0.481*** (0.135)	0.496*** (0.136)	0.523*** (0.133)
GDPRPC(-2)	0.091 (0.106)	0.191* (0.111)	0.170 (0.105)	0.143 (0.104)	0.126 (0.112)	0.084 (0.112)	0.095 (0.112)	0.095 (0.107)	0.114 (0.110)	0.137 (0.112)
GFIRPC	0.254*** (0.038)	0.257*** (0.046)	0.248*** (0.053)	0.282*** (0.045)	0.225*** (0.056)	0.242*** (0.050)	0.238*** (0.050)	0.246*** (0.049)	0.239*** (0.050)	0.237*** (0.051)
GFIRPC(–1)	–0.065 (0.040)	–0.045 (0.042)	–0.081* (0.047)	–0.058 (0.037)	–0.077 (0.049)	–0.056 (0.044)	–0.068 (0.045)	–0.063 (0.044)	–0.070 (0.045)	–0.089** (0.045)
GFIRPC(–2)	–0.012 (0.035)	–0.033 (0.036)	–0.030 (0.038)	–0.002 (0.036)	–0.002 (0.040)	0.008 (0.039)	0.000 (0.040)	0.000 (0.038)	–0.003 (0.039)	–0.019 (0.041)
EDERPC	–0.017 (0.047)	0.020 (0.064)	0.025 (0.050)	0.013 (0.053)	0.051 (0.054)	0.066 (0.058)	0.047 (0.052)	0.046 (0.051)	0.049 (0.055)	0.038 (0.055)
EDERPC(-1)	0.041 (0.027)	0.086*** (0.029)	0.076*** (0.024)	0.075*** (0.024)	0.055** (0.026)	0.069** (0.027)	0.067*** (0.025)	0.068*** (0.024)	0.069*** (0.025)	0.076*** (0.025)
EDERPC(-2)	0.047* (0.025)	0.037 (0.026)	0.055** (0.026)	0.045* (0.024)	0.056** (0.027)	0.063** (0.029)	0.061** (0.026)	0.057** (0.026)	0.057** (0.028)	0.053** (0.027)
Number of obs.	93	94	93	93	93	93	93	93	93	93
Wald test GFIRPC (p-level)	0.255	0.441	0.203	0.280	0.275	0.393	0.296	0.344	0.286	0.132
Wald test EDERPC (p-level)	0.105	0.008	0.002	0.003	0.016	0.013	0.003	0.004	0.006	0.002

GDPRPC = real per-capita GDP growth rates (five-year averages), GFIRPC = real per-capita gross fixed investment growth rates (five-year averages), EDERPC = real per-capita public education expenditure growth rates (five-year averages)
Standard errors are in parenthesis. *, ** and *** denote significance at the 10, 5 and 1 percent level, respectively. Estimates for constant terms not shown.