This is an Accepted Manuscript of an article published by Taylor & Francis in Journal of International Trade & Economic Development on 02/02/2020, available online: https://www.tandfonline.com/doi/full/10.1080/09638199.2019.1704843

Testing for causality between FDI and economic growth using heterogeneous panel data

Melisa Chanegriha
Middlesex University, London, UK

Chris Stewart
Kingston University, London, UK

Christopher Tsoukis
Keele University, Staffordshire, UK

Abstract

The causal relationship between FDI inflows and growth is of great policy interest, yet the state of concrete knowledge on the issue is rather poor. Our contribution is to investigate the causal relationship between the ratio of FDI to GDP (FDIG) and economic growth (GDPG) using a battery of cutting-edge methods and an extensive data set. We employ the heterogeneous-panel tests of the Granger non-causality hypothesis based on the works of Hurlin (2004a), Fisher (1932, 1948) and Hanck (2013). Our panel data set is compiled from 136 developed and developing countries over the 1970-2006 period. According to the Hurlin and Fisher tests, FDIG unambiguously Granger-causes GDPG for at least one country. However, the results from these tests are ambiguous regarding whether GDPG Granger-causes FDIG for at least one country. Using a test based upon Hanck (2013), both with and without one structural break in the vector autoregression, we are able to determine whether and for which countries there is Granger-causality. This test suggests that at most there are six countries (Estonia, Guyana, Poland, Switzerland, Tajikistan and Yemen) where FDIG Granger-causes GDPG and at most four countries (Dominican Republic, Gabon, Madagascar and Poland) where GDPG Granger-causes FDIG.

JEL classification: C33; F21; F43; N10;

Keywords: Granger-causality Tests; Panel Data; FDI; Economic Growth.

Addresses:
Chanegriha: Economics and International Development, Business School, Middlesex University, Hendon campus, The Burroughs, London NW4 4BT, UK; m.chanegriha@mdx.ac.uk.
Stewart: School of Law, Social and Behavioural Sciences, Kingston University, Penrhyn Road, Kingston upon Thames, Surrey KT1 2EE, UK; c.stewart@kingston.ac.uk.
Tsoukis: Economics and Finance, Keele Management School, Keele University, Staffordshire, ST5 5BG, UK; email: c.tsoukis@keele.ac.uk

1. Introduction

With the sharp rise in Foreign Direct Investment (FDI) flows since the 1990s, questions have arisen as to its relation to host countries' output and growth (Chowdhary and Mavrotas, 2005; Ghosh and Wang, 2009). A range of analyses have emphasised the beneficial effects of incoming FDI: It can potentially contribute to economic growth through new capital investment, technology transfer, development of human capital and skills, integration into global economic networks and strengthening of the competitive environment in a host country (De Mello, 1997, 1999; Blomstrom et al., 1992, 1996; Borensztein et al., 1998).[1]At the same time, the host country's GDP and market size is one of the key determinants of incoming FDI itself (Chanegriha, Stewart and Tsoukis, 2017). Understanding the direction of causality between GDP and FDI is crucial for formulating public policies that encourage private investors in developing countries. A finding that FDI has a positive impact on growth would imply that policy makers should focus on policies that have been shown to promote FDI such as school attainment, openness to international trade, lower taxes and inflation (Asiedu, 2002; Chakrabarti, 2001; Chanegriha, Stewart and Tsoukis, 2017); whereas, if FDI does not cause growth, such policies would need to be reconsidered. In terms of theory, a non-causality result would also cast doubt on the validity of the theories that have stressed the beneficial effects of FDI for the host country. While there is a pool of empirical studies regarding the relationship between FDI and economic growth, discussed below, the results are mixed. We still concur with Caves (1996) who early on suggested that "the causal relationship between FDI and economic growth is a matter on which we totally lack trustworthy conclusions". This is an unsatisfactory state of affairs for both theory and public policy.

This paper tests the direction of causality between FDI and economic growth. Our work contributes to the literature in the following ways. First, we apply the tests to a larger panel of countries than previously considered in the literature. Our panel analysis uses pooled data from 136 developed and developing countries for 1970–2006. Existing studies that test Granger non-causality (GNC) between GDP and FDI apply this test on time-series data for a single or small group of countries. By contrast, this paper analyses pooled data for a large number of countries over a relatively long period to exploit both cross-sectional and time-series dimensions of the

[1] There may also exist drawbacks for the host country, e.g. a deterioration of the trade balance (the flip side of the improvement of the capital account) and crowding out of domestic investment.

data. Second, in addition to applying standard time-series GNC tests, we also apply a battery of panel GNC tests by utilising recent advances in the relevant methodology. These include the traditional Fisher (1932, 1948) method and the recent Hurlin (2004a) test. We are not aware of any previous application of Hurlin's (2004a) method to the causality between FDI and growth in the literature. We also adapt the panel method applied by Hanck (2013) within the context of unit root testing to test for GNC. This panel method is robust in the face of cross-sectional dependence and can identify which individual units (countries) reject the null hypothesis of interest and those that do not. We are not aware of any previous application of this method to GNC testing. The battery of tests and the large sample aim at obtaining an holistic view and are both motivated by the conflicting results in the extant literature. Finally, in all panel tests that we employ, we allow for the least restrictive specification, thus avoiding erroneous general inferences.

Empirical work on the FDI-growth relationship has utilised a variety of samples, methodologies and conditioning factors (e.g., financial markets, technological development, openness, regulatory environment, human capital, labour markets and more). The studies may be grouped into three categories according to their results. The first category finds a positive unconditional effect of FDI on growth – Blomstrom et al. (1996); Gao (2001) and Lensink and Morrissey (2006). The second finds an ambiguous role for FDI alone on economic growth and find an important role for various conditioning factors that promote the beneficial effects of FDI – Borensztein et al. (1998), Campos and Kinoshita (2002), OECD (2002), Alfaro et al. (2004), Busse and Groizard (2008), Agrawal (2000). The problem with this class of studies is that they do not reach any consensus as to the most important conditioning, or facilitating, factors. A third category does not find any positive effect of FDI on growth, even taking into account conditioning factors as above – Carkovic and Levine (2005) and Mencinger (2003).[2] Thus, all considered, the lack of any robust conclusions is the only safe conclusion on the FDI-growth relationship. In addition, the role of economic growth as an important determinant of FDI inflows into host countries, mentioned above, suggests a possible dual causality of FDI to growth (Choe, 2003).

[2] See Chowdhury and Mavrotas (2005) and Ozturk (2007) for surveys of the FDI and growth relationship. Mody and Murshid (2002) discusses the relationship between domestic investment and FDI. See Asiedu (2003) for an excellent discussion of the relationship between policy reforms and FDI in the case of Africa. Gorg and Greenaway (2004) analyse the effects of FDI on domestic firms.

A fourth strand of literature investigates Granger-causality between GDP and FDI. These causality tests have also failed to reach unanimous conclusions. There seem to be those that find causality to run (mostly) from FDI to GDP, such as Chan (2000), Dutt, Duttaray and Mukhopadhyay (2008), Zang (2001), OECD (2002). However, the strength of the causal effect varies considerably, as do the conditioning factors. Other studies report reverse causality, from GDP to FDI, e.g. Chakraborty and Basu (2002), Choe (2003), Ozturk and Kalyoncu (2007), Sooreea-Bheemul and Sooreea (2013). Again, the details vary considerably, e.g. some may find mixed results across different countries, or bi-directional causality with one direction more prominent, etc. Yet others find no significant causality (Ericsson and Irandoust, 2001), or very mixed results (Gursoy et al., 2013) or even negative causality (Mencinger, 2003, found a negative causal relationship between FDI and GDP implying that FDI hampered the real convergence of Eastern European countries with the rest of the EU). One conclusion that may follow from such disparate results is the need to continue testing by employing larger data sets and more general methods; this motivates our study.

In addition to diverse country experiences and samples, differences in empirical methods may account for such discrepancies in the results. Various criticisms of the empirical approaches include the following. First, the use of time averaged data, resulting in loss of information and bias (Greene, 2000). Second, the reliance on GDP growth rates, i.e. first differences, resulting in misleading inferences regarding long run relationships (Ericsson et al., 2001). Third, the potential of endogeneity bias resulting from reverse causality (see Parsons and Titman, 2007).

Our methodology applies panel GNC tests to exploit the enhanced power of panel data methods. These methods are based on Fisher (1948), Hurlin (2004a), and Hanck (2013). Endogeneity is not an issue in our causality tests as the regressors are all lagged variables. Further, we do not average our data and therefore avoid the issues associated with this. In our analysis we assume that FDI/GDP and GDP growth are stationary. In the former case, we do not expect FDI and GDP to diverge without bound while in the latter we believe that GDP growth is intrinsically stationary. Owing to the relatively short time-series for many countries we cannot consider error correction models and so limit the analysis to two stationary series. Finally, our use of panel data should help increase the power of the tests.[3]

[3] In our estimation we do not distinguish between developed and underdeveloped countries because Hanck's (2013) method allows us to identify whether Granger-causality exists for each individual country.

The rest of the paper is organised as follows. Section 2 provides a review of the literature while Section 3 outlines the econometric methodology and data. Section 4 presents and discusses the empirical results; we conclude in Section 5.

2. Review of the Literature

Neoclassical models of growth as well as endogenous growth models provide the basis for most empirical work on the FDI–growth relationship. The relationship has been studied by explaining four main channels: (i) determinants of growth, (ii) determinants of FDI, (iii) the role of MNCs in host countries and (iv) the direction of causality between FDI and growth (Chowdhury and Mavrotas, 2005).

According to neoclassical growth theory, economic growth generally comes from two sources: factor accumulation and total factor productivity (TFP) growth (Felipe, 1997). Within this theory, technological progress and labour growth are exogenous to the economic system and inward FDI merely increases the investment rate, leading to a transitional increase in per capita income growth without any long-run growth effect (Hsiao, 2006). The extent to which FDI affects output growth is limited by the assumptions of the model (Sass, 2003). The potential impact of FDI on growth is only in the short run, the magnitude and duration of which depends on the transitional dynamics to the steady-state growth path. FDI only exerts an effect on the level, not the growth rate, of per capita output, that is, it does not alter the growth of output in the long-run (Calvo and Robles, 2003).

In the framework of endogenous growth models, three main channels can be detected through which FDI affects growth. First, FDI increases capital accumulation in the receiving country by introducing new inputs and technologies (Blomstrom et al., 1996; Borensztein et al., 1998). In the case of new technologies, FDI is expected to be a potential source of productivity gains via spillover to domestic firms. Empirically, Blomstrom et al. (1996) found that positive growth effects are caused by increasing FDI using FDI inflows in a developing country as a measure of its interchange with other countries. They also found that FDI has a significant effect on promoting growth in exporting countries rather than in importing countries. This implies that the impact of FDI varies across countries and trade policy can affect the role of FDI in economic growth.

Second, FDI raises the level of knowledge and skills in the host country through labour and manager training (De Mello, 1997, 1999). Influenced by Mankiew et al.'s (1992) pioneering research, most recent empirical models have added education to the standard growth equation as a proxy for human capital. Borensztein et al. (1998) suggested that the level of human capital determines the ability to adopt foreign technology. Thus, larger endowments of human capital are assumed to induce higher growth rates given the amount of FDI. They suggested further that countries might need a minimum threshold stock of human capital in order to experience the positive effects of FDI.

Bengoa and Sanchez-Robles (2003) showed that FDI is positively correlated with economic growth, however host countries require human capital, economic stability and liberalised markets in order to benefit from long-term FDI inflows. Developed countries are expected to have a higher level of human capital and thereby to benefit more from FDI than developing countries. This seems to be confirmed by Xu (2000).

Third, FDI increases competition in the host country's industry by overcoming entry barriers and reducing the market power of existing firms. As a consequence of endogenous growth theory, FDI has a newly perceived potential role in the growth process. In the new theory of economic growth FDI not only affects the level of output per capita it also influences its growth rate. This literature has developed various hypotheses to explain why FDI may enhance the growth rate of per capita income in the host country (Calvo and Robles, 2003).

A consensus has been that FDI tends to have a significant effect in promoting economic growth through multiple channels such as capital formation, technology transfer and spillover and human capital enhancement – see Barro and Sala-i-Martin (1995) and De Mello (1997). Econometric models of endogenous growth have been combined with studies of the diffusion of technology in an attempt to show the effect of FDI on the economic growth of several economies (Lucas, 1988; Barro,1991). In these models, technology plays an important role in economic development. As a result, and in contrast to the neoclassical theory, monetary and fiscal policies are deemed to play a substantive role in advancing growth in the long-run.

Two strands of research have emerged, one that discusses the effects of FDI on economic growth and the other that recognises these effects and subsequently tries to identify the determinants of FDI flows to receiving countries. The possibility of a two-way causality between FDI and a host country's economic growth identifies a third line of research in the FDI literature (Choe, 2003). Countries with fast economic growth generate more demand for FDI and offer opportunities for making profits. By contrast, inward FDI flows may enhance growth through positive direct and indirect effects on variables that affect growth. This suggests bidirectional causality between FDI and growth. Despite the considerable volume of research on the subject, there is conflicting evidence on the (dual) direction of causality between FDI and economic growth. We seek to provide evidence on this issue for a large number of countries using time-series and panels data methods, which we outline in the next section.

3. Econometric Methodology

We test for GNC (Granger, 1969, 1980) between two variables, the FDI-GDP ratio and GDP growth, using heterogeneous panel data. First, we apply standard time-series GNC tests for each country. Second, our panel tests are based upon pooling the time-series results to exploit the panel properties of data and allow the coefficients to vary across countries. Within this broad framework, we apply three panel GNC tests developed by Hurlin and Venet (2001) and Hurlin (2004a); Fisher (1948); and Hanck (2013). These panel tests develop Holtz-Eakin et al's (1988) method by allowing the coefficients to be different across sections. We consider the most general case of heterogeneous slopes and intercepts, thus avoiding the pitfall of making erroneous generalised inferences across the entire cross-section which might in fact be true only in a subset of countries (Hood and Irwin, 2006).

3.1.1 The Hurlin method

Hurlin and Venet (2001) and Hurlin (2004a, 2004b, 2008) and develop Granger-causality tests to take into account cross-sectional heterogeneity in panel data (unbalanced or balanced). Hence, they distinguish between the heterogeneous non-causality (HENC) and homogeneous non-causality (HNC) hypotheses. Hurlin and Venet (2001) and Hurlin (2008) consider two covariance stationary variables, denoted $x_{i,t}$ and $y_{i,t}$, observed on $t = 1,2, \ldots, T_i$ periods and $i = 1,2, \ldots., N$ individuals (where for a balanced panel $T_i = T$) in a linear bivariate heterogeneous panel vector autoregression (VAR) of the following form:

$$x_{i,t} = \alpha_i + \sum_{H=1}^{H_i} \gamma_i{}^{(H)} x_{i,t-H} + \sum_{H=1}^{H_i} \beta_i^{(H)} y_{i,t-H} + \varepsilon_{i,t} \tag{1}$$

The lag-length H_i can be different for different cross-sectional units, however, when $H_i = H$ the lag-lengths are identical for every cross-section. Individual coefficients, α_i, are considered fixed while the slope coefficients, $\gamma_i{}^{(H)}$ and $\beta_i^{(H)}$, vary across units. Equation (1) is estimated by ordinary least squares (OLS) for each cross-sectional unit. The time-series GNC null for each individual unit is $\beta_i^{(1)} = \cdots = \beta_i^{(H_i)}$ and can be tested using the standard time-series F-statistic, F_i, which has an $F(H_i, T_i - 2H_i - 1)$ distribution. Hurlin and Venet (2001, p. 14) demonstrate that the corresponding time-series Wald statistic is $W_i = H_i F_i$ that asymptotically (as $T_i \to \infty$) has a $\chi^2(H_i)$ distribution.

The corresponding null hypothesis (H_0) for the whole panel is homogeneous non-causality (HNC), which is expressed as:

$$H_0\colon \beta_i^{(1)} = \cdots = \beta_i^{(H_i)} = 0, \qquad \forall\, i \tag{2}$$

The alternative hypothesis (H_1) is that $y_{i,t}$ Granger-causes $x_{i,t}$ for at least one cross-section. That is, there are N_1 $(< N)$ individual units with no causality from $y_{i,t}$ to $x_{i,t}$, and $N - N_1$ individuals where $y_{i,t}$ Granger-causes $x_{i,t}$, thus:

$$H_1\colon \begin{cases} \beta_i^{(1)} = \cdots = \beta_i^{(H_i)} = 0, & i = 1, \dots, N_1 \\ \beta_i^{(1)} \neq 0 \cup \dots \cup = \beta_i^{(H_i)} \neq 0 & i = N_1 + 1, \dots, N \end{cases} \tag{3}$$

When $0 < N_1 < N$ the causality relationship is heterogeneous across individual units.

Hurlin (2004a, p. 14) demonstrates that provided $T_i > 5 + 2H_i$ the following normalised average Wald statistic has a standard normal distribution as N tends to infinity and is appropriate for fixed (small) T (semi-asymptotic):[4]

[4] The panel test statistic is not always positive, although it is based on individual Wald statistics that are all positive, because the expected value of these statistics is subtracted in constructing the normalised Z statistics. Nevertheless, the test is one-tailed because only very small values of Wald statistics will fall in the extreme left-hand tail and these will indicate non-rejection of the null. Hence, the rejection region only occurs in the right-hand tail. For extensive and full derivations of asymptotic and semi-asymptotic distributions see Hurlin (2008).

$$\tilde{Z}_{N;T}^{HNC} = \sqrt{N}\left[W_{N,T}^{HNC}(\varphi) - N^{-1}\sum_{i=1}^{N} H_i \times \frac{(T_i-2H_i-1)}{(T_i-2H_i-3)}\right] \times \left[N^{-1}\sum_{i=1}^{N} 2H_i \times \frac{(T_i-2H_i-1)^2\times(T_i-2H_i-3)}{(T_i-2H_i-3)^2\times(T_i-2H_i-5)}\right]^{-\frac{1}{2}} \tag{4}$$

where,

$$W_{N,T}^{HNC}(\varphi) = \frac{1}{N}\sum_{i=1}^{N} W_{i,T} \tag{5}$$

such that $W_{i,T}$ is the Wald statistic for cross-section i.

The above statistic, $\tilde{Z}_{N;T}^{HNC}$, is appropriate when the panel is unbalanced (T_i varies across units) and the lag lengths (H_i) in each cross-section's VAR are different.[5]

When $\tilde{Z}_{N;T}^{HNC}$ exceeds its critical value the HNC null is rejected and the alternative that at least one cross-sectional unit exhibits Granger-causality (GC), cannot be rejected. Otherwise, the HNC null cannot be rejected and all cross-sectional units satisfy GNC.

Hurlin (2008, pp. 15 – 17) reports Monte Carlo simulation experiments that demonstrate that the semi-asymptotic panel statistic (fixed T and large N), $\tilde{Z}_{N;T}^{HNC}$, is virtually correctly sized for *all* values of T and N. Further, these semi-asymptotic panel statistics, which are approximately correctly sized, exhibit substantially greater power than the Wald statistics that are calculated for a single time-series. This is true even when N is small.[6] "This improvement in power can be intuitively understood as follows. Individual statistics are bounded from below (by zero) but may take arbitrarily large values. Hence, when averaging among individual Wald statistics, the 'abnormal' realisations (realisations below the chi-squared critical value) are annihilated by the realisations on the true side (large)." Hurlin (2008, p. 16). The power of the panel statistic is slightly lower when there is Granger-causality for some cross-sectional units in the panel and not others. Nevertheless, power is regarded as "reasonable" even when T and N are small and when there is causality for only a very small percentage of cross-sections (which is the worse-

[5] When the panel is balanced, and the lag lengths are the same in each cross-section's VAR a simplified panel test statistic may be employed – see Hurlin (2008).

[6] This is suggested to be true, for example, when $N = 5$. This is so even when the time-series is around 50 observations, a typical size for annual macroeconomic time-series.

case scenario in terms of power).

When T and N are very small there is some slight size distortion as H rises which means that the statistics are not very near to the standard normal distribution and critical values from this distribution can be improved upon. Hurlin (2008, p. 18) suggests that the critical value, $\tilde{C}_{N,T}(\alpha)$, for the semi-asymptotic panel statistics (based on a balanced panel and constant H across sections) can be approximated by the following expression:

$$\tilde{C}_{N,T}(\alpha) = Z_\alpha \times \frac{(T-2H-1)}{(T-2H-1)} \times \sqrt{\frac{2H}{N} \times \frac{(T-2H-3)}{(T-2H-5)}} + \frac{H \times (T-2H-1)}{(T-2H-3)} \tag{6}$$

where, Z_α is the critical value taken from the standard normal distribution for the α level of significance.

Hurlin (2004a) does not provide the formula for calculating critical values when the panel is unbalanced (T_i varies across units) *and* the lag lengths (H_i) in each cross-section's VAR are different (for the panel statistic given by (4)). However, from the equations reported in Hurlin (2004a) as (13), (14), (17) and (20) the formula for obtaining the critical value can be obtained as:[7]

$$\tilde{C}_{N,T}(\alpha) = Z_\alpha \sqrt{N^{-2} \sum_{i=1}^{N} \left[2H_i \times \frac{(T_i-2H_i-1)^2 \times (T_i-2H_i-3)}{(T_i-2H_i-3)^2 \times (T_i-2H_i-5)}\right]} + N^{-1} \sum_{i=1}^{N} \left[H_i \times \frac{(T_i-2H_i-1)}{(T_i-2H_i-3)}\right] \tag{7}$$

This panel GNC test's advantages include improved efficiency due to the increased sample size of the test and substantially greater power compared to its time-series counterpart (even for small T and N).[8] The testing procedure is simple to implement being based on averages of

[7] Alternatively, one can group countries into the value of N, T_i and H_i used in the test and identify the critical value appropriate for each group using (6). To obtain the critical value for the whole panel one can take the weighted average of these group critical values where the weights reflect the proportion of cross-sectional units from the whole panel appearing in each group.

[8] Hurlin and Venet (2008, p. 11) provide the following commentary within the context of bivariate GNC tests between financial development and GDP growth. "What is the main advantage of this Granger non-causality test? For instance, let us assume that there is no causality from financial development to growth for all of the N countries. Given the Wald statistics properties in small sample[s], the analysis based on N individual tests is likely to be inconclusive. With a small T sample, some of the realizations of the individual Wald statistics are likely to be superior to the asymptotic critical values of the chi-square distribution. These 'large' values of individual statistics lead to wrongly reject the null hypothesis of non-causality for at least some countries. The conclusions are then no[t] clear cut. On the contrary, in our panel average statistic, these "large" values of individual Wald statistics are crushed by the others which converge in probability to zero. When N tends to infinity, the cross-

Wald statistics obtained from time-series regressions and the model allows for heterogeneity in all coefficients across the individual units and for heterogeneity in terms of which cross-sections exhibit GNC. The two main drawbacks of this procedure are as follows. “Firstly, the rejection of the null of Homogeneous Non-causality does not provide any guidance as to the number or the identity of the particular members for which the null of non-causality is rejected. Secondly, the asymptotic distribution of our statistics is established under the assumption of cross-section independence. As for panel unit root tests, it is now necessary to develop second generation panel non-causality tests that allow for general or specific cross-section dependencies. This is precisely our objective for future researches.” Hurlin (2008, p. 20). Based on Hanck (2013) we present a procedure that addresses both of these drawbacks below.

3.1.2 The Fisher method

The Fisher panel test (1932, 1948), denoted λ, is:

$$\lambda = -2\sum_{i=1}^{N} ln(p_i) \sim \chi^2(2N) \qquad (8)$$

where p_i is the probability value for the F or Wald test for (in the current context) the GNC null for the ith cross-sectional unit and ln denotes the natural logarithm operator. This tests the null hypothesis of GNC for all N cross-sections against the alternative that there is Granger-causality for at least one individual unit. If the λ exceeds the critical value given by the χ^2 distribution with $2N$ degrees of freedom the null is rejected.

This test is subject to the same criticisms as the Hurlin method because it does not account for any cross-sectional dependence and that when the null is rejected it does not indicate for how many or which cross-sectional units the null is rejected for.

3.1.3 The Hanck (2013) method

Hanck (2013) proposed an intersection panel unit root test, making use of earlier work by Simes (1986) and Hommel (1988). The test is robust to general patterns of cross-sectional

sectional average is likely to converge to zero. The null hypothesis of [the] homogeneous non-causality hypothesis will not be rejected.” Our comments are given in squared parentheses.

dependence, is straightforward to implement and can identify which cross-sectional units in the panel reject the null and which do not.[9] However, we apply this intersection test within the context of GNC (rather than unit roots). We are not aware of this procedure being applied within the context of GNC tests. This can be justified because the procedure is based on probability values from time-series tests and is not restricted to any specific class of tests.[10]

Within the GNC context the Simes-type panel test is based upon the estimated time-series equations for each cross-sectional unit as specified by (1). The HNC null hypothesis is re-expressed as follows:

$$H_0 = \cap_{i=1,2,\dots,N} H_{i,0} \tag{9}$$

where, $\cap_{i=1,2,\dots,N}$ denotes the intersection over the individual cross-sectional units for $i = 1, 2, \dots, N$ and H_{i0}: $\beta_i^{(1)} = \cdots = \beta_i^{(H_i)} = 0$ for one particular i. If the null is rejected there is at least one cross-section that exhibits Granger-causality (GC), that is:

$$H_1 = \cup_{i=1,2,\dots,N} H_{i,1} \tag{10}$$

where, $\cup_{i=1,2,\dots,N}$ denotes the union over the individual cross-sectional units for $i = 1, 2, \dots,$ N and $H_{i,1}$: $\beta_i^{(1)} \neq 0 \cup \dots \cup \beta_i^{(H_i)} \neq 0$ for one particular i.

The test is based upon the probability values, p_i, of time-series F or Wald GNC tests for the null $H_{i,0}$ obtained from the estimation of equation (1) for each of the i cross-sectional units. These N probability values are arranged in ascending order, thus, $p_1 \leq p_2 \leq \cdots \leq p_N$, where p_1 is associated with the cross-sectional unit that is most likely to reject $H_{i,0}$.

[9] In being able to account for general forms of cross-sectional dependence Hanck (2008) argues that this has advantages over many second-generation panel unit root tests where non-trivial decisions are required by the user in the implementation of the tests that may affect the outcome. Such decisions are not required in the application of the intersection unit root test. Hanck (2008, pp. 4 – 5) shows that the intersection test controls size for patterns of cross-sectional dependence often assumed in panel models with dynamics.

[10] The procedure is appropriate for probability values based on test statistics that are multivariate totally positive of order two. This contains a large class of distributions including the absolute valued multivariate normal, absolute valued central multivariate t and central multivariate F, see Hanck (2013). Given that GNC tests can be based on t, F and chi-squared distributions this would make this an appropriate test for use with Hanck's (2013) procedure.

The intersection test rejects the null for any individual cross-section in the panel at the α level of significance only if the following condition holds:

$$p_j \leq \frac{j\alpha}{N} \qquad \text{for some } j = 1, 2, \dots, N \tag{11}$$

The N ordered probability values are compared with ever increasing critical points, defined by $\frac{j\alpha}{N}$, and if at least one p_j exceeds its critical point the null is rejected for the whole panel (hence, at least one cross-section exhibits GC) otherwise GNC is inferred for all individual units.

To identify which individual cross-sections in the panel reject, or fail to reject, the GNC null, we follow Hanck (2013) in applying Hommel's (1988) procedure. The first step is to calculate r such that the following condition holds (for all q for a given i):

$$r = max\left\{p_{(N-i+q)} > \frac{q\alpha}{i}\right\} \qquad \text{for } q = 1, 2, \dots, i \qquad \text{where} \qquad i = 1, 2, \dots, N \tag{12}$$

The second step is to use r to determine which cross-sections reject the GNC null and which do not. In particular, if $r = 0$ the GNC null is rejected for all cross-sectional units – $H_{i,0}$ is rejected for all i. Whereas if $r > 0$, reject the GNC null for all cross-sectional units where $p_j \leq \frac{\alpha}{r}$ and do not reject the null for all units where this condition is not satisfied.[11]

This panel GNC testing approach is referred to as the Hanck (2013) GNC intersection test. The ability of the Hanck (2013) procedure to identify which countries exhibit GNC and which do not and to deal with cross-sectional dependence should make its inference superior to that obtained from the Hurlin and Fisher panel tests. The panel nature of the Hanck (2013) procedure should make its inference superior to that of time-series tests, too.

3.1.4 The Hanck (2013) method allowing for structural breaks

To allow for the possibility of a single structural break for each country we augment the

[11] In identifying which cross-sectional units in the panel reject the null and those which do not using a large number of tests Hommel (1998) proves that the above procedure controls for the "Familywise Error Rate" (FWER). That is, in choosing the level of significance for an individual test to be α, the above procedure ensures that the size of the test for at least one unit's $H_{i,0}$ is α.

bivariate VAR specified by equation (1) to allow all coefficients to shift at a break date set in the middle of the time-series sample, denoted T_i^{Yr}, thus:[12]

$$x_{i,t} = \alpha_i + \sum_{H=1}^{H_i} \gamma_i^{(H)} x_{i,t-H} + \sum_{H=1}^{H_i} \beta_i^{(H)} y_{i,t-H} \\ + \delta_i D_{i,t} + \sum_{H=1}^{H_i} \mu_i^{(H)} D_{i,t} x_{i,t-H} + \sum_{H=1}^{H_i} \rho_i^{(H)} D_{i,t} y_{i,t-H} + \varepsilon_{i,t} \tag{13}$$

where, $D_{i,t} = 0$ for $t = 1,2,\ldots,(T_i^{Yr} - 1)$ and $D_{i,t} = 1$ for $t = T_i^{Yr}, 2, \ldots, T_i$.

If the system Schwarz Information Criteria (SIC) for the VAR allowing a break, equation (13), is less than the VAR without a break, equation (1), we take this as evidence of a structural break for country i. For each country where there is evidence of a structural break we apply time-series GNC tests for both sub-samples using null hypotheses that are specified as follows (for country i):

$$H_0^1: \beta_i^{(1)} = \cdots = \beta_i^{(H_i)} = 0 \tag{14}$$

$$H_0^2: \beta_i^{(1)} + \rho_i^{(1)} = 0 \cap \beta_i^{(2)} + \rho_i^{(2)} = 0 \cap \ldots \cap = \beta_i^{(H_i)} + \rho_i^{(H_i)} = 0 \tag{15}$$

If (14) is rejected there is evidence that $y_{i,t}$ Granger-causes $x_{i,t}$ for country i in the pre-break sub-sample, otherwise GNC is inferred in the first period. Similarly, if (15) is rejected there is evidence that $y_{i,t}$ Granger-causes $x_{i,t}$ for country i in the post-break sub-sample, otherwise GNC is inferred for the second sub-sample. We use both F- and Wald versions of the GNC tests to assess whether (14) and (15) can be rejected.

We then apply the Hanck (2013) panel method to the probability values of the time-series F and Wald GNC test statistics using all countries where there is evident parameter change in the panel. Based on these panel tests we determine whether $y_{i,t}$ Granger-causes $x_{i,t}$ for country i in each sub-sample.

[12] When there are an even number of time-series observations for country i both sub-samples have the same number of observations. When there are an odd number of time-series observations the first sub-sample has one more data point than the second sub-sample. T_i^{Yr} denotes the date of the first observation of the second sub-sample.

3.2 Data Description and Sources

An unbalanced panel dataset of 136 countries (see column 1 of Table 1) covering the period 1970–2006 (annually) is used. The data were extracted from the WDI 2006 edition. The two variables employed are net inflows of FDI as a percentage of GDP (denoted FDIG), and real per-capita GDP growth (denoted GDPG). The unit of measurement for both variables (prior to transformation) is US dollars.

4. Empirical Results

We test the GNC null in bivariate VARs for both FDIG to GDPG and the reverse causality relationship of GDPG to FDIG. The three panel tests discussed above are applied as well as standard time-series tests. Results based on both Wald and F statistics are given. Table 1 and Table 2 report the GNC test results from the time-series and Hanck (2013) methods while the Hurlin and Fisher tests are presented in Table 3. In Table 4 we present the time-series and Hanck (2013) tests that allow for a structural break for those countries where an evident structural break is found.

4.1 Time-series test results

The lag lengths of the VAR are chosen for both variables in each country according to Schwarz's Information Criteria (SIC), with a maximum of 3 lags, (when the SIC favoured zero lags GNC tests were applied in a VAR with 1 lag). The probability values relating to the F and Wald versions of the time-series GNC tests are reported under the columns headed P_j of Table 1 and Table 2, respectively. Test statistics for GDPG causing FDIG are given under the heading GDPG to FDIG whereas test statistics for FDIG causing GDPG are given under the heading FDIG to GDPG.

According to the time-series F-test there is evidence of GC from GDPG to FDIG at the 5% level for 15 of the 136 countries (Algeria, Bangladesh, Gabon, Greece, Iran, Ireland, Israel, Jordan, Kuwait, Niger, Norway, Oman, Spain, Turkey and Vietnam). Similarly, the time-series Wald test suggests that there is evidence of GC from GDPG to FDIG at the 5% level for 17 countries (the same 15 countries as identified by the F-test plus Kenya and Macedonia). The F-test indicates evidence of GC from FDIG to GDPG at the 5% level for 14 countries (Algeria, Burkina Faso, Chile, El Salvador, Estonia, Guinea Bissau, Guyana, Honduras, South Korea, Lithuania, Malaysia, Mauritius, Poland and Tanzania). The Wald test identifies evidence of GC from FDIG to GDPG at the 5% level for 15 countries (the same 14 countries indicated by the F-test plus Barbados). Where there is evidence of GC it is unidirectional except for Algeria where bidirectional causality is suggested. Whilst there is evidence of GC for a small number of countries, the time-series results indicate no causality in either direction for the vast majority of countries – 108 or 79.4% according to the F-test and 105 or 77.2% using the Wald test.

[Insert table 1 and 2]

4.2 Fisher and Hurlin panel test results

The rows labelled Hurlin and Fisher at the bottom of the Table 3 give the Hurlin and Fisher panel test statistics with associated probability values. The probability values for Fisher statistics based on the F (Wald) version of the GNC test are 0.089 (0.005) for GDPG to FDIG and 0.001 (0.000) for FDIG to GDPG. The test results cause us to reject the GNC null hypothesis for all countries at the 5% level of significance for all tests except that based on the F-version for GDPG causing FDIG, where the null can only be rejected at the 10% level. These results unambiguously suggest that FDIG Granger-causes GDPG for at least one country. While the evidence is ambiguous as to whether GDPG Granger-causes FDIG for at least one country this (alternative) hypothesis is not convincingly rejected and we cannot discount the possibility that GC exists in this direction as well for at least one country.

[Insert table 3]

The one tailed probability values based on the Normal distribution for Hurlin's (2004a, 2004b) panel test, presented at the bottom of Table 1, are only available for the Wald version of the test, see equation (4). The probability value for Granger-causality $\tilde{Z}_{N;T}^{HNC}$ from GDPG to FDIG is 0.086 which suggests that GDPG does not Granger-cause FDIG at the 5% level for any of the 136 countries in the panel – if it is rejected at the 10% level. In contrast, the probability for GNC from FDIG to GDPG is 0.000 which rejects the null hypothesis at all conventional levels of significance and unambiguously indicates that FDI Granger-causes GDPG for at least one country in the panel.

Hence, the Fisher and Hurlin panel tests unambiguously indicate that FDI Granger-causes GDPG for at least one country, however, they both show some ambiguity as to whether GDPG Granger-causes FDI for any country.

4.3 Panel Hanck (2013) test results

This section considers the GNC test results from the Hanck (2013) panel method (not allowing for structural breaks) based upon probability values from the time-series F-tests and Wald tests, reported in Table 1 and Table 2, respectively.

Based upon the F-test for Granger-causality from GDPG to FDIG we find that r = 136 (see equation (12)) because the probability values, P_j, are greater than $\frac{\propto}{r}$ (with $\alpha = 0.05$), being 0.00037, for all 136 countries. This suggests that the null hypothesis that GDPG does not Granger-cause FDIG cannot be rejected for all countries. The F-test for Granger-causality from FDIG to GDPG indicates that r = 134 because the probability values are greater than 0.00037 for 134 countries. Thus, the only two countries where there is evidence that FDIG Granger-causes GDPG are Estonia and Guyana. FDIG does not Granger-cause GDP for the remaining 134 countries.

Using the Wald test for Granger-causality from GDPG to FDIG we find that r = 136 which suggests that the null hypothesis that GDPG does not Granger-cause FDIG cannot be rejected for all countries. This is consistent with the Hanck (2013) results from the F-test. The Wald test for Granger-causality from FDIG to GDPG indicates that r = 133. Thus, for only 3 of the 136 countries is there evidence that FDIG Granger-causes GDPG being Estonia, Guyana and Poland. The only difference from the Hanck (2013) F-test results is that Poland is added to the countries where there is evident Granger-causality.

We now consider the Hanck (2013) method allowing for structural breaks in the VAR. For 131 of the 136 countries we could estimate VAR models that allow all coefficients to change in the middle of the sample.[13] For twelve countries (that are identified in Table 4) the VAR with a

[13] The five countries where there were insufficient degrees of freedom to estimate VAR models that allow a break were Bangladesh, Burkina Faso, Kyrgyz Repubic, Mongolia and the Slovak Republic.

break is preferred to that without a break as the former had a lower SIC compared to the latter (the unreported results are available from the authors upon request). We apply the panel Hanck (2013) panel test procedure to the GNC hypothesis for both the pre-break and post-break sub-samples for these twelve countries and report the results in Table 4. For none of the twelve countries is there evidence that GDPG Granger-causes FDIG in the pre-break sub-sample according to the Hanck (2013) method based on either the F or Wald statistics. However, there is evidence that GDPG Granger-causes FDIG in the post-break sub-sample according to the Wald (F) tests for four (one) countries, being: Dominican Republic, Gabon, Madagascar and Poland (Gabon). There are three (two) countries where FDIG Granger-causes GDPG in the pre-break period according to the Wald (F) based test: Switzerland, Tajikistan and Yemen (Tajikistan and Yemen). In the post-break period FDIG Granger-causes GDPG according to the Wald (F) based test for the following four (three) countries; Estonia, Poland, Tajikistan and Yemen (Estonia, Tajikistan and Yemen).

[Insert table 4]

Hence, allowing for a structural break in the VAR increases the number of countries where Granger-causality is evident according to the Hanck (2013) method. Taking the results from both the no break and break-point VAR, we therefore conclude that GDPG Granger-causes FDIG for at least one part of the sample period for four of the 136 countries (Dominican Republic, Gabon, Madagascar and Poland). Similarly, FDIG Granger-causes GDPG for at least one part of the sample period for six of the 136 countries (Estonia, Guyana, Poland, Switzerland, Tajikistan and Yemen). Nevertheless, we find no Granger-causality for the vast majority of countries considered.

5. Conclusion

The main objective of our work is to investigate the issue of causality across a sample of 136 diverse countries for the period 1970 – 2006 by applying time-series and three panel Granger-causality tests based on Hurlin (2004a, 2004b), Fisher (1948), and Hanck (2013). We also apply Hanck (2013) based GNC tests for twelve countries where we find evidence of parameter change. As argued, the data set is larger than previous similar studies and the methods are the most advanced and general available. In particular, they can accommodate heterogeneous intercepts and slopes, thus allowing us to make country-by-country inferences and not make possibly erroneous generalised inferences across the cross-section. We argue that this is an appropriate approach in view of the disparate and conflicting results of existing empirical studies.

The results can be summarised as follows. According to the Hurlin and Fisher panel tests FDIG unambiguously Granger-causes GDPG for at least one country. However, the results from these tests are ambiguous regarding whether GDPG Granger-causes FDIG for at least one country. Using the Hanck (2013) panel test we are able to determine whether and for which countries there is Granger-causality. This test suggests that at most there are six countries (Estonia, Guyana, Poland, Switzerland, Tajikistan and Yemen) where FDIG Granger-causes GDPG and at most four countries (Dominican Republic, Gabon, Madagascar and Poland) where GDPG Granger-causes FDIG. The results from Hanck's (2013) test are broadly consistent with those based on Fisher (1948) and Hurlin (2004a, 2004b), however, the former are illuminating in that they suggest that there is evidence of Granger-causality for very few (nine) of the 136 countries. We regard the panel tests as more reliable than the individual time-series tests, which also suggest evidence of Granger-causality for relatively few countries (if more than is indicated by the panel tests).

We note that the nine countries (Dominican Republic, Estonia, Gabon, Guyana, Madagascar, Poland, Switzerland, Tajikistan and Yemen) where there is evident Granger-causality according to the Hanck (2013) method have different histories of macroeconomic episodes, policy regimes and growth patterns. For instance, according to the World Bank, Estonia and Poland are European economies in transition which have policy decisions that attract even more FDI and their locations and growth prospects thus favour them.

Our finding that in only 6 out of 136 countries is there significant Granger-causality from FDIG to GDPG suggests that there is no impact of FDI on economic growth for virtually all countries. However, it maybe that the share of FDI inflows to GDP have been quantitatively too small to have a high and significant impact on economic growth or that the relationship between the two variables is too complex to be identified in a bivariate Granger-causality framework. Further, the relationship between FDI and economic growth may well depend on the determinants of FDI. If the determinants have a strong link with growth in the host country growth may be found to cause FDI, while output may grow faster when FDI takes place under other circumstances.

Overall, the empirical evidence reported in this paper would lend support to a conclusion that there is little causality between FDI inflows and economic growth in either direction (excepting 9 countries out of 136). Thus, while there is much attention in policy and academia on FDI, our evidence questions whether FDI is substantively related with the growth process.

Regarding possible future research we suggest the consideration of causality in systems involving more than two variables as well as replacing FDI with foreign capital inflows such as foreign direct investment, foreign aid, foreign portfolio investment and foreign loans – see, for example, Ehigiamusoe and Lean (2019). Consideration of other testing frameworks, such as the ARDL method, would also be desirable.

References

Agarwal, P. (2000) Economic impact of foreign direct investment in South Asia, mimeo. Bombay: Indira Gandhi Institute of Development Research, Gen. A.K. Bombay; India

Alfaro, L., Chanda, A., Kalemli-Ozcan, S. and Sayek, S. (2004). FDI and economic growth: The role of local financial markets, Journal of International Economics. 64: 89-112.

Asiedu, E. (2002). On the Determinants of Foreign Direct Investment to Developing Countries: Is Africa Different?, World Development, 30 (1): 107-19.

Asiedu, E. (2003). Policy Reform and Foreign Direct Investment to Africa: Absolute Progress but Relative Decline, Development Policy Review. 22 (1): 41-48.

Barro, R J., (1991), "Economic Growth in a Cross Section of Countries," Quarterly Journal of Economics Vol. 106, No. 2, pp. 407 - 443.

Barro, R J., and Xavier Sala-I-Martin, (1995), Economic Growth, McGraw-Hill, Boston.

Bengoa,M.and Sanchez-Robles,B.(2003), " Foreign direct investment,economic freedom and growth: new evidence from Latin America", European journal of political,Economy,vol.19,pp.529-45.

Blomström, M., Kokko, A., and M. Zejan (1992). Host Country Competition and Technology Transfer by Multinationals, NBER Working Papers 4131, National Bureau of Economic Research, Inc.

Blomström, M., Lipsey, R.E. and Zejan, M. (1996). Is Fixed Investment the Key to Economic Growth?, Quarterly Journal of Economics. 111(1): 269-276.

Borenztein,E., DeGregorio, J. and Lee, J.W. (1998). How does foreign direct investment affect economic growth, Journal of International Economics. 45: 115-135.

Busse, M. and Groizard, J.L. (2008). FDI, Regulations and Growth, The World Economy. 31(7): 861-886.

Calvo, M.B., Robles, B.S. (2003), "Foreign direct investment, economic freedom and growth: new evidence from Latin-America", *European Journal of Political Economy*, Vol. 19 No.3, pp.529-45.

Campos, N. F. and Kinoshita, Y. (2002). Foreign Direct Investment as Technology Transferred: Some Panel Evidence from the Transition Economies, The Manchester School. 70, (3): 398-419.

Carkovic, M., and Levine, R. (2005). Does Foreign Direct Investment Accelerate Economic Growth? in T.H. Moran, E.M. Graham, and M. Blomstrom. Does Foreign Direct Investment Promote Development? Washington, DC: Institute for International Economics and Center for Global Development. Washington, DC, 195–220.

Caves R. (1996): Multinational enterprises and Economic analysis. Cambridge: Cambridge University Press

Chakrabarti, A. (2001). The Determinants of Foreign Direct Investment: Sensitivity Analyses of Cross-Country Regressions, Kyklos. 54 (1): 89-113.

Chakraborty, C., and Basu, P. (2002). Foreign Direct Investment and growth in India: a cointegration approach, Applied Economics. 34: 1061-1073.

Chan V. (2000): Foreign Direct Investment and Economic Growth in Taiwan's Manufacturing Industries; in Kruger K. and Takatoshi I. (eds.): The Role of FDI in East Asia Economic Development. Chicago: Chicago University Press.

Chanegriha, M., Stewart, C, Tsoukis C. (2017): Identifying the robust economic, geographical and political determinants of FDI: an Extreme Bounds Analysis, Empirical Economics. 52, 2: 759-776.

Choe, J.I. (2003). Do foreign direct investment and gross domestic investment promote economic growth?, Review of Development Economics. 7: 44-57.

Chowdhury, A., and Mavrotas, G. (2005). FDI and Growth: A Causal Relationship, United Nations University, WIDER, Research Paper No: 2005/25.

De Mello, L.F. Jr. (1997), Foreign Direct Investment in Developing Countries and Growth: A Selective Survey, Journal of Development Studies. 34 (1): 1-34.

De Mello, L. F., Jr. (1999). Foreign Direct Investment-led Growth: Evidence from time Series and Panel Data, Oxford Economic Papers. 51 (1): 133-151.

Dutt, A., Duttaray, M. and Mukhopadhyay, K. (2008). Foreign direct investment and economic growth in less developed countries: An empirical study of causality and mechanisms, Applied Economics. 40: 1927-1939.

Ehigiamusoe, K. and Lean, H. H. (2019). Foreign Capital Inflows and Economic Growth in Nigeria: Any Nexus, forthcoming in Journal of African Business.

Ericsson, J., and Irandoust, M., (2001). On the Causality between Foreign Direct Investment and Output: A Comparative Study, The International Trade Journal. 15(1): 1-26.

Felipe, J. (1997). Total Factor Productivity Growth in East Asia: A Critical survey. ADB EDRC Report Series N. 65.

Fisher, R.A. (1932). Statistical Methods for Research Workers, Edinburgh: Oliver & Boyd.

Fisher, R.A. (1948). Combining Independent Tests of Significance, American Statistician. 2 (5): 30-31.

Gao, T. (2005). Foreign direct investment and growth under economic integration, Journal of International Economics. 67: 157-174.

Ghosh, M. and Wang, W. (2009). Does FDI accelerate economic growth? Global Economy Journal, 9 (4): 1-21.

Gorg, H., and D. Greenaway (2004).Much Ado about Nothing? Do Domestic Firms Really Benefit from Foreign Direct Investment?, World Bank Research Observer, 19: 171-97.

Greene, W.H. (2000). Econometric analysis. Englewood Cliffs, NJ: Prentice-Hall International.

Gursoy, F., Sekreter, A. and Kalyoncu, H. (2013). FDI and economic growth relationship based on cross-country comparison, International Journal of Economics and Financial Issues. 3(2): 519-524.

Hanck, C. (2008). An Intersection Test for Panel Unit Roots, Mimeo, Department of Economics, University of Dortmund.

Hanck, C. (2013). An intersection test for panel unit roots, Econometric Reviews. 32(2): 183–

203.
Hommel, G. (1988). A Stagewise Rejective Multiple Test Procedure Based on a Modified Bonferroni Test, Biometrika. 75 (2): 383-386.
Hood III, M.V., Kidd, Q. and Morris, I. L. (2008). Two Sides of the Same Coin? Employing Grange Causality Tests in a Time Series Cross-Section Framework, Political Analysis. 16: 324–344.
Hsiao, F. S. T. "FDI, Exports and GDP in East and Southeast Asia — Panel Data versus Time-series Causality Analyses," Journal of Asian Economics, 17(4), 2006, pp. 1082-1106
Hurlin, C. (2004a). Testing Granger Causality in Heterogeneous Panel Data Models with Fixed Coefficients, Mimeo, University of Orléans.
Hurlin, C. (2004b). A Note on Causality Tests in Panel Data Models with Random Coefficients, Miméo, University of Orléans.
Hurlin, C. and Venet, B. (2001). Granger Causality Tests in Panel Data Models with Fixed Coefficients, Mimeo, University Paris IX.
Hurlin, C., (2008). Testing for Granger Non-causality in Heterogeneous Panels, Working Paper LEO, No. 2007-10, University of Orléans.
Lensink, R., and Morrissey, O. (2006). Foreign Direct Investment: Flows, Volatility, and the Impact on Growth, Review of International Economics. 14(3): 478–493.
Lucas, R., (1988). On the Mechanics of Economic Development. Journal of Monetary Economics, 22 (1), 3-42
Mankiw,N.G, Romer, D and Weil, D.(1992) Acontribution to the empirics of economic growth. Quaterly Journal of Economics,107,407-437.
Mencinger, J., (2003). Does Foreign Direct Investment Always Enhance Economic Growth?, Kyklos, 56(4), 491-509.
OECD (2002). Foreign direct investment for development: Maximising benefits, minimizing costs , Paris, OECD publishing.
Ozturk, I. (2007). Foreign Direct Investment-Growth Nexus: A Review of the recent Literature, International Journal of Applied Econometrics and Quantitative Studies. 4(2): 79-98.
Ozturk, I. and Kalyoncu, H. (2007). Foreign Direct Investment and Growth: An Empirical Investigation Based on Cross-Country Comparison, Economia Internazionale. 60 (1): 75-82.
Simes, R. J. (1986). An improved Bonferroni Procedure for Multiple Tests of significance, Biometrika. 73, 3: 751-754.
Sooreea-Bheemul, B. and Sooreea, R. (2013). Missing causality links between foreign direct investment, exports, domestic investment and economic growth, International Journal of Business and Emerging Markets. 5 (4): 322-340.
Xu, B., (2000). Multinational enterprises, technology diffusion, and host country productivity growth, Journal of Development Economics, 62: 477-93.
Zhang, K.H. (2001). How does Foreign Direct Investment affect Economic Growth in China?, Economics of Transition. 9 (3): 679–693.

Table 1: Time-series and Hanck (2013) GNC tests (F-statistic)

GDPG to FDIG

Country	P_j	Country	P_j	Country	P_j	Country	P_j
Oman	*0.00407**	USA	0.19981	Swaziland	0.50615	Iceland	0.73854
Algeria	*0.00802**	China	0.21266	Korea	0.51001	Croatia	0.74221
Niger	*0.01522**	Madagascar	0.21290	New Zealand	0.53996	Armenia	0.76413
Ireland	*0.01660**	Indonesia	0.21506	Cyprus	0.56445	Nicaragua	0.76947
Kuwait	*0.01843**	El Salvador	0.22315	Zambia	0.57261	Singapore	0.77647
Iran	*0.02353**	Brazil	0.22858	Fiji	0.57957	Zimbabwe	0.77677
Turkey	*0.02712**	Mozambique	0.24091	Yemen	0.58545	Ghana	0.78207
Gabon	*0.02831**	Guyana	0.24464	Uruguay	0.60697	Senegal	0.83143
Spain	*0.03250**	Argentina	0.26016	Uganda	0.60700	Kyrgyz Rep	0.83208
Jordan	*0.03501**	Moldova	0.26038	Equatorial	0.61902	Denmark	0.84976
Vietnam	*0.03824**	Uzbekistan	0.27409	France	0.62188	Belarus	0.85270
Bangladesh	*0.04144**	Portugal	0.29978	Peru	0.63007	Belize	0.86255
Norway	*0.04484**	Mongolia	0.33817	Bulgaria	0.63497	Austria	0.86480
Greece	*0.04857**	Ethiopia	0.34389	Switzerland	0.64267	Liberia	0.87705
Israel	*0.04904**	Australia	0.36163	Syrian	0.65463	Bolivia	0.88163
Kenya	0.05850	Romania	0.36166	Sudan	0.65807	Tonga	0.88176
Macedonia	0.07140	Japan	0.37770	Rwanda	0.66332	Kazakhstan	0.88276
Tunisia	0.07464	Thailand	0.39462	Costa Rica	0.66981	Ivory Cost	0.88793
Canada	0.08068	Morocco	0.39467	Congo Rep	0.67120	Slovenia	0.89119
Belgium	0.08558	Sierra leon	0.39829	Guinea	0.67264	Guatemala	0.90013
Chad	0.09241	Nigeria	0.40402	Cambodia	0.67401	Barbados	0.90073
Mexico	0.10784	Congo Dem	0.40428	Guinea Bissau	0.67554	Lesotho	0.90329
Netherland	0.11117	Mauritius	0.41412	Albania	0.67974	Ecuador	0.90380
UK	0.12233	Grenada	0.41555	Paraguay	0.69126	Slovak Rep	0.90741
South Africa	0.14309	India	0.41885	Tanzania	0.69150	Italy	0.91439
Poland	0.14844	Honduras	0.42053	Finland	0.69169	Burkina Faso	0.92351
Colombia	0.15317	Hungary	0.42097	Sweden	0.69342	Czech Rep	0.92572
Haiti	0.15513	Benin	0.45678	Mali	0.69820	Sri Lanka	0.94118
Dominican Rep	0.16857	Estonia	0.48167	Togo	0.70062	Panama	0.94401
Central Africa	0.17150	Malaysia	0.48170	Germany	0.70707	Philippines	0.95820
Lithuania	0.17522	Angola	0.48204	Egypt	0.71466	Pakistan	0.96614
Vanuatu	0.17710	Malawi	0.48223	Jamaica	0.72368	Somalia	0.96712
Tajikistan	0.17788	Mauritania	0.48418	Venezuela	0.72702	Nepal	0.96837
Burundi	0.18354	Botswana	0.49965	Djibouti	0.73561	Chile	0.99030

Table 1: Time-series and Hanck (2013) GNC tests (F-statistic) continued

FDIG to GDPG

Country	P_j	Country	P_j	Country	P_j	Country	P_j
Estonia	***0.00000*****	Jamaica	0.24437	Angola	0.43401	Dominican	0.70277
Guyana	***0.00028*****	Cyprus	0.24742	Sri Lanka	0.43696	Djibouti	0.71720
Honduras	*0.00161**	Germany	0.25516	Bolivia	0.44615	Romania	0.72008
Poland	*0.00254**	Portugal	0.25871	Ghana	0.45092	Sierra Leon	0.73788
Algeria	*0.00713**	Egypt	0.26035	Mozambique	0.46557	Argentina	0.75688
Lithuania	*0.01573**	Togo	0.26067	Yemen	0.47407	New Zealand	0.76084
El Salvador	*0.02281**	Uruguay	0.26287	Malawi	0.47421	Vanuatu	0.76264
Chile	*0.02285**	Senegal	0.27362	Denmark	0.48289	India	0.80901
Tanzania	*0.02694**	Singapore	0.28075	Switzerland	0.48440	Venezuela	0.81089
Mauritius	*0.02706**	USA	0.29176	Oman	0.51308	Iran	0.81999
Korea	*0.03358**	Syrian	0.29852	Slovak Rep	0.51523	Croatia	0.82296
Guinea Bissau	*0.03519**	Netherland	0.30226	Fiji	0.51651	Spain	0.83109
Burkina Faso	*0.04276**	France	0.30602	Vietnam	0.51695	Uganda	0.83150
Malaysia	*0.04850**	Mexico	0.30893	Ethiopia	0.52185	Jordan	0.83199
Barbados	0.05280	Czech Rep	0.31139	Botswana	0.52256	South Africa	0.85139
Finland	0.06744	Sweden	0.31832	Zambia	0.52422	Rwanda	0.85199
Haiti	0.07249	Canada	0.33207	Paraguay	0.52994	Philippines	0.86277
Indonesia	0.08319	Turkey	0.33396	Kenya	0.53162	Tajikistan	0.86570
Japan	0.09495	Ireland	0.33662	Albania	0.54220	Liberia	0.87773
Lesotho	0.12259	Burundi	0.33992	China	0.54548	Kuwait	0.87907
Nigeria	0.13545	Zimbabwe	0.34977	Australia	0.54954	Belgium	0.88109
Tunisia	0.13915	Austria	0.35797	Thailand	0.58584	Kyrgyz Rep	0.88844
Slovenia	0.14347	Peru	0.35853	Bangladesh	0.60362	Mongolia	0.89703
Congo Dem	0.14643	Morocco	0.37840	Sudan	0.61283	Hungary	0.90669
Moldova	0.16158	Pakistan	0.37870	Madagascar	0.61849	Mali	0.92423
Guinea	0.16759	Panama	0.37991	Somalia	0.63957	Central Africa	0.93305
Gabon	0.18329	Costa Rica	0.38082	Belarus	0.64134	Tonga	0.95354
Iceland	0.19860	Ecuador	0.38961	Bulgaria	0.64851	Nepal	0.96147
Mauritania	0.20005	Guatemala	0.39679	Benin	0.65577	Kazakhstan	0.96584
Uzbekistan	0.20344	Chad	0.40585	Congo Rep	0.65777	Equatorial	0.97720
Nicaragua	0.20662	Ivory cost	0.40617	Swaziland	0.65821	Cambodia	0.97998
Grenada	0.23086	Brazil	0.40674	Colombia	0.65966	Israel	0.98438
Italy	0.23539	UK	0.41084	Niger	0.67626	Armenia	0.99268
Macedonia	0.23926	Norway	0.42515	Greece	0.69267	Belize	0.99968

Table 1 notes:

The column headed Country identifies the country to which the row refers to. The column headed P_j gives the probability value for each individual country's time-series GNC test arranged in ascending order of magnitude. According to the Hanck (2013) test, when P_j is below $\frac{\propto}{r} = 0.00037$ the GNC null is rejected for that country, where $\frac{\propto}{r}$ gives the nominal level of significance ($\propto = 0.050$) divided by r, where r = 136 (GDPG to FDIG) and r = 134 (FDIG to GDPG). A single asterisk and italic (*) statistic indicates that a time-series test statistic is significant at the 5% level. A double asterisk and bold (**) statistic indicates that the Hanck (2013) test statistic is significant at the 5% level.

Table 2: Time-series and Hanck (2013) GNC tests (Wald statistic)

GDPG to FDIG

Country	P_j	Country	P_j	Country	P_j	Country	P_j
Bangladesh	*0.00045**	Burundi	0.17403	Swaziland	0.50100	Iceland	0.73439
Oman	*0.00192**	USA	0.19048	Korea Rep	0.50523	Croatia	0.74004
Algeria	*0.00326**	China	0.20349	New Zealand	0.53494	Armenia	0.75839
Niger	*0.01032**	Indonesia	0.20593	Cyprus	0.54970	Nicaragua	0.76729
Ireland	*0.01100**	Uzbekistan	0.20887	Zambia	0.56862	Singapore	0.77450
Kuwait	*0.01191**	El Salvador	0.21352	Fiji	0.57331	Zimbabwe	0.77492
Iran	*0.01739**	Brazil	0.21834	Yemen	0.57419	Ghana	0.78016
Gabon	*0.02049**	Mozambique	0.22750	Uganda	0.59876	Senegal	0.83007
Turkey	*0.02056**	Guyana	0.22905	Uruguay	0.60340	Kyrgyz Rep	0.83020
Spain	*0.02418**	Mongolia	0.23170	Equatorial	0.61558	Denmark	0.84587
Vietnam	*0.02466**	Moldova	0.23284	France	0.61599	Belarus	0.85124
Jordan	*0.02579**	Argentina	0.24926	Peru	0.62676	Belize	0.86124
Norway	*0.02891**	Portugal	0.29042	Bulgaria	0.63171	Austria	0.86372
Kenya	*0.03274**	Ethiopia	0.32046	Switzerland	0.63730	Liberia	0.87580
Israel	*0.03962**	Romania	0.34295	Syria	0.65138	Tonga	0.88008
Greece	*0.04031**	Australia	0.35470	Sudan	0.65486	Bolivia	0.88070
Macedonia	*0.04375**	Japan	0.37102	Rwanda	0.66037	Kazakhstan	0.88162
Belgium	0.05648	Thailand	0.38820	Costa Rica	0.66692	Ivory Cost	0.88704
Tunisia	0.06536	Morocco	0.38825	Congo Dem	0.66833	Slovenia	0.88842
Canada	0.07064	Sierra Leon	0.39193	Guinea Bissau	0.66948	Guatemala	0.89904
Chad	0.08279	Mauritius	0.39605	Guinea	0.66978	Barbados	0.89924
Mexico	0.09806	Nigeria	0.39684	Albania	0.67074	Lesotho	0.90139
Dominican	0.09819	Congo Rep	0.39801	Cambodia	0.67098	Ecuador	0.90305
Netherlands	0.10137	Guatemala	0.40738	Tanzania	0.68490	Slovak Rep	0.91125
UK	0.11250	India	0.41138	Paraguay	0.68859	Italy	0.91371
Poland	0.12016	Honduras	0.41450	Finland	0.68865	Czech Rep	0.92498
South Africa	0.13329	Hungary	0.41495	Sweden	0.69078	Burkina Faso	0.92731
Colombia	0.13960	Benin	0.44793	Mali	0.69561	Sri Lanka	0.94074
Haiti	0.14539	Estonia	0.47058	Togo	0.69805	Panama	0.94358
Tajikistan	0.14726	Angola	0.47228	Germany	0.70456	Philippines	0.95790
Madagascar	0.15128	Malaysia	0.47580	Egypt	0.71223	Pakistan	0.96590
Central Africa	0.16188	Malawi	0.47708	Jamaica	0.72135	Somalia	0.96663
Vanuatu	0.16317	Mauritania	0.47889	Venezuela	0.72471	Nepal	0.96803
Lithuania	0.16564	Botswana	0.49474	Djibouti	0.73297	Chile	0.99016

Table 2: Time-series and Hanck (2013) GNC tests (Wald statistic) continued

FDIG to GDPG

Country	P_j	Country	P_j	Country	P_j	Country	P_j
Estonia	***0.00000*****	Italy	0.22653	Angola	0.42295	Greece	0.69002
Guyana	***0.00001*****	Jamaica	0.23564	Sri Lanka	0.43117	Romania	0.71369
Poland	***0.00010*****	Germany	0.24658	Bolivia	0.44049	Djibouti	0.71436
Honduras	*0.00057**	Portugal	0.24860	Ghana	0.44495	Sierra Leon	0.73568
Burkina Faso	*0.00150**	Egypt	0.25185	Mozambique	0.45733	Argentina	0.75428
Algeria	*0.00276**	Togo	0.25218	Yemen	0.45854	New Zealand	0.75857
Lithuania	*0.01074**	Uruguay	0.25441	Denmark	0.46618	Vanuatu	0.75977
Mauritius	*0.01179**	Senegal	0.26532	Slovak Rep	0.46626	India	0.80709
Tanzania	*0.01267**	Singapore	0.27201	Malawi	0.46895	Venezuela	0.80935
El Salvador	*0.01640**	USA	0.28374	Switzerland	0.47573	Iran	0.81854
Chile	*0.01680**	France	0.28809	Ethiopia	0.50684	Croatia	0.82153
Guinea Bissau	*0.02061**	Syria	0.29008	Vietnam	0.50807	Uganda	0.82837
Korea Rep	*0.02638**	Netherland	0.29440	Oman	0.50819	Spain	0.82947
Barbados	*0.03738**	Mexico	0.30118	Fiji	0.50897	Jordan	0.83024
Malaysia	*0.03911**	Czech Rep	0.30150	Kenya	0.51703	South Africa	0.85020
Finland	0.05711	Sweden	0.31071	Botswana	0.51795	Rwanda	0.85079
Haiti	0.06326	Canada	0.32418	Zambia	0.51963	Philippines	0.86168
Indonesia	0.07372	Turkey	0.32661	Paraguay	0.52542	Tajikistan	0.86223
Japan	0.08529	Ireland	0.32854	Albania	0.52805	Liberia	0.87648
Lesotho	0.09866	Burundi	0.33265	China	0.54117	Kuwait	0.87787
Slovak Rep	0.11238	Zimbabwe	0.34266	Australia	0.54527	Belgium	0.87803
Nigeria	0.12423	Austria	0.35098	Bangladesh	0.56733	Kazakhstan	0.88720
Tunisia	0.12933	Peru	0.35156	Thailand	0.58201	Mongolia	0.90137
Moldova	0.13065	Morocco	0.37173	Madagascar	0.60018	Hungary	0.90593
Uzbekistan	0.13313	Pakistan	0.37204	Sudan	0.60908	Mali	0.92362
Congo Dem	0.13664	Panama	0.37327	Somalia	0.63350	Central Africa	0.93253
Guinea	0.15794	Costa Rica	0.37419	Belarus	0.63741	Tonga	0.95292
Gabon	0.17202	Ecuador	0.38312	Bulgaria	0.64540	Nepal	0.96103
Iceland	0.18098	Guatemala	0.38788	Benin	0.65089	Kazakhstan	0.96547
Mauritania	0.19042	Brazil	0.39960	Congo Rep	0.65476	Equatorial	0.97696
Nicaragua	0.19605	Chad	0.39960	Swaziland	0.65500	Cambodia	0.97977
Macedonia	0.21077	Ivory Cost	0.39993	Colombia	0.65549	Israel	0.98414
Cyprus	0.21612	UK	0.40467	Niger	0.67344	Armenia	0.99252
Grenada	0.21894	Norway	0.41208	Dominican	0.68997	Belize	0.99968

Table 2 notes:
See notes to Table 1 except r = 136 for the GDPG to FDIG test, r = 133 for the FDIG to GDPG test and $\frac{\propto}{r} = 0.00038$ for the FDIG to GDPG test.

Table 3: Hurlin and Fisher GNC tests

	GDPG to FDIG		FDIG to GDPG		GDPG to FDIG		FDIG to GDPG	
Test	F-test	PF	F-test	PF	W	PW	W	PW
Hurlin					1.369	0.086	4.502	0.000
Fisher	303.867	0.089	348.541	0.001	335.918	0.005	432.065	0.000

Table 3 notes:
Hurlin denotes Hurlin's panel GNC Wald test allowing for heterogeneous T and H (which is appropriate for finite T and large N) and the corresponding (one-tail) asymptotic (normal) p-values are beneath PW. Asymptotic (one-tail normal distribution) 1%, 5% and 10% critical values for Hurlin's test are, respectively: 2.326, 1.645 and 1.282. Semi-asymptotic (one-tail) 1%, 5% and 10% critical values for Hurlin's test are, respectively: 1.664, 1.550 and 1.489. Fisher denotes Fisher's panel GNC tests (both and F and Wald versions below their associated headings) with corresponding chi-squared (with 2N degrees of freedom) probability values beneath PF and PW. The 1%, 5% and 10% critical values for the Fisher-type panel GNC test are: 329.181, 311.467 and 302.286.

Table 4: Time-series and Hanck (2013) GNC tests for countries with structural breaks (F and Wald statistics)

F-statistic

GDPG to FDIG				FDIG to GDPG			
Pre-break sub-sample		Post-break sub-sample		Pre-break sub-sample		Post-break sub-sample	
Country	P_j	Country	P_j	Country	P_j	Country	P_j
Dominican Rep	0.160	Gabon	***0.000******	Tajikistan	***0.000******	Tajikistan	***0.000******
Estonia	0.197	Poland	*0.014**	Yemen	***0.001******	Estonia	***0.000******
Poland	0.247	Dominican Rep	*0.050**	Switzerland	*0.013**	Yemen	***0.004******
Uzbekistan	0.272	Madagascar	0.053	Estonia	*0.020**	Poland	*0.032**
Colombia	0.333	Switzerland	0.215	Moldova	0.204	Switzerland	*0.039**
Gabon	0.378	Estonia	0.418	Madagascar	0.293	Moldova	0.407
Switzerland	0.378	Uzbekistan	0.438	Uzbekistan	0.470	Madagascar	0.412
Yemen	0.540	Colombia	0.534	Armenia	0.581	Dominican Rep	0.488
Moldova	0.643	Tajikistan	0.606	Gabon	0.602	Uzbekistan	0.572
Madagascar	0.674	Moldova	0.629	Poland	0.634	Colombia	0.795
Armenia	0.888	Yemen	0.738	Colombia	0.649	Armenia	0.830
Tajikistan	0.947	Armenia	0.990	Dominican Rep	0.715	Gabon	0.870
$r = 12, \frac{\propto}{r} = 0.0042$		$r = 11, \frac{\propto}{r} = 0.0045$		$r = 10, \frac{\propto}{r} = 0.0050$		$r = 9, \frac{\propto}{r} = 0.0056$	

Wald statistic

GDPG to FDIG				FDIG to GDPG			
Pre-break sub-sample		Post-break sub-sample		Pre-break sub-sample		Post-break sub-sample	
Country	P_j	Country	P_j	Country	P_j	Country	P_j
Dominican Rep	*0.005**	Dominican Rep	***0.000******	Tajikistan	***0.000******	Estonia	***0.000******
Uzbekistan	0.068	Gabon	***0.000******	Yemen	***0.000******	Tajikistan	***0.000******
Estonia	0.174	Madagascar	***0.000******	Switzerland	***0.005******	Yemen	***0.000******
Poland	0.212	Poland	***0.000******	Estonia	*0.007**	Poland	***0.004******
Colombia	0.321	Switzerland	0.184	Madagascar	0.149	Switzerland	*0.018**
Switzerland	0.365	Uzbekistan	0.199	Moldova	0.161	Madagascar	0.239
Gabon	0.369	Estonia	0.393	Uzbekistan	0.324	Dominican Rep	0.285
Yemen	0.522	Colombia	0.523	Armenia	0.565	Moldova	0.359
Moldova	0.629	Tajikistan	0.583	Gabon	0.597	Uzbekistan	0.436
Madagascar	0.637	Moldova	0.609	Poland	0.620	Colombia	0.793
Armenia	0.885	Yemen	0.729	Colombia	0.644	Armenia	0.826
Tajikistan	0.945	Armenia	0.990	Dominican Rep	0.672	Gabon	0.869
$r = 12, \frac{\propto}{r} = 0.0042$		$r = 8, \frac{\propto}{r} = 0.0063$		$r = 9, \frac{\propto}{r} = 0.0056$		$r = 8, \frac{\propto}{r} = 0.0063$	

Table 4 notes: see notes to Table 1 except the columns headed "Pre-break sub-sample" refer to GNC tests conducted over the sub-sample prior to the break date and the columns headed "Post-break sub-sample" refer to GNC tests conducted over the sub-sample starting from the break date (the second sub-period). The level of significance is set as $\propto = 0.050$.