Original Research

Does Increasing Product Complexity and Diversity Cause Economic Growth in the Long-Run? A GMM Panel VAR Evidence

Ejike Udeogu[1] [iD], Shampa Roy-Mukherjee[1],
and Uzochukwu Amakom[2] [iD]

Abstract

Using aggregate data from 31 Organization for Economic Co-operation and Development (OECD) countries covering periods from 1982 to 2017, this study examines the notion that the level of product complexity is a good determinant of economic growth in the long run. We use the impulse-response function (IRF) computed from the consistent generalized method of moment panel vector autoregressive (GMM pVAR) model to estimate the response of the real output growth to a change in the economic complexity index. The IRF shows that the economic complexity index has a significant impact on economic growth; a 1 standard deviation shock to the economic complexity index at time 0 contributes around 2.34 percentage points to the average rate of growth of output within the first period. The point estimates are positive and significant up to the third period. The cumulative IRF shows that the aggregate impact on economic growth is about 4.4% in the long run. Compared to some widely used innovation proxies such as the gross expenditure on research and development and secondary school enrollment, the economic complexity index performs relatively better in our model in determining economic growth in the long run.

Keywords

capital accumulation, economic complexity, endogenous growth, Granger causality, panel vector autoregression

Introduction

The view that economic complexity is the main determinant of economic growth in the long run is undoubtedly the highlight of the debate started by the pioneers of economic development theories in the 1940s. Many of the seminal studies implied that the level of sophistication and diversity of a country's produce/exports—*á la* economic complexity—is the primary reason why there are variations in living standards across countries (cf. Hirschman, 1958; Lewis, 1954; Rosenstein-Rodan, 1943; Singer, 1950). The consensus is that when economies move from high dependency on agriculture and extractive industrial products to technologically advanced manufacturing and services—a process often referred to as "structural transformation"—they tend to witness an acceleration of economic development in the long-term.

Interestingly, this view has also been echoed by both neoclassical and modern growth theorists who see economic complexity—in the form of "technological advancement" or "innovation"—as the main determinant of economic growth (see Aghion and Howitt, 1992, 1998a, 1998b; Lucas, 1988; Romer, 1986, 1990; Solow, 1956, 1957, for example). In

particular, the "innovation-based" growth models of Romer (1990) and Aghion and Howitt (1992, 1998a, 1998b)—the main architects of the modern (endogenous) growth theory—suggest that long-run economic growth derives from innovations that come in the form of product complexity and variety, process, and organizational innovation.

The crucial difference between the neoclassical view and the modern view centers around the critical drivers of technological change/innovation or economic complexity. While the neoclassicals see exogenous shocks as driving technological change (cf. Solow, 1956, 1957), the modern growth theorists maintain that an increase in physical and human capital investments, investments in research and development, the level of factor endowments, geographical and demographic factors, institutional quality, the level of social

[1]University of East London, UK
[2]University of Nigeria, Nsukka, Enugu, Nigeria

Corresponding Author:
Uzochukwu Amakom, Institute for Development Studies, University of Nigeria, Enugu Campus, Enugu 40006, Nigeria.
Email: uzochukwu.amakom@unn.edu.ng

capital and networks, including historical trajectories, determine the rate of technological change/innovation or the level of product complexity that affect economic growth in the long run (cf. Acemoglu et al., 2014, 2019; Acemoglu & Robinson, 2012; Aghion & Howitt 1992, 1998a, 1998b; Grossman and Helpman 1991a, 1991b, 1994; Levine et al., 2000; Lucas, 1988; Rebelo, 1991; Romer, 1986, 1987, 1990).

Despite the huge theoretical support for the modern growth hypothesis over the years, consistent empirical evidence, particularly in a panel context, seem to be in short supply. Indeed, a majority of the studies on this subject tend to be cross-sectional, with very few time-series studies. However, given the inconsistency of cross-sectional models to control for individual effects, the outcomes of most of these studies have been seen to produce inconsistent results and conclusions. For instance, while some studies (see Coe & Helpman, 1995; Griliches, 1998; Mohnen, 1990; Mohnen & Lépine, 1991; Zachariadis, 2003 for instance) appear to show a statistically large significant relationship between technical progress (proxied either by expenditure on R&D, patent count, number of researchers or scientists per capita or human capital indicators) and output growth (or productivity, in some cases), many have also tended to produce contradictory results (cf. Bilbao-Osorio & Rodríguez-Pose, 2004; Hall et al., 2010; Jones, 1995).

Remarkably, the inconsistencies in the results of studies examining the innovation–growth nexus have been remarked to stem from the widespread use of "unsatisfactory" innovation or technological progress proxies (cf. Griliches, 1994, 1998; Hall et al., 2010; Hidalgo & Hausmann, 2009). To be sure, many of the few leading empirical studies on the subject have tended to adopt traditional measures, such as the human capital index, gross expenditure on research and development, years of schooling, or the number of scientists per population (see, for example, Acemoglu et al., 2014, 2019; Aghion & Howitt, 1992; Barro, 1991, 1997, 2001; Griliches & Mairesse, 1990; Guellec & Potterie, 2003; Levine et al., 2000; Mankiw et al., 1992; Romer, 1990) to proxy technological change or economic complexity. These measures have however been shown to be inadequate because technological change or economic complexity involves a variety of factors, from an economy's level of factor endowments, geography, institutional quality, level of social capital and networks, historical trajectories, changes in technology and return on capital, that are not properly captured by these simplistic traditional indicators (Hall et al., 2010; Hidalgo & Hausmann, 2009). So, the reduction of technological progress or innovation to a simple static aggregate that does not account for the complementary effects of several underlying factors is seen to result in the inconsistencies of these past results. The inadequacy of the simple ordinary least squares (OLS) estimators in the context of dynamic models with endogenous covariates and the use of cross-sectional data that do not control for both country and period effects could also be said to add to the inconsistencies of past results.

This article contributes to the growing literature on technological progress/innovation and growth by addressing some of the aforementioned gaps in the literature. To begin with, it uses a relatively new and robust index to proxy technological progress or innovation. The paper uses the recently developed economic complexity index (ECI) by Hidalgo and Hausmann (2009) in a dynamic panel setting to estimate economic growth in the long run. Second, it applies an efficient and consistent econometric technique that is robust to endogeneity bias that often arise from possible omission of relevant variables in the model, variables measurement errors, or simultaneous causality. To control for the effect of political and economic institutions that often change slowly over time, exogenous governance indicators were further employed as both control and instrumental variables in the model. Overall, the study offers a substantial quantitative improvement over past studies that have examined the technological progress/innovation–growth nexus.

The remainder of the paper is structured as follows. Section "Conceptual Framework" contains the discussion on the ECI and the theoretical framework. Section "Data sources, variables description, and empirical method" presents the data and methods used for the empirical analysis while the discussion of the empirical results is presented in section "Discussion of Results." Section "Concluding Remarks" provides the concluding remarks and recommendations.

Conceptual Framework

The ECI developed by two prominent Harvard scholars, Ricardo Hausman and Cesar Hidalgo, has remarkably attracted huge attention in mainstream economics in recent years with more than 1700 citations according to Google Scholar. In the main, it is seen as a superior reflector of the level of technological advancement or innovation in an economy and a better determinant of economic growth *cum* development in the long run (cf. Abdon & Felipe, 2011; Caldarelli et al., 2012; Cristelli et al., 2015; Hartmann et al., 2017; Hausmann et al., 2014; Kemeny & Storper, 2014).

To begin with, the index is derived from a process called the method of reflections, which is basically interpreting trade data as a bipartite network in which the diversity and complexity of a country's products (exports) is seen as reflecting the diversity of that country's nontradable capabilities (the knowledge, know-how, social capital, institutions, etc.) and their interactions (cf. Hidalgo & Hausmann, 2009). Essentially, the level of complexity or uniqueness of the product a country exports is deemed as reflecting the level of available "capabilities" involved in the production process. Equally, the diversity of "capabilities" present in a country and their interactions is viewed as determining the complexity or the level of technology embodied in the goods and services produced in the country. In all, the ECI is viewed as measuring the relative knowledge intensity of an economy

by considering the knowledge intensity of the products the country exports.

Capabilities, on the contrary, following the explanation by Hausmann et al. (2014), involve the set of human and physical capital—tacit and explicit knowledge, available tools and machines, and relevant infrastructure, including the domestic and international institutions that exist in a place and time, which encourage capital accumulation or affects individual firm's capacity for continual capital accumulation. The institutions that support the process of capital accumulation, as Kotz et al. (1994) explained, includes political, socio-cultural, and economic institutions such as the state of labor—that is, management relations, the organization of work processes and the character of industrial organization—, the role of money and banking and their relation to industry, the role of the state in the economy, the line-up of political parties, the state of race and gender relations, and the character of the dominant culture and ideology.

Implicitly, it is assumed that low-productivity and low-wage activities (á *la* production of nontechnologically advanced commodities or agricultural produce) reflect the existence of small degree of tacit knowledge and less distributed knowledge, including the unavailability of effective and relevant institutions in that economy. On the contrary, high-productivity and high-wage activities is seen to reflect the presence of a large degree of tacit knowledge and more distributed knowledge, including the availability of effective and relevant institutions in the economy. The wider implication being that development or growth will be slow for countries with productive structures geared toward noncomplex (nontechnologically advanced) and competitive products and fast for countries with productive structures geared toward complex (technologically advanced) and quasi-monopolistic products (cf. Hartmann et al., 2017; Hausmann et al., 2014; Hidalgo & Hausmann, 2009).

Remarkably, several seminal studies on economic development have tacitly (or perhaps even explicitly) conveyed similar view that the level of economic complexity explains the development–underdevelopment dichotomy. For instance, several Marxian theories have argued that the prevalence of complex (or quasi-monopolistic) production that attracts higher rates of return (profit), which leads to increase in output/income for the exporting country, in "core" economies is the reason for the continued prosperity of these developed countries. While the prevalence of noncomplex (competitive) production, which often attract lower profits because of the presence of high-price competition in the world market that puts pressure on the price–cost margin, leading to a decrease or stagnation of output *cum* income for the exporting country, in "peripheries" explains the development of underdevelopment in these poor countries (cf. Dos Santos, 1970; Frank, 1967; Prebisch, 1963; Singer, 1950; Wallerstein, 2004; Williamson, 2013).

This theoretical characterization of the economic complexity–growth nexus could be illustrated schematically as follows (Figure 1).

The positive relation between economic complexity and growth seem also to be reflected in available macroeconomic data. Indeed, available historical data shows that economies exporting complex and diverse products tend, in many cases, to be rich (developed) countries. Whereas countries exporting noncomplex and ubiquitous products unsurprisingly tend to be low-income (developing) countries (Tables 1 and 2).

To conclude, this study tests the hypothesis that the ECI is a good reflection of the level of technological advancement (capabilities) in an economy and a better predictor of economic growth in the future. In addition, we use an empirical framework that has been shown to be robust to the endogeneity bias problem that often arise from the omission of relevant variables from a model, an error in the measurement of the variable(s) or due to simultaneous causality. The use of exogenous control variables and instruments allow for the capturing of the quality of political and economic institutions available while the panel transformations allow for the control of unobservable country and period effects.

Data Sources, Variables Description, and Empirical Method

For the analysis, annual data from 1982 to 2017 was collected for 31 Organization for Economic Co-operation and Development (OECD) countries. The availability of data and the statistical requirement for the panel to be relatively homogeneous for a GMM panel VAR model restricted our choice of countries and time period. All the data used in the analysis comes from the OECD statistical database, The World Bank's World Governance Indicators (WGI), and the Harvard University's Center for International Development "Atlas of Economic Complexity" statistical databases. The governance indicators used as control variables were extracted from The World Bank Governance Indicator database, while the real gross domestic product (GDP), total gross investment (gross-fixed capital formation), and the employment figures were extracted from the OECD database. The data for the ECI (*eci*) was obtained from the Atlas of economic complexity database. See Table A1 in the Appendix for the long definition of these variables and further information about their sources.

This article follows the traditional production function approach of dividing the output (the real GDP) and the capital accumulation (real gross-fixed capital formation) variables by the total number of people employed (cf. Barro et al., 2017) to generate proportional indicators. The new variables generated are the real GDP per worker (*rgdp*) and the gross-fixed capital formation per worker (*gfcf*). To linearise the relationship in the model, the natural logarithms of these variables is estimated. Since the GMM estimator used
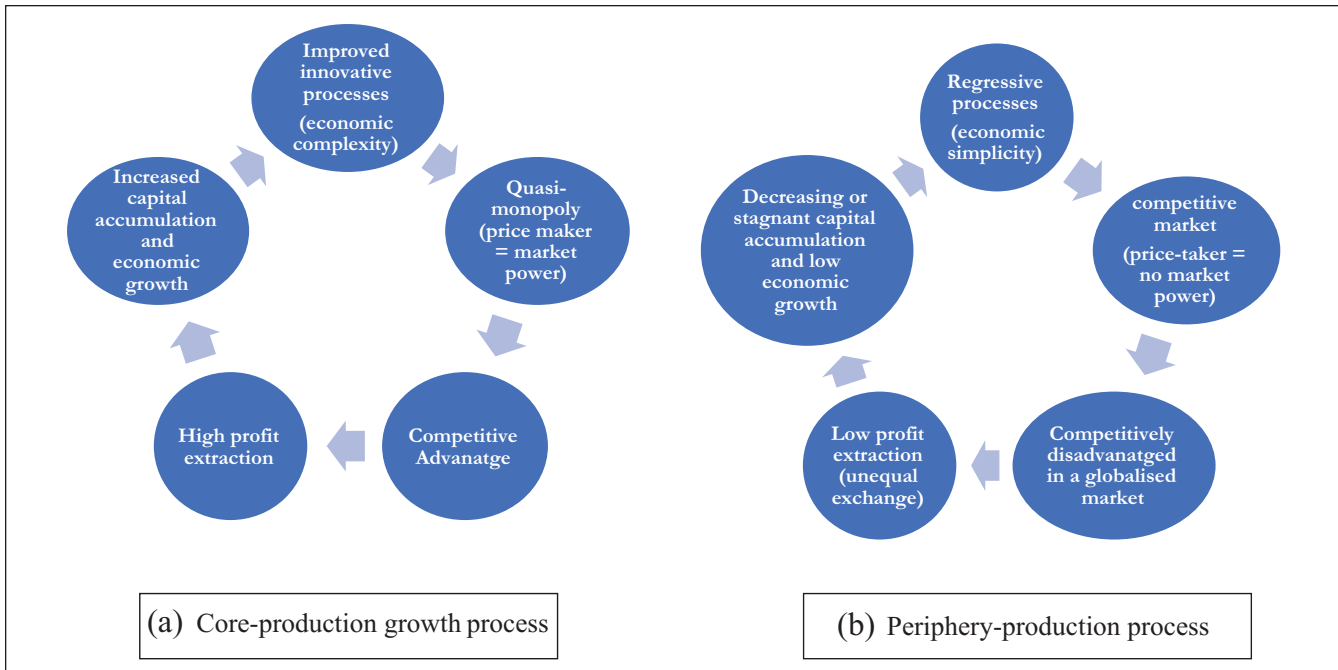
Figure 1. Theoretical Schemata of the complexity–growth nexus.
Note. Based on author's own analysis.

Table 1. Complexity of Economic Output of Countries in Different Income Groups.

Development category	Income group	Complexity level					
		Low-tech	Lower mid-tech	Mid-tech	Upper mid-tech	High-tech	Total
Developing	300–4,999	44	4	4	0	0	52
Emerging	5,000–19,999	3	12	6	4	1	26
Developed	20,000+	0	3	6	13	3	25

Note. Low-tech = ECI < 0, lower mid-tech = 0 ≤ ECI < 0.5, mid-tech = 0.5 ≤ ECI < 1, upper mid-tech = 1 ≤ ECI < 1.5, and high-tech = ECI ≥ 1.5. Authors' own calculation based on data from The World Bank's World Development Indicator (WDI) database and The Atlas of Economic Complexity database of the Center for International Development at Harvard University, http://www.atlas.cid.harvard.edu. ECI = economic complexity index.

in panel vector autoregressions suffer from weak instrument problems when the variable being modeled is near unit root (cf. Abrigo & Love, 2016), we sidestep the issue by specifying the reduced-form VAR model using variables in first differences (growth form). Using Stata's built-in xtunitroot command, we run panel unit-root tests on the transformed variables to confirm their assumed stationarity. We also conducted some other preliminary analyses, to test for individual effects, cross-sectional independence, and slope homogeneity because these determine whether an ordinary OLS regression, instead of a far more sophisticated model, suffices. To save space, the results of these analyses are moved to the Appendix.

Finally, we converted all the variables into 3-yearly non-overlapping averages to minimize the impact of short-term movements and ensure that the entities sampled are close as possible to their steady state throughout the study period. This is necessary because, as Roodman (2009) explained, the validity of the GMM technique depends on the assumption

that changes in the instrumenting variables are uncorrelated with the fixed effects: that is, they require that throughout the study period, individuals sampled are not too far from their long-run means.

This latter transformation also controls for period effects. Table 3 below contains the summary statistics of the transformed variables.

$$\Delta rgdp_{i,t} = \alpha_{i,11}\Delta rgdp_{i,t-1} + \alpha_{i,12}\Delta gfcf_{i,t-1} \\ + \alpha_{i,13}\Delta eci_{i,t-1} + Z_{i,1t} + \mu_{i,1t} + v_{i,1t} \quad (1)$$

$$\Delta gfcf_{i,t} = \alpha_{i,21}\Delta gfcf_{i,t-1} + \alpha_{i,22}\Delta rgdp_{i,t-1} \\ + \alpha_{i,23}\Delta eci_{i,t-1} + Z_{i,2t} + \mu_{i,2t} + v_{i,2t} \quad (2)$$

$$\Delta eci_{i,t} = \alpha_{i,31}\Delta eci_{i,t-1} + \alpha_{i,32}\Delta gfcf_{i,t-1} \\ + \alpha_{i,33}\Delta rgdp_{i,t-1} + Z_{i,3t} + \mu_{i,3t} + v_{i,3t} \quad (3)$$

Table 2. Top and Bottom 10 Countries by Economic Complexity.

Top 10	Country	GDP per capita (cons 2010 US$)	ECI
1	Austria	44,445.30	1.76
2	Japan	43,797.71	1.63
3	Czech Republic	17,679.84	1.54
4	Singapore	40,243.15	1.51
5	Switzerland	70,795.72	1.45
6	Germany	40,219.52	1.39
7	Korea, Rep.	18,817.82	1.33
8	Sweden	48,501.92	1.29
9	Ireland	47,486.89	1.27
10	United Kingdom	37,901.37	1.27
Bottom 10			
1	Nigeria	1,900	−2.07
2	Angola	2,964.89	−1.76
3	Sudan	1,317.25	−1.62
4	Guinea	616.89	−1.6
5	Cameroon	1,262.01	−1.44
6	Mauritania	1,131.44	−1.42
7	Republic of the Congo	2,548.67	−1.35
8	Democratic Rep. of the Congo	330.64	−1.22
9	Cote d'Ivoire	1,286.74	−1.22
10	Ghana	1,192.04	−1.2

Note. ECI & income average from 1996 to 2015. Authors' own calculation based on data from The World Bank's World Development Indicator (WDI) database and the "The Atlas of Economic Complexity," Center for International Development at Harvard University, http://www.atlas.cid.harvard.edu. The data cover over 103 countries. GDP = gross domestic product; ECI = Economic Complexity Index.

Table 3. Summary Statistics.

Variable	Panel section	M	SD	Minimum	Maximum	Observations
eci_g	Overall	0.88	13.40	−31.53	84.04	N = 328
	Between		3.40	−5.50	8.54	n = 31
	Within		12.99	−32.29	83.27	Avg. T = 10.58
lngfcf_g	Overall	10.86	12.29	−32.97	77.23	N = 294
	Between		4.76	3.07	28.26	n = 31
	Within		11.59	−41.17	69.03	Avg. T = 9.61
lngdp_g	Overall	11.83	5.25	−7.71	36.42	N = 294
	Between		3.15	6.29	20.38	n = 31
	Within		4.48	−9.38	28.69	Avg. T = 9.61

Source. Based on the authors' own calculation.
Note. To analyze the long-run relationship between the economic complexity index (Δ eci), capital per worker (Δ gfcf), and real output per worker (Δ rgdp), the following standard vector autoregressive model of order 1 is specified:

where $i = (1, \ldots, 31)$ and $t = (1, \ldots, 13)$. $\mu_{i,kt}$ and $v_{i,kt}$ ($k = 1, \ldots, 3$) are the unobserved time-invariant individual effects (such as geography, climate, belief system, colonial history, etc.) and idiosyncratic errors, respectively. The $Z_{i,t}$ are the exogenous control variables, which include the proxy for political stability (*polstab*), regulatory quality (*regq*), rule of law (*rule*), and corruption control (*coc*). These are not time invariant and do not change rapidly. Instead, they change slowly over time. Equations (1), (2), and (3) specify that $\Delta rgdp$, $\Delta gfcf$, and Δeci are linear functions of their own lags and the lags of the other endogenous variables as well as a

function of the contemporary exogenous and time-invariant factors, including the unobserved idiosyncratic errors. The maximum lag order of one is chosen based on the result of the consistent moment and model selection criteria (MMSC) proposed by Andrews and Lu (2001) for GMM models. Detailed discussion of the lag selection process is relegated to the Appendix to save space.

The specification of equations (1–3), which are a reduced form VAR(1), is shown to be consistent under the assumption that the projection errors are mean zero, uncorrelated with the lagged dependent and explanatory variables and are

serially uncorrelated (cf. Hansen, 2018). These assumptions can be stated formally as:

$$E(v_{i,kt}) = 0; \ E(v_{i,kt} \mid y_{i,t-1}) = 0; E(v_{i,kt-1} \mid v_{i,kt}) = 0 \quad (4)$$

Here, y is taken to include all the endogenous variables in the models.

Equations (4) infer that the projection errors are white noise processes and are seen to be strictly stationary and ergodic if $y_{i,t}$ are also strictly stationary. This assumption can be tested by applying the techniques proposed by Hamilton (1994) and Lutkepohl (2005) that show the VAR(1) process $y_{i,t}$ is strictly stationary and ergodic if the maximum value of the ith eigenvalues is less than 1. The result of this diagnostic is discussed in the next section.

As mentioned earlier, exogenous contemporaneous indicators are included in the panel VAR as control variables. These untransformed variables also serve as additional instrument for the model. We assume the contemporaneous conditioning variables, $z_{it} = \{z_{i,1}, \ldots, z_{i,T}\}$, to be appropriately strictly exogenous. The Hansen's J test of over-identification with a null hypothesis of exogenous and valid instruments is used to test this exogeneity assumption as well as the overall validity of our specifications. The extended equation modeled in the study is summarized as follows:

$$\Gamma = \left(y_{it} \mid y_{i,t-1}, z_{i,kt}, \mu_{i,kt}, v_{i,kt}\right) \quad (5)$$

To properly identify the parameters of equation (5), it is also important to deal with the presence of the assumed individual (country) effects. Studies (see Acemoglu et al., 2014, 2019; Acemoglu & Robinson, 2012, for example) have shown that several time-invariant factors such as geography and colonial history) are key determinants of economic growth in the long run.

Traditionally, a fixed-effect model would have sufficed to deal with the panel effect. However, in the context of a dynamic model, which is characterized by the presence of a lagged dependent variable among the regressors, it has been shown that the traditional OLS and within-group models would be inconsistent and inefficient in estimating the unknown parameters in the dynamic context. In particular, the inconsistency with the fixed-effect model, especially in a small-T large-N panel like ours, is seen to arise because the demeaning process which subtracts the cross-sectional mean from each variable, to account for the individual fixed effects from all the variables, creates a correlation between the regressors and the idiosyncratic error term, which then renders the within-group estimator inconsistent (see Baltagi, 2013 for the proof).

To control for individual effects in a dynamic panel equation, the difference-GMM approach proposed by Anderson and Hsiao (1982) and Arellano and Bond (1991) that involves first-differencing the equations to eliminate the individual

effects and subsequently instrumenting with lagged levels of the variables has often been used (see Acemoglu et al., 2014, 2019; Bond et al., 2010). However, the problem with this transformation is that it magnifies the gap in unbalanced panels and, by construction, introduces serial correlation in the model (cf. Abrigo & Love, 2016).

To reduce the potential bias and imprecision associated with the usual difference techniques in a dynamic context, the Arellano and Bover (1995) and Blundell and Bond (1998, 2000) proposed forward orthogonal deviation (FOD) approach is often adopted as an alternative transformation. The FOD, instead of using deviations from past realizations as with the first-difference approach, uses the average of all available future observations, which it subtracts from each preceding observation, thereby minimizing data loss. In essence, only the most recent observations are omitted in the estimation. Following Hansen (2018), if we denote the original variables as $y_{i,t}$, then the first difference transformation imply that $y^*_{i,t} = y_{i,t} - y_{i,t-1}$, while for the FOD $y^*_{i,t} = f_{i,t}$, $y_{i,t} - 1/T_i - t \ (y_{i,t+1} + \ldots + y_{i,Ti})$, where T_i is the number of available future observations for panel i at time t and $f_{it} = (T_i - t)/(T_i - t + 1)$. This transformation can be applied to all but the final observation, which is lost. Essentially, $y^*_{i,t}$ subtracts from $y_{i,t}$ the average of the remaining values, and then rescales so that the variance is constant under the assumption of homoscedastic errors. At the level of the individual, this transformation can be written as $y^*_i = \Lambda_i y_i$. Where Λ_i is the $(T_i-1) \times T_i$ orthogonal deviation operator. Roodman (2009) and Hansen (2018) contain an excellent discussion of the FOD GMM estimation in a dynamic panel setting. Abrigo and Love (2016) also contain a detailed explanation of the forward orthogonal transformation technique.

Applying the FOD transformation Λ_i to equation (5), we obtain

$$y^*_{i,t} = A_o + A_1 y^*_{i,t-1} + \beta z^{*'}_{it} + v^*_i \quad (6)$$

The orthogonality conditions of the transformed equation imply

$$E(v_i \mid y_{i,s}, z_i) = 0 \ \left(s < (t-1)\right) \quad (7)$$

Overall, the vector of GMM-style instrumental variables employed to identify the parameters of the transformed equation is then

$$\Pi = [y_{i,t-4}, y_{i,t-3}, y_{i,t-2}, z_{i,t}] \quad (8)$$

The FOD transformation eliminates any time-invariant (country-specific effect) that might have been correlated with the explanatory variables. Furthermore, since the equations uses current values of the dependent variables against the past values of the endogenous explanatory variables, the simultaneity of the system is also eliminated.

Table 4. pVAR Coefficients.

Variables	Δlngfcf	Δlngdp	Δeci
Δlngfcf$_{i,t-1}$	0.019 (0.121)	−0.031 (0.039)	0.312* (0.161)
Δlngdp$_{i,t-1}$	0.470 (0.348)	0.366** (0.167)	1.071* (0.620)
Δeci$_{i,t-1}$	0.038 (0.109)	0.124*** (0.039)	−0.058 (0.197)
polstab	−5.489 (15.949)	16.765*** (5.920)	−58.293** (26.866)
regq	−0.660 (21.130)	7.973 (9.282)	78.446** (37.005)
rule	3.653 (30.822)	−16.260 (11.611)	−96.249* (51.777)
coc	13.093 (19.875)	5.497 (7.352)	−92.182** (38.703)
Observations	235	235	235

Note. Based on the authors' estimates. Results are from the stata *pVAR* code developed by Abrigo and Love (2016). pVAR = panel vector autoregressive. Standard errors in parentheses.
*$p < .1$. **$p < .05$. ***$p < .01$.
Instruments: L(1/3).(*lngfcf_g lngdp_g eci_g*) polstab regq rule coc.
Hansen's *J* chi-square (18) = 13.883 (*p* = .737) .

Potential problems with serial correlation in the errors of each equation are subsequently eliminated by the incorporation of appropriate number of lags for each variable.

The above outlined FOD GMM panel VAR method has been shown to be consistent and efficient (see proof in Abrigo & Love, 2016) and have been widely applied in different contexts (see Carpenter & Demiralp, 2012; Head et al., 2016; Mora & Logan, 2012; Neumann et al., 2010; for instance). Additional tests were nevertheless conducted to ensure our chosen model is appropriate and yields consistent and efficient estimates, given the true nature of the panel data set. Discussion of some of these additional tests, such as the test for stationarity of the variables, cross-sectional independence, and slope homogeneity are relegated to the Appendix to save space.

Discussion of Results

In this article, we use the GMM panel VAR technique to investigate the dynamic impact of product complexity and diversity on economic growth. In practice, the panel vector autoregression estimates are rarely interpreted. Instead, researchers are often interested in the impact of changes in each independent variable on the dependent variable in the panel VAR system. For this, the impulse-response function (IRF) is often used. In our case, we are mainly interested in the response of the change in real output per worker to a change in the ECI. We also computed the Granger causality test for the first-order GMM panel VAR, to confirm whether past values of the ECI are useful in predicting future values of output per worker, conditional on the past rate of output growth. The coefficient estimates and robust standard errors of the VAR(1) regression are reported in Table 4 while the Granger Causality Wald test, with the null hypothesis that the coefficients on all the lags of an endogenous variable are jointly equal to 0 are reported in Table 5. The IRFs are reported in Figure 3.

According to the Granger causality test results in Table 5 above, changes in economic complexity and diversity can be

Table 5. Granger Causality Test.

Equation	Excluded	Chi-square	df	prob
Δ lngfcf				
	Δ lngdp	1.822	1	0.177
	Δ eci	0.123	1	0.726
	ALL	2.217	2	0.330
Δ lngdp				
	Δ lngfcf	0.656	1	0.418
	Δ eci	9.857	1	0.002
	ALL	9.908	2	0.007
Δ eci				
	Δ lngfcf	3.745	1	0.053
	Δ lngdp	2.988	1	0.084
	ALL	7.729	2	0.021

Note. Based on the authors' estimates. pVAR Granger stata code used. H_0 = excluded variable does not Granger-cause equation variable; H_1 = excluded variable Granger-causes equation variable.

said to Granger-cause real GDP per capita growth at the usual confidence levels (row two). The results also confirm that the coefficients of the two independent variables are jointly different from 0, as such cannot be excluded in the equation of economic growth. Likewise, we find that growth in output per worker and growth in capital per worker Granger-causes changes in economic complexity, at the 90% confidence interval (row 3).

Given that the coefficients on the reduced-form panel VARs cannot be interpreted as causal influences without imposing identifying restrictions on the parameter, we resort to the use of IRFs, which have been shown to have known interpretation when the panel VAR model is stable. Stability of the VAR allows it to be reformulated as an infinite-order vector moving average (VMA), on which assumptions about the error covariance matrix may be imposed (cf. Abrigo & Love, 2016; Hansen, 2018)

The result of the stability test, presented in Figure 2 and Table 6 below, expectedly confirm that our panel VAR is
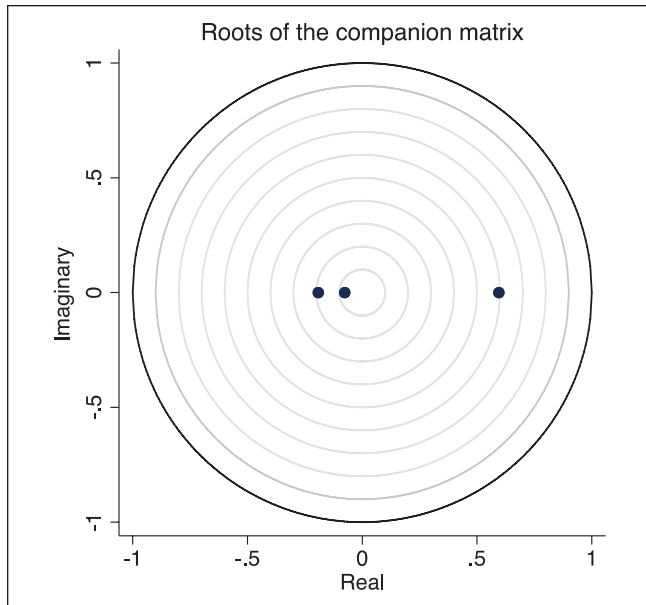
Figure 2. Stability test plot.
Note. All the eigenvalues lie inside the unit circle. Thus, pVAR satisfies stability condition. pVAR = panel vector autoregressive.

invertible (consistent) and has an infinite-order VMA representation, given that the moduli of the companion matrix are all less than one.

For the IRFs, we compute the orthogonalised IRF and the cumulative IRF based on Cholesky decomposition. Confidence intervals are computed using 200 Monte Carlo draws on the estimated model. Since the GMM panel VAR is stable, the shocks are expected to converge to zero in the long run. That is, the shocks are anticipated to be temporary, and over the long run, the series are forecasted to return to their deterministic (long-run) trend.

As mentioned, the endogenous variables are ordered as per the Cholesky decomposition. Specifically, we contend, based on intuition, that increase in capital per worker affects output growth contemporaneously, while shocks on the ECI affect output growth with a lag. The result of the Cholesky forecast-error variance decomposition (FEVD) lends support to this ordering, as it shows the impact of the complexity index to be noncontemporaneous while the impact from economic growth on *eci* is relatively contemporaneous. The result also shows a relatively contemporaneous impact of capital accumulation (investment) on output growth.

Furthermore, given that the contributions of the exogenous variables included in the panel VAR model are disregarded when computing the forecast-error variance, we will only be focusing on the impact of the ECI and the capital accumulation variable on output growth in the rest of the discussion.

The result of the IRF (Figure 3) show the response of economic (output per worker) growth to a shock in *eci*. The IRF suggest that the *eci* has a significant impact on economic

Table 6. Eigenvalue Stability Condition Table.

Eigenvalue		
Real	Imaginary	Modulus
.595189	0.0000	0.595189
−.1916031	0.0000	0.1916031
−.0761707	0.0000	0.0761707

Source. Based on the authors' own calculation using the pVAR stable program provided by Abrigo and Love (2016).
Note. All the eigenvalues lie inside the unit circle. Thus, pVAR satisfies stability condition.

growth; a 1 standard deviation shock to the ECI at time 0 contributes around 2.34% to the average rate of growth of real GDP per worker within the first period. The point estimates are positive and significant for most horizons up to the 6th period. These results confirm the hypothesis that the ECI is a good predictor of economic growth. The cumulative IRF shows the long-run impact of a standard deviation shock to the ECI on the real output per worker. The result shows that the cumulative increase in the rate of economic growth would be circa 4.4% by the 10th period (Figure 4).

The effect of capital accumulation on economic growth was also examined. The response of economic growth to a shock in capital per worker is contemporaneous; economic growth reaches a level approximately 1.8 percentage points from the mean in the same period and stays nearly 0.14 percentage points higher up to the fifth period if there is a 1 standard deviation shock to capital per worker (Figure 5).

Finally, we compared the performance of the ECI to those of the two widely used technical change proxies—the rate of increase in the gross expenditure on research and development (*gerd*) and the secondary school enrollment (*sec*) rate. The Granger causality test results for the *sec* and *gerd* equations are reported in Tables 7 and 8, respectively. The VAR coefficients are not reported, to save space. From Table 7, we find that the *sec* term is not statistically different from zero, suggesting that the rate of secondary school enrollment does not Granger cause growth in output per worker. Similarly, results in Table 8 indicate that the rate of gross expenditure on research and development does not Granger cause growth in economic output. These results are undoubtedly plausible, given that not all expenditure on research and development produce significant outputs and likewise increasing secondary school enrollment does not necessarily translate into increasing productivity and output, especially when the quality of the education is not properly accounted for in the enrollment figures.

Concluding Remarks

Using data from over 31 OECD countries, covering periods from 1982 to 2017, this study tests the hypothesis that the
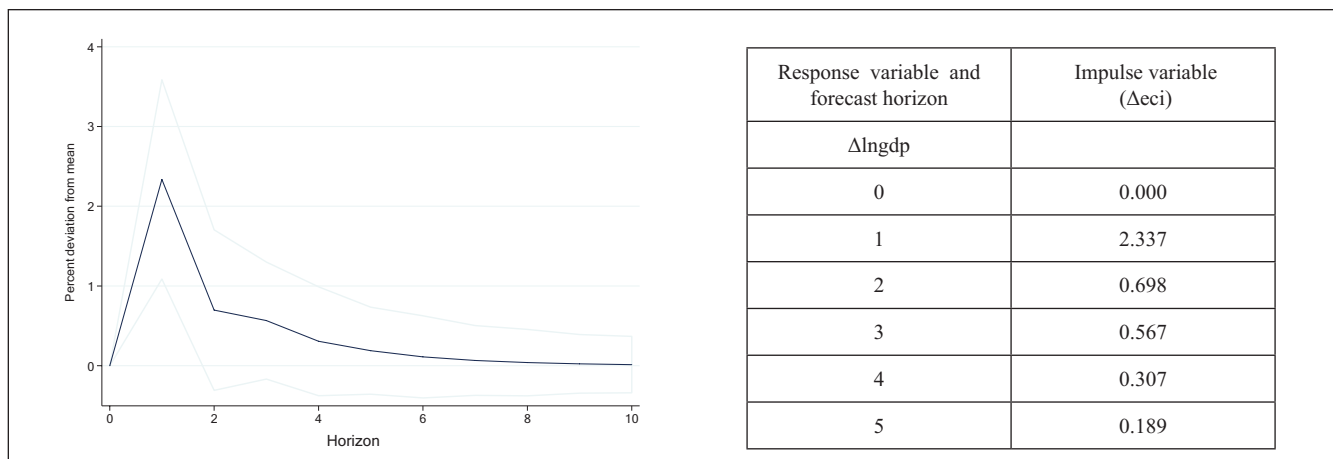
Response variable and forecast horizon	Impulse variable (Δeci)
Δlngdp	
0	0.000
1	2.337
2	0.698
3	0.567
4	0.307
5	0.189

Figure 3. Response of GDP growth to a shock to economic complexity.
Note. Based on the authors' estimate. The orthogonalized IRF chart and table show the response to Cholesky 1 SD innovations ± 2 SE. GDP = gross domestic product; IRF = impulse-response function.
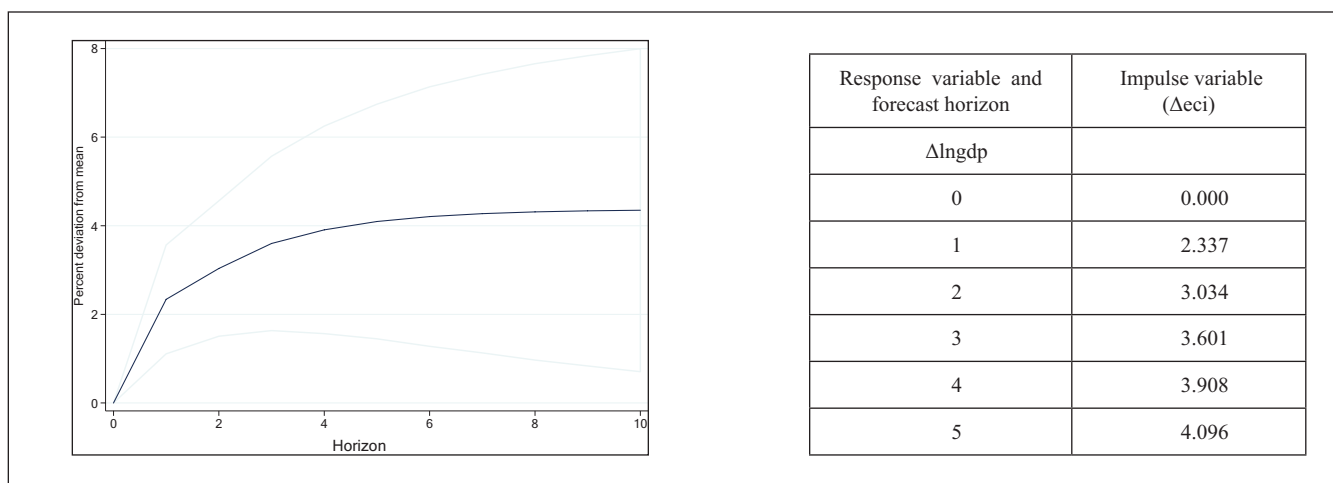

Response variable and forecast horizon	Impulse variable (Δeci)
Δlngdp	
0	0.000
1	2.337
2	3.034
3	3.601
4	3.908
5	4.096

Figure 4. Cumulative response of output growth to a shock to economic complexity.
Note. Based on the authors' estimate. The cumulative orthogonalized IRF chart and table show the response to Cholesky 1 SD innovations ± 2 SE. IRF = impulse-response function.
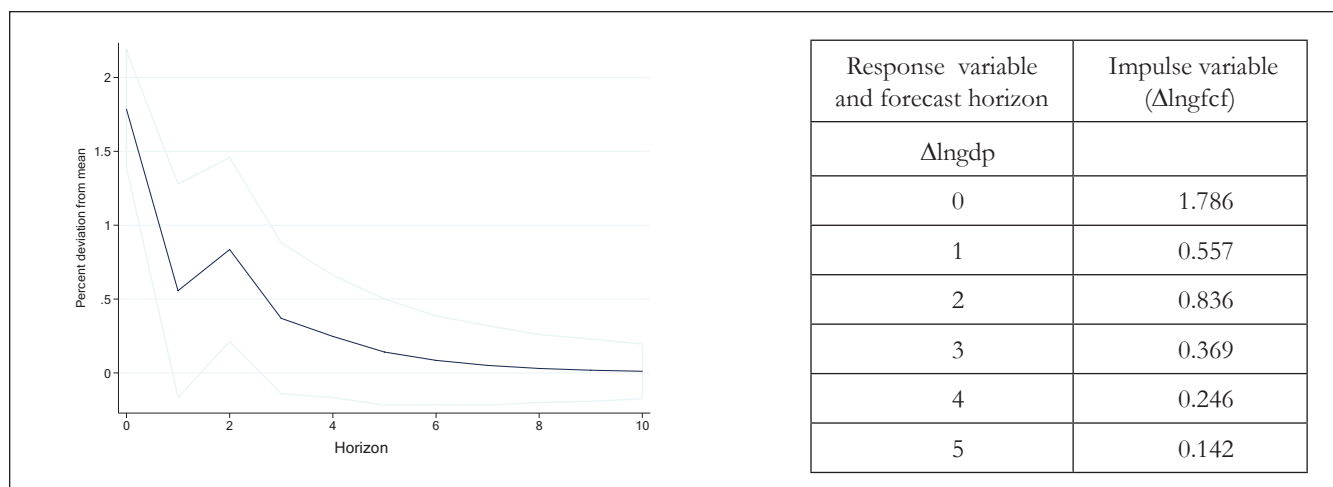

Response variable and forecast horizon	Impulse variable (Δlngfcf)
Δlngdp	
0	1.786
1	0.557
2	0.836
3	0.369
4	0.246
5	0.142

Figure 5. Response of GDP growth to a shock to capital per worker.
Note. Based on the authors' estimate. The orthogonalized IRF chart and table show the response to Cholesky 1 SD innovations ± 2 SE. GDP = gross domestic product; IRF = impulse-response function.

Table 7. Granger Test (sec).

Equation	Excluded	Chi-square	df	Prob
Δ *lnsec*				
	Δ *lngfcf*	0.059	1	.809
	Δ *lngdp*	4.161	1	.041
	ALL	5.908	2	.052
Δ *lngfcf*				
	Δ *lnsec*	2.026	1	.155
	Δ *lngdp*	3.55	1	.06
	ALL	4.657	2	.097
Δ *lngdp*				
	Δ *lnsec*	0.508	1	.476
	Δ *lngfcf*	2.821	1	.093
	ALL	2.914	2	.233

Table 8. Granger Test (Gerd).

Equation	Excluded	Chi-square	df	Prob
Δ *lngerd*				
	Δ *lngfcf*	0.331	1	.565
	Δ *lngdp*	1.381	1	.24
	ALL	2.42	2	.298
Δ *lngfcf*				
	Δ *lngerd*	1.897	1	.168
	Δ *lngdp*	6.447	1	.011
	ALL	6.784	2	.034
Δ *lngdp*				
	Δ *lngerd*	0.258	1	.611
	Δ *lngfcf*	3.982	1	.046
	ALL	4.132	2	.127

Note. Authors' own estimate. H_0 = excluded variable does not Granger-cause equation variable; H_1 = excluded variable Granger-causes equation variable.

level of product complexity is a good determinant of economic growth in the long run. The regression is estimated using the generalized method of moment panel vector autoregressive method, which has been shown to be consistent and efficient when endogeneity bias—caused by the omission of relevant variables from a model or an error in the measurement of the variable or due to simultaneous causality between the dependent and independent variable—is present in a model. The paper uses 3-yearly nonoverlapping averages, instead of the annual series, to minimize the impact of short-term shocks. The forward orthogonal transformation is used to control the country effects, and exogenous governance indicators are used to control the effect of factors that change

slowly over time, such as the role of law, political stability, regulatory quality, and corruption control.

The results of the empirical analyses show that changes in the level of product complexity and diversity, as measured by the ECI, does indeed determine economic growth in the long run. The IRF shows that a 1 standard deviation shock to the ECI at time 0 contributes around 2.34% to the average rate of growth of real output per worker within the first period. The point estimates are positive and significant for most horizons up to the sixth period. Furthermore, the cumulative IRF shows that the cumulative increase in the rate of economic growth would be around 4.4% by the fifth period.

Compared to traditional technical change proxies such as the gross expenditure on research and development and secondary school enrollment, the ECI performs relatively better in determining economic growth in the long run. The coefficients on both the secondary school enrollment variable and the expenditure on research and development variable appear statistically indifferent from 0 in the empirical result.

Overall, this study offers a substantial quantitative improvement over past studies that have examined the innovation/complexity–growth nexus. The model developed in this article generates dynamics quantitively consistent with several development theories and observations that indicate product complexity and diversity or innovation determines the pace of economic growth *cum* development in the long run.

The results of this study also contain some policy implications that are certainly noteworthy. The finding is particularly significant for economies both in the OECD and other regions looking for effective ways to stimulate meaningful growth in their economies. The results confirm that by restructuring the economy, particularly by mobilizing resources toward knowledge-intensive manufacturing as well as diversifying exports, countries can stimulate the prerequisite output growth capable of instigating sustainable economic development in the long run. Perhaps the most telling is the confirmation that improving the set of human capital—both the tacit and explicit knowledge, which can be achieved through quality education, nutrition and health care, and investing in physical capital, which includes improvement in available tools and machines and relevant infrastructure (roads, railways, sea and air ports, electricity, internet, etc.), including the enhancement of institutions (economic, political, and social) that exist in a place and time and encourages capital accumulation as well as accelerates economic complexity, are the key engines of growth. This finding helps to take the guesswork away from effective industrial policy design.

Appendix A

Table A1. Variables Definition and Sources.

Variables	Indicator name	Long definition	Sources
rgdp	GDP (constant 2010 US$)	GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in constant 2010 U.S. dollars. Dollar figures for GDP are converted from domestic currencies using 2010 official exchange rates. For a few countries where the official exchange rate does not reflect the rate effectively applied to actual foreign exchange transactions, an alternative conversion factor is used.	OECD National Accounts data files, https://data.oecd.org
gfcf	Gross-fixed capital formation (constant 2010 US$)	Gross-fixed capital formation (formerly gross domestic investment) consists of outlays on additions to the fixed assets of the economy plus net changes in the level of inventories. Fixed assets include land improvements (fences, ditches, drains, and so on); plant, machinery, and equipment purchase; and the construction of roads, railways, and the like, including schools, offices, hospitals, private residential dwellings, and commercial and industrial buildings. Inventories are stocks of goods held by firms to meet temporary or unexpected fluctuations in production or sales, and "work in progress." This is according to the 1993 SNA; net acquisitions of valuables are also considered capital formation. Data are in constant 2010 U.S. dollars.	OECD National Accounts data files, https://data.oecd.org
eci	Economic complexity index	ECI is derived from the method of reflections of Hidalgo and Hausmann (2009). The index combines information on the diversity of a country (the number of products it exports), and the ubiquity of its products (the number of countries that export that product) to reflect the diversity and complexity of capabilities (that is, the knowledge, know-how, social capital, institutions, etc.) available in a country and their interactions.	"The Atlas of Economic Complexity," Center for International Development at Harvard University, http://www.atlas.cid.harvard.edu.
polstab	Political stability and absence of violence/terrorism: estimate	Political stability and absence of violence/terrorism measures perceptions of the likelihood of political instability and/or politically motivated violence, including terrorism. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, that is, ranging from approximately −2.5 to 2.5.	The World Bank's WGI database, https://databank.worldbank.org/source/worldwide-governance-indicators
regq	Regulatory quality: estimate	Regulatory quality captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, that is, ranging from approximately −2.5 to 2.5.	The World Bank's WGI database, https://databank.worldbank.org/source/worldwide-governance-indicators
rule	Rule of law	Rule of law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, that is ranging from approximately −2.5 to 2.5.	The World Bank's WGI database, https://databank.worldbank.org/source/worldwide-governance-indicators
Coc	Control of corruption	Control of corruption captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, that is ranging from approximately −2.5 to 2.5.	World Bank's WGI database, https://databank.worldbank.org/source/worldwide-governance-indicators

Note. GDP = gross domestic product; OECD = Organization for Economic Co-operation and Development; ECI = Economic Complexity Index; WGI = World Governance Indicator.

Appendix B

Unit Root Testing

The results from the Fisher-type unit-root tests also confirm the stationarity of $y_{i,t}$ thus stationarity and ergodicity of the white noise processes. The results of the tests are presented below.

Table B1. Test of Stationarity (Fisher-Type Unit Root/Phillips–Perron Tests).

(4a) Unit-root test for $\Delta\,rgdp_{i,t}$	Statistic	p value
Inverse chi-squared (62), P	658.291	.000
Inverse normal, Z	−21.666	.000
Inverse logit t(159), L*	−32.667	.000
Modified inv. chi-squared, Pm	53.549	.000
(4b) Unit-root test for $\Delta\,gfcf_{i,t}$	**Statistic**	**p value**
Inverse chi-squared(62), P	408.659	.000
Inverse normal, Z	−15.650	.000
Inverse logit t(159), L*	−20.237	.000
Modified inv. chi-squared, Pm	31.131	.000
(4c) Unit-root test for $\Delta\,eci_{i,t}$.	**Statistic**	**p value**
Inverse chi-squared (62), P	760.443	.000
Inverse normal, Z	−24.212	.000
Inverse logit t(159), L*	−37.769	.000
Modified inv. chi-squared, Pm	62.722	.000

Note: AR parameter is panel specific; panel means and time trend are included; Newey–West lags used is 1 lag; H_0: all panels contain unit roots vs. H_1: At least one panel is stationary; number of panels = 31; Avg. number of periods = 31.23. Source: Based on the authors' own calculation.

Appendix C

The Breusch and Pagan (1980) proposed that Lagrange multiplier (LM) test is used to test the presence of panel effect in the model—that is, whether there is significant differences across the units. The null hypothesis of the LM test is that variances across the entities is zero. So, the failure to reject the null hypothesis means the OLS would be a more appropriate estimator to use. The Pesaran (2007) test is used to assess whether there is a contemporaneous correlation among the residuals. One of the assumptions of the GMM model is that there is no cross-sectional dependence. The null of the

Pesaran cross-section dependence (CD) test is that residuals are not correlated. To be consistent, we would expect to accept the null of the Pesaran CD test. To test for homoskedasticity or constant variance, we applied the modified Wald test for groupwise heteroskedasticity. In a fixed-effect regression model, the test has a null of homoskedasticity or constant variance. Results of these preliminary tests show that the use of the GMM pVAR is relevant and appropriate in this context, given the presence of individual effects and cross-sectional independence. The presence of heteroskedasticity in the model requires the use of robust standard errors, which we applied in the main model accordingly.

Table C1. Preliminary Analysis.

Variables	FE	RE
lngfcf_g	0.186*** (0.022)	0.224*** (0.023)
L.lngfcf_g	−0.017 (0.024)	0.032 (0.024)
eci_g	−0.017 (0.018)	−0.015 (0.019)
L.eci_g	0.061** (0.026)	0.091*** (0.026)
polstab	2.801* (1.514)	1.950*** (0.553)
regq	0.362 (2.240)	−0.145 (1.394)
rule	−4.300 (3.322)	−0.605 (1.560)
coc	−1.036 (2.020)	−0.992 (1.099)
Constant	10.948*** (3.485)	7.033*** (1.234)

(continued)

Table C1. (continued)

Variables	FE	RE
R^2 (within)	.556	.529
R^2 (between)	.213	.621
R^2 (overall)	.415	.533
F test	0.000	0.000
rho	0.452	0.034
Breusch–Pagan LM test for fixed effect	0.000	N/A
Breusch-Pagan LM test for random effect	N/A	0.004
Pesaran CD test	N/A	0.127
Wald groupwise heteroscedasticity test	0.000	N/A
Observations	266	266
Number of *cid*	31	31

Note. Calculations are based on the authors' own analysis. Standard errors in parentheses. LM = Lagrange Multiplier; CD = Cross-Section Dependence.
*$p < .1$. **$p < .05$. ***$p < .01$.

Appendix D

Lag Length Selection Criteria

A consistent panel VAR analysis is remarked to be dependent on choosing the optimal lag order in both panel VAR specification and moment condition (cf. Abrigo & Love, 2016). So, to ensure consistency of the model applied in this study, the three consistent MMSC proposed by Andrews and Lu (2001) for GMM models are used to select the appropriate lag lengths for the proposed model. The MMSC, as noted by Abrigo and Love (2016), is analogous to most of the commonly used maximum likelihood-based model selection criteria such as the Akaike Information Criteria (AIC) (Akaike,

1969), the Bayesian Information Criteria (BIC) (Akaike, 1977; Schwarz, 1978), and the Hannan–Quinn Information Criteria (HQIC) (Hannan & Quinn, 1979). The MMSC models include the Moments of Bayesian Information Criteria (MBIC), the Moments of Akaike Information Criteria (MAIC), and the Moments of Hannan–Quin Information Criteria (MQIC) techniques. The result of the MMSC is presented below.

Based on the three model selection criteria, a first-order panel VAR is the preferred model because it has the smallest coefficient of determination (CD), MBIC, MAIC, and MQIC. Accordingly, based on the Andrews and Lu selection criteria, a first-order panel VAR model is fitted.

Table D1. Lag Selection Criteria.

Lag	CD	J	J p value	MBIC	MAIC	MQIC
1	0.6153130	50.46843	.266227	−148.9183	−39.53157	−83.50413
2	0.7786297	39.48731	.3168689	−120.0221	−32.51269	−67.69074
3	0.8298611	33.48552	.1815167	−86.14654	−20.51448	−46.89802
4	0.5502893	20.4248	.3094184	−59.32990	−15.5752	−33.16422

Source. Based on the authors' own calculation. CD = coefficient of determination; MBIC = Moments of Bayesian Information Criteria; MAIC = Moments of Akaike Information Criteria; MQIC = Moments of Hannan-Quin Information Criteria.

Appendix E

Test of Slope Homogeneity

In empirical studies, for there to be consistent and efficient estimates from standard aggregating or pooling methods—such as the fixed and random effects and the generalized method of moments methods—the slope coefficients need to homogeneous across individual units, particularly for dynamic panels (Pesaran & Smith, 1995). Pesaran and Smith showed that if the true model consists of heterogeneous slopes, imposing slope homogeneity, by pooling or aggregating the heterogenous dynamic panel, produces very misleading estimates. In essence, testing for slope homogeneity is

important to ensure the selection of appropriate econometric method and the estimation of consistent results.

In panel data four procedures are widely used to estimate the unknown parameters. These are pooling, aggregating, averaging group estimates, and cross-section regressions. In the static case, if the coefficients differ randomly, studies have shown that all four procedures give unbiased estimates of coefficient means regardless. In the dynamic case, however, when the true panel consists of heterogenous coefficients, pooling, and aggregating, such as with fixed- and random-effect and GMM models, have been shown to give inconsistent and potentially highly misleading estimates of the coefficients, though the cross-section can provide consistent estimates of the long-run effects.

Using the modified version of Swamy's test of slope homogeneity proposed by Pesaran and Yamagata (2008), for panel data models where the cross-section dimension (N) could be large relative to the time-series dimension (T), we test the null hypothesis of homogeneous slopes that implies all slope coefficients are identical across cross-sectional units. For our GMM pVAR estimates to be consistent and efficient, we would expect to accept the null of homogeneous slopes. A rejection of the null will however mean alternative panel procedures, such as the cross-section or mean group method, would suffice in our analysis.

The result of the slope homogeneity test, presented below, is computed using stata's xthst command. We use the HAC consistent test static derived by Blomquist and Westerlund (2013) to ensure robust standard errors, given the presence of heteroskedasticity in our model.

Table E1. Testing for Slope Heterogeneity.

	Delta	*p* value
Model I—non-adj. delta	0.453	.651
—adjusted delta	1.114	.265
Modell II—non-adj. delta	−1.237	.216
—adjusted delta	—	—

Source. Based on the authors' calculations, using stata's xthst command.
Note. H_0: slope coefficients are homogeneous. HAC Kernel: bartlett. Variables partialled out: constant.

Model I is a static model that does not include any lags of the dependent, independent, and control/exogenous variables. Model II however is specified as an ARDL (1,0) model that included the first lag of the dependent variable. In both specifications, only the constant is partialled out—assumed to be heterogenous in the panel. The test results for both specifications failed to reject the null of slope homogeneity, pointing to the consistency of our GMM pVAR estimates, which are efficient and robust under the assumption of slope homogeneity in a $N > T$ dynamic panel data settings.

Availability of Data and Material

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declaration of Conflicting Interests

Funding

ORCID iDs

Ejike Udeogu https://orcid.org/0000-0003-1626-2796
Uzochukwu Amakom https://orcid.org/0000-0002-7036-4278

References

Abdon, A., & Felipe, J. (2011). *The product space: What does it say about the opportunities for growth and structural transformation of Sub-saharan Africa?* (Levy Economics Institute Working Paper Archive No, wp_670). https://ideas.repec.org/p/lev/wrkpap/wp_670.html

Abrigo, M. R. M., & Love, I. (2016). Estimation of panel vector autoregression in stata. *Stata Journal*, *16*(3), 778–804.

Acemoglu, D., Gallego, F. A., & Robinson, J. A. (2014). Institutions, human capital, and development. *Annual Review of Economics*, *6*(1), 875–912.

Acemoglu, D., Naidu, S., Restrepo, P., & Robinson, J. A. (2019). Democracy does cause growth. *Journal of Political Economy*, *127*(1), 47–100.

Acemoglu, D., & Robinson, J. (2012). *Why nations fail: The origins of power, prosperity, and poverty*. Crown.

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

Aghion, P., & Howitt, P. (1998a). Capital accumulation and innovation as complementary factors in long-run growth. *Journal of Economic Growth*, *3*, 111–130.

Aghion, P., & Howitt, P. (1998b). *Endogenous growth theory*. MIT Press.

Akaike, H. (1969). Fitting Autoregressive Models for Prediction. *Annals of the Institute of Statistical Mathematics*, *21*, 243–247. https://doi.org/10.1007/BF02532251

Akaike, H. (1977). Likelihood of a model and information criteria, *Journal of Econometrics*, *16*(1), 3–14.

Anderson, T. W. and Hsiao, C. (1982). Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, *18*(1), 47–82.

Andrews, D. W. K., & Lu, B. (2001). Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models. *Journal of Econometrics*, *101*(1), 123–164.

Arellano, M., & Bond, S. (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 variable estimation of error-components models. *Journal of Econometrics*, *68*, 29–51.

Baltagi, B. H. (2013). *Econometric analysis of panel data*. John Wiley.

Barro, R. (1991). Economic growth in a cross section of countries. *Quarterly Journal of Economics*, *106*(2), 407–443.

Barro, R. (1997). *Determinants of economic growth: A cross-country empirical study*. MIT Press.

Barro, R. (2001). Human capital and growth. *American Economic Review*, *91*(2), 12–17.

Barro, R., Chu, A., & Cozzi, G. (2017). *Intermediate macroeconomics*. Cengage.

Bilbao-Osorio, B., & Rodríguez-Pose, A. (2004). From R&D to innovation and economic growth in the EU. *Growth and Change*, *35*, 434–455.

Blomquist, J., & WesterLund, J. (2013). Testing slope homogeneity in large panels with serial correlation. *Economic Letters*, *121*(3), 374–378.

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

Blundell, R., & Bond, S. (2000). GMM estimation with persistent panel data: An application to production functions. *Econometric Reviews*, *19*(3), 321–340.

Bond, S., Leblebicioglu, A., & Schiantarelli, F. (2010). Capital accumulation and growth: A new look at the empirical evidence. *Journal of Applied Econometrics*, *25*(7), 1073–1099.

Breusch, T. S., & Pagan, A. R. (1980). The lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies*, *47*(1), 239–253.

Caldarelli, G., Cristelli, M., Gabrielli, A., Pietronero, L., Scala, A., & Tacchella, A. (2012). A network analysis of countries' export flows: Firm grounds for the building blocks of the economy. *PLoS ONE*, *7*(10), Article e47278.

Carpenter, S., & Demiralp, S. (2012). Money, reserves, and the transmission of monetary policy: Does the money multiplier exist? *Journal of Macroeconomics*, *34*, 59–75.

Coe, D. T., & Helpman, E. (1995). International R&D spillovers. *European Economic Review*, *39*(5), 859–887.

Cristelli, M., Tacchella, A., & Pietronero, L. (2015). The heterogenous dynamics of economic complexity. *PLoS ONE*, *10*(2), Article e0117174.

Dos Santos, T. (1970). The structure of dependence. *American Economic Review*, *60*(2), 231–236.

Frank, G. (1967). *Capitalism and underdevelopment in Latin America*. Monthly Review Press.

Griliches, Z. (1994). Productivity, R&D, and the data constraint. *American Economic Review*, *84*, 1–23.

Griliches, Z. (1998). *R&D and productivity: The econometric evidence*. University of Chicago Press.

Griliches, Z., & Mairesse, J. (1990). R&D and productivity growth: Comparing Japanese and U.S. manufacturing firms. In C. R. Hulten (Ed.), *Productivity growth in Japan and the United States* (pp. 317–348). Chicago University Press. https://doi.org/10.7208/9780226360607-014

Grossman, G. M., & Helpman, E. (1991a). *Innovation and growth in the global economy*. MIT Press.

Grossman, G. M., & Helpman, E. (1991b). Trade, knowledge spillovers, and growth. *European Economic Review*, *35*, 517–526.

Grossman, G. M., & Helpman, E. (1994). Endogenous innovation in the theory of growth. *Journal of Economic Perspectives*, *8*(1), 23–44.

Guellec, D., & Potterie, B. (2003). The impact of public R&D expenditure on business R&D. *Economics of Innovation and New Technology*, *12*, 225–243.

Hall, B. H., Mairesse, J., & Mohnen, P. (2010). Measuring the returns to R&D. In H. B. Hall & N. Rosenberg (Eds.), *Handbook of the economics of innovation* (Vol. 2., pp. 1033–1082). Elsevier.

Hannan, E. J. & Ouinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, *41*(2), 190–195.

Hansen, E. B. (2018). *Econometrics*. University of Wisconsin.

Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.

Hartmann, D., Guevara, R. M., Jara-Figueroa, C., Aristaran, M., & Hidalgo, A. C. (2017). Linking economic complexity, institutions, and income inequality. *World Development*, *93*, 75–93.

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Chung, S., Jimenez, A., Simoes, A., & Yildirim, M. (2014). *The atlas of economic complexity*. Puritan Press.

Head, A., Lloyd-Ellis, H., & Sun, H. (2016). Search, liquidity, and the dynamics of house prices and construction. *American Economic Review*, *106*, 1214–1219.

Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, *106*(26), 10570–10575.

Hirschman, A. O. (1958). *The strategy of economic development*. Yale University Press.

Jones, C. (1995). Time series tests of endogenous growth models. *The Quarterly Journal of Economics*, *110*, 495–525.

Kemeny, T., & Storper, M. (2014). Is specialization good for regional economic development? *Regional Studies*, *49*(6), 1003–1018.

Kotz, D. M., McDonough, T., & Reich, M. (1994). *Social structures of accumulation: The political economy of growth and crisis*. Cambridge University Press.

Levine, R., Loayza, N., & Beck, T. (2000). Financial intermediation and growth: Causality and causes. *Journal of Monetary Economics*, *46*(1), 31–77.

Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The Manchester School*, *22*(2), 139–191.

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

Lutkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, *107*(2), 407–437.

Mohnen, P. (1990). New technologies and inter-industry spillovers. *STI Review, (OECD)*, 131–147.

Mohnen, P., & Lépine, N. (1991). R&D, R&D spillovers and payments for technology: Canadian evidence. *Structural Change and Economic Dynamics*, *2*, 213–228.

Mora, N., & Logan, A. (2012). Shocks to Bank Capital: Evidence from UK Banks at Home and Away. *Applied Economics*, *44*, 1103–1119.

Neumann, T. C., Fishback, P. V., & Kantor, S. (2010). The dynamics of relief spending and the private urban labour market during the new deal. *Journal of Economic History*, *70*, 195–220.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, *22*(2), 265–312.

Pesaran, M. H., & Smith, R. (1995). Estimating long-run relationships from dynamic heterogenous panels. *Journal of Econometrics*, *68*(1), 79–113.

Pesaran, M. H., & Yamagata, T. (2008). Testing slope homogeneity in large panels. *Journal of Econometrics*, *142*(1), 50–93.

Prebisch, R. (1963). *Towards a dynamic development policy for Latin America*. United Nations.

Rebelo, S. (1991). Long-run policy analysis and long-run growth. *Journal of Political Economy*, *99*, 500–521.

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

Romer, P. M. (1987). Growth based on increasing returns due to specialization. *American Economic Review*, *77*, 56–62.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*, 71–102.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal*, *9*(1), 86–136.

Rosenstein-Rodan, P. N. (1943). Problems of industrialisation of eastern and south-eastern europe. *The Economic Journal*, *53*(210/211), 202–211.

Schwarz, G. (1978). *The Annals of Statistics*, 6(2), 461–464.

Singer, H. W. (1950). The distribution of gains between investing and borrowing countries. *American Economic Review*, *40*(2), 473–485.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, *70*(1), 65–94.

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

Wallerstein, I. (2004). *World-systems analysis: An introduction*. Duke University Press.

Williamson, J. G. (2013). *Trade and poverty: When the third world fell behind*. MIT Press.

Zachariadis, M. (2003). R&D, innovation, and technological progress: A test of the schumpeterian framework without scale effects. *Canadian Journal of Economics*, *36*(3), 566–586.