CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

New results and a model of scale effects on growth

Kul Luintel and Panayiotis Pourpourides (ID)

Cardiff Business School, Cardiff University, Cardiff, UK
Corresponding author: Panayiotis Pourpourides; Email: pourpouridesp@cardiff.ac.uk

Abstract

The consensus view in the growth literature is that R&D scale effects are absent in mature industrialized economies but may be present in emerging economies undergoing transition. Scale effects imply a proportional relationship between a stationary $I(0)$ regressand (growth rates of real per capita GDP and/or TFP) and a non-stationary $I(1)$ regressor (the scale of R&D), which gives rise to the problem of unbalanced regression and spurious parameter estimates. This issue has not been adequately addressed in the existing literature. Furthermore, emerging economies have received relatively little attention in this context. We address these issues by (i) accurately measuring R&D scale and (ii) adopting an appropriate econometric specification and estimator. We find significant scale effects in a panel of emerging countries, but not in developed countries. We propose an endogenous growth model that captures these properties—presence of scale effects during growth transitions, but not at the long-run equilibrium—thereby reconciling our results. Our model predicts that the long-run growth rates of per capita real GDP and TFP are driven by the growth rates of technological innovation and aggregate employment—although, in the case of emerging economies, only technological innovation significantly contributes to TFP growth.

Keywords: endogenous technical change; scale effects; panel integration and cointegration; emerging economies; economic growth

JEL classifications: O3; O4; O14; O33; O47

1. Introduction

The scale effects, predicted by the first generation of the R&D-based endogenous growth models (Romer, 1990b; Aghion and Howitt, 1992), posit a proportional relationship of growth rates of knowledge (technology) and real per capita GDP to the scale of research and development (R&D) activity. However, the lack of empirical support for scale effects across developed countries is established as a stylized fact. Jones (1995b) eloquently summarizes it by stating that the increasing trend of either R&D labor or real R&D expenditure bears no relation to TFP (total factor productivity) growth.[1]

The theoretical and empirical conundrum of scale effects led theorists to develop two classes of second-generation R&D-based endogenous growth models: (i) semi-endogenous growth models (e.g., Jones, 1995a; Kortum, 1997; Segerstrom, 1998), and (ii) fully-endogenous "Schumpeterian" growth models (henceforth Schumpeterian; e.g., Young, 1998; Dinopoulos and Thompson, 1998; Howitt, 1999). In semi-endogenous growth models, scale effects are primarily supply-side effects, while in fully-endogenous growth models, scale effects can arise from both the supply side and the demand side. These two classes of growth models eliminate scale effects through different mechanisms. Semi-endogenous models assume decreasing returns to knowledge stock, which weakens scale effects as the economy advances and accumulates knowledge. Fully endogenous growth models maintain constant returns to aggregate R&D, but they offset scale effects through two main mechanisms. First, R&D at the firm level faces diminishing returns, meaning that adding

more researchers to a single firm does not lead to a proportional increase in technological progress, thereby weakening the direct link between population growth and innovation. In other words, each firm undertakes its own R&D instead of contributing to a shared knowledge pool, which would otherwise scale proportionally with population growth. Second, as population grows, the number of firms and product varieties expand, maintaining a competitive market where demand per firm does not systematically increase with population size. As more firms enter, competition increases, preventing any one firm from enjoying higher per-product demand or excess profitability, which, in turn, removes the incentive for population-driven R&D growth. Following these theoretical advances—namely, the advent of second-generation growth models—empirical scrutiny of scale effects has taken a back seat. The empirical literature has mainly concentrated on developed economies, comparing the two types of second-generation R&D-based growth models by testing either technology production functions or the models' predictions.[2]

We aim to contribute to the literature both empirically and theoretically. Empirically, we obtain valid estimates of scale parameters by appropriately measuring R&D scale and employing a suitable regression specification and estimator that address issues of unbalanced regression and potential spurious parameter estimates. Our novel empirical approach uncovers significant R&D scale effects in emerging countries, but finds them absent (i.e., insignificant) in OECD countries. Theoretically, in light of the markedly different results between the developed and emerging country panels, we propose an endogenous growth model that incorporates R&D labor and capital in a distinct manner. We then analyze the dynamics of scale effects, offering a theoretical framework that reconciles our empirical findings.

To provide context, most existing empirical assessments of scale effects use either labor input (Z) in the R&D sector or total R&D expenditure (R) as proxies for the scale of R&D. These studies typically compare the time-series properties of these proxies with the growth rates of technology (TFP) and/or per capita real GDP (e.g., Jones, 1995a), analyzing the trends between variables measured in levels versus those measured in growth rates. Variables measured in growth rates are unequivocally stationary, $I(0)$, while those measured in levels are non-stationary, $I(1)$, (see Section 2). As a result, they exhibit very different data patterns (trends), which led Jones (1995a) to conclude that scale effects are "counterfactual." Although these data patterns are insightful, they still leave room for valid estimation and testing of scale parameters.

The scale of R&D activities encompasses two key components: the labor employed (Z) and the real capital expenditure incurred (E) in the R&D sector. A focus on Z alone, as the scale variable, suffers from the problem of omitting a relevant variable (measure of R&D scale), namely, E, and vice versa. The use of R as the scale measure captures both R&D labor and capital expenditures as a single aggregate measure, however, the downsides of employing the aggregate measure are (i) it does not allow for the potentially different effects (roles) of Z and E—as two distinct components of the scale of R&D—on the growth rates of per capita real GDP and/or technology (TFP), and (ii) any attempt to estimate the scale parameters in a bivariate setting—by employing either R or Z as the scale measure—is likely to suffer from the problem of unbalanced regression (non-standard distribution) and spurious parameter estimates. This is because the theory of scale effects associates $I(0)$ dependent variables measured in growth rates with $I(1)$ covariates measured in levels (scales of R&D), which gives rise to the problem of an unbalanced regression (relevant tests in Section 2). To our knowledge, this issue has not been formally addressed while testing the scale effects. We address this issue by incorporating both Z and E as covariates in estimating the scale effects. Our trivariate approach not only captures the scale of R&D appropriately and distinctly but also resolves the issue of unbalanced regression and provides valid estimates of scale parameters, so long as the two scale variables (covariates) are mutually cointegrated, and this is what we find. Thus, our estimates of the scale effects are based on a more realistic measure of the scales of R&D than has been utilized hitherto, and on an estimation strategy that addresses the issue of non-standard distribution.

We conduct separate but parallel estimates of the scale effects across developed (DE) and emerging (EME) country panels. The DE panel includes 19 OECD countries from 1960 to 2016, while the EME panel comprises 26 emerging economies from 1988 to 2016. Over the last three decades or so, EME countries have made significant headway in their R&D activities. Both the scales of R&D and of patenting activities have gone up significantly across EME countries, and, as is well known, they have also outperformed DE countries in terms of growth rates (see Luintel and Khan, 2017). Nonetheless, there is little doubt that EME countries are in growth transitions and are yet to mature. In this context, a parallel scrutiny of scale effects across DE and EME countries would be interesting from the perspectives of both the first- and the second-generation growth models. This is because if DE countries operate close to their long-run equilibrium, as is widely concurred, and if scale effects are indeed the phenomena associated with growth transitions, then one would expect to find evidence of scale effects across EME countries but not across DE countries. This is exactly what we find, which underpins both the first- and the second-generation of growth models. Our results from the DE panel are consistent with the core findings in the literature that scale effects are missing in these economies as they operate close to their long-run equilibrium. Likewise, the significance of scale effects across EME economies is also consistent with the view that these countries are in growth transitions.

Theoretically, we propose a semi-endogenous growth model that explains why scale effects are present during economic transition but disappear as economies approach their long-run equilibrium. To the best of our knowledge, no formal analysis has examined the dynamics of scale effects in an endogenous growth framework that treats R&D labor and R&D capital as distinct inputs in the production function, each serving different roles in the innovation process. Our model draws from Acemoglu (1998) and is augmented by Jones' (1995a) technology production function.[3]

The frontier innovation that underpins scientific progress and contributes to the enhancement of final goods production is driven by R&D activity within a perfectly competitive research sector. This sector employs scientists and engineers—i.e., R&D labor—who initiate the discovery of new innovations. When a research firm innovates, it patents its discovery and becomes a monopolist of the new technology. Then, the monopolist translates/improves the new technology into a tangible form by embedding it in a device for a profit using R&D capital.[4] In our theoretical model, we treat R&D labor as the primary input driving breakthroughs in fundamental knowledge, which aligns with much of the existing literature (e.g., Romer, 1990b; Jones, 1995b; Ha and Howitt, 2007; Luintel and Khan, 2009). We model R&D capital as the key input for transforming basic innovations (knowledge) into tangible, device-embedded technologies—an approach consistent with recent findings by Growiec et al. (2023), who emphasize the critical role of R&D capital in fostering innovation and potentially reversing the decline in U.S. "ideas TFP".[5] Our theoretical construct seeks to capture the subtle differences between basic innovations and their subsequent progression into applied and experimental innovations, without explicitly modeling this transition. Research firms without the patent can conduct research in enhancing the technology and sell the outcome of their research to the firm possessing the patent.[6]

We show that scale effects are a function of an economy's position along its path to long-run equilibrium, and that they are non-monotonically related to the shares of new technologies, which are proportional to the growth rate of technology. We characterize developed economies as those having relatively small shares of new technologies due to their large accumulated knowledge stocks (the denominator of the share). Emerging economies, on the other hand, are characterized as having relatively large shares of new technologies due to their small accumulated knowledge stocks. Small shares of new technologies across DE countries imply small scale effects, whereas large shares across EME countries imply large scale effects, unless the latter are at the very early stage of development with little or no accumulated R&D capital stock. This characteristic is a fundamental aspect of semi-endogenous growth models, in which economies experience diminishing marginal returns to innovation capacity. As a result, scale effects diminish as the level of accumulated technology rises, aligning with the notion that innovation capacity becomes progressively less sensitive to additional technological accumulation over time.

Our model further shows that when the economy converges to its balanced growth path (BGP), the rate of economic growth is driven by the growth rates of aggregate employment and technological innovations. EME countries, being at a further distance from their BGP relative to DE economies, take longer to converge. We empirically evaluate the model's long-run growth predictions by approximating technological innovations using the flow of patent filings—a widely used measure in the literature—and find that the results are largely consistent with the model's predictions.

The rest of the paper is organized as follows. Section 2 presents empirical estimates of scale effects, Section 3 presents the endogenous growth model, and Section 4 tests the long-run predictions of the model. Section 5 concludes.

2. Estimates of scale effects

In this section, we discuss our sample and data, lay out our econometric model and estimation method, and report the parallel results of scale effects obtained from DE and EME country panels. We have an unbalanced panel of DE countries, which consists of 19 OECD countries with data on almost all relevant variables spanning the period 1965–2016. The exceptions are (i) the data on TFP which cover the period of 1965–2014, and (ii) the employment data for Austria and Denmark which span 1969–2016. The DE panel has a minimum of 950 to a maximum of 988 country years of data points spanning at least 50 years. Our EME country panel is also unbalanced, consisting of 26 countries with data points covering a minimum of 609 to a maximum of 719 country years, spanning a maximum of 29 (1988–2016) to a minimum of 23 (1994–2016) years. Of the 26 sample EME countries, three (Belarus, Cuba, and Pakistan) do not have data on TFP. Table A1 of Appendix A lists data sources and sample countries.

In principle, using adjusted measures of TFP such as those developed by Basu et al. (2006) would be preferable, as they better isolate technological progress by accounting for cyclical variations in factor utilization. However, constructing a reliable adjusted TFP series across the broad panel of emerging market economies in our sample is almost infeasible due to data limitations—in particular, the unavailability of consistent information on capital services, labor effort, and capacity utilization. Given these constraints, we rely on the TFP growth measure from the Penn World Table (PWT), primarily due to its broad coverage, cross-country consistency, and widespread use in empirical growth research.[7]

Summary statistics for some of the key R&D variables for both country panels are reported in Table A2 of Appendix A. There is considerable heterogeneity in the growth rates of real per capita income, domestic patent filings, R&D and research intensities, and research productivity both within and across these country panels. Our panel of EME countries shows an average annual growth rate of 2.69% during the sample period. China records the highest average annual growth rate of 8.19% and the Russian Federation the lowest (0.006%). Likewise, the average annual growth rate is 2.05% across the DE panel: Ireland records the highest (3.80%) and Switzerland the lowest (1.16%). The average number of annual domestic patent filings is 26,594 in the DE panel which is 2.6 folds higher than in the EME panel (10,232). China is dominant in patent filings across the EME countries, and the USA is dominant across the DE countries. However, Chinese average annual domestic patent filings are 61% higher than those of the USA. Our empirical results, reported below, are robust to the exclusions of big and/or small countries from the panel. The sample average R&D intensity is 0.75% across EME countries, which is much lower than that of DE countries (1.69%). Singapore shows the highest R&D intensity of 1.90% across the EMEs, and Sweden across the DEs (2.54%). The average research intensity across EME countries is about two-thirds of the DE level. The average research productivity in DE countries is a lot higher (over 4 folds) than in EME countries. The USA shows quite a low proportion of R&D capital expenditure relative to its total R&D expenditure, which may reflect its mature R&D sector.

The theory of scale effects posits a proportional relationship of growth rates of productivity (*TFP*) and real per capita GDP (*x*) to the level (scale) of R&D activity. Specifically, we estimate the average cross-country-time scale effects across EME and DE countries, as measured by the average cross-country and time semi-elasticities $\varepsilon_{M,Z} = \frac{1}{nT} \sum_i^n \sum_t^T \varepsilon_{M,Zi,t}$ and $\varepsilon_{M,E} = \frac{1}{nT} \sum_i^n \sum_t^T \varepsilon_{M,E_{i,t}}$ for $M \in \{x \,; TFP\}$. To this end, we specify an auxiliary regression of the following form:

$$g_{M,i,t} = \rho_i + \gamma_t + \varepsilon_{M,Z} \ln Z_{i,t} + \varepsilon_{M,E} \ln E_{i,t}$$

$$+ \sum_{j=-2}^{2} \eta_j \Delta \ln Z_{i,t-j} + \sum_{j=-2}^{2} \mu_j \Delta \ln E_{i,t-j} + e_{i,t}, \tag{1}$$

where $g_{M,i,t}$ is the growth rate of *M*, for country *i* at time *t*. Equation (1) is a fixed effects linear-log model in the Dynamic OLS (DOLS, Stock and Watson, 1993) framework, where ρ_i captures the country-specific fixed effects and γ_t captures the time effects. The scale effects relationship is between the dependent variable, $g_{M,i,t}$, measured in growth rates, and the covariates, $Z_{i,t}$ and $E_{i,t}$, measured in log levels. Panel unit root tests confirm that growth rates of per capita real GDP, $g_{x,i,t}$, and total factor productivity, $g_{TFP,i,t}$, are $I(0)$, while scale variables, $lnZ_{i,t}$ and $lnE_{i,t}$, are $I(1)$.[8] This very presence of stationary $I(0)$ dependent variable and non-stationary $I(1)$ covariates may give rise to the problem of unbalanced regression while testing the scale effects.

Any estimation of scale effects through bivariate regressions which employs a single proxy of the scale of R&D—whether *lnZ* or *lnE* or *lnR* as they all are $I(1)$—suffers from the problem of an unbalanced regression and non-standard distribution because the regression residuals will be non-stationary. However, our specification is a trivariate one, and so long as the two $I(1)$ covariates, *lnZ* and *lnE*, are mutually cointegrated, they provide a sensible specification for the $I(0)$ dependent variable by making the estimating equation balanced.[9] Our specification, therefore, has two clear advantages: (i) it provides valid estimates of scale parameters as *lnZ* and *lnE* are cointegrated, and (ii) the inclusion of both *lnZ* and *lnE* captures the scale of R&D activities distinctly and more accurately. If the parameters of $lnZ_{i,t}$ and $lnE_{i,t}$ are both positive and significant, this would support the presence of scale effects in both measures of R&D scale. Alternatively, a positive and significant coefficient on either $lnZ_{i,t}$ or $lnE_{i,t}$ alone would also indicate evidence of an R&D scale effect in the corresponding scale measure.

The DOLS is a powerful and efficient estimator of a cointegrating relationship when the regression model contains a mixture of stationary and non-stationary variables. This approach augments the estimating equation by the suitably differenced leads and lags of non-stationary regressors, which eliminate endogeneity. Stock and Watson (1993) allow for covariates with different orders of integration—e.g., $I(0)$, $I(1)$, and $I(2)$—in the regression equation; however, they always maintain the dependent variable as $I(1)$. Our dependent variable is stationary, therefore, our trivariate specification, which incorporates two mutually cointegrated regressors, is important for a valid estimation and inference of scale effects.

Table 1 reports the estimates of scale parameters for the EME and DE country panels. The first two columns of results pertain to the full EME panel of 26 countries. Results from the full EME panel show significant scale effects under both measures of R&D scale, as both *lnZ* and *lnE* appear positive and significant in explaining $g_{x,i,t}$, and $g_{TFP,i,t}$. The estimated scale parameter of *lnZ* is significant at 1% in explaining $g_{x,i,t}$ while the scale parameter of $lnE_{i,t}$ is significant at 10% or better. Likewise, *lnE* appears significant at 1% in explaining $g_{TFP,i,t}$ but *lnZ* is significant at 10% or better. The Levin et al. (2002) t-tests (t_{llc}) reject the non-stationarity of the error correction term at a very high level of precision, confirming the estimated scale effect relationships for the EME panel are indeed cointegrated.

China and India are two major emerging countries, with China notably leading in R&D and patenting within the emerging world. To assess the robustness of our findings on significant scale effects in the EME panel, we exclude China and India from the sample and re-estimate the relationship. Results, reported in the middle two columns of Table 1, show a significant scale effect

Table 1. Panel DOLS estimates of scale effects

	EME (1988–2016)		EME excl. China & India		OECD (1965–2016)	
Regressors	Dependent variables					
	g_x	g_{TFP}	g_x	g_{TFP}	g_x	g_{TFP}
ln Z	0.045	0.024	0.014	0.012	0.013	−0.002
	(0.018)	(0.014)	(0.019)	(0.016)	(0.009)	(0.007)
	{0.012}	{0.095}	{0.473}	{0.597}	{0.139}	{0.827}
ln E	0.016	0.027	0.020	0.024	−0.0004	−0.004
	(0.009)	(0.011)	(0.009)	(0.011)	(0.015)	(0.003)
	{0.089}	{0.011}	{0.046}	{0.038}	{0.939}	{0.297}
	Panel cointegration tests					
t_{llc}	{0.000}	{0.000}	{0.00}	{0.00}	N/A	N/A
N [OBS]	26 [501]	23 [446]	24 [461]	21 [406]	19 [890]	19 [893]

Notes: Numbers in parentheses are standard errors, and those within curly braces are *p*-values of Wald tests under the null that the estimated coefficient is zero, which are $\chi^2(1)$. t_{llc} {*p*-value reported} are the Levin, Lin, and Chu (ibid.) t-test of the null of unit root in the panel error correction term (i.e., the null of non-cointegration of the estimated relationships). Since all parameter estimates of OECD panels are statistically insignificant, it makes no sense to conduct cointegration tests; hence, "N/A." The second-order leads and lags are used for augmentations, and constant and linear trends are maintained as individual deterministic components. Belarus, Cuba, and Pakistan do not have data on g_{TFP}. N [OBS] denotes the number of countries [data points] used in each estimation after accounting for the leads and the lags. Numbers beyond three decimal places are reported as 3.5e-4 = 0.00035.

in *lnE* while the coefficient of *lnZ* is positively signed but imprecisely estimated. These results are consistent with the findings of Growiec et al. (2023), albeit from a slightly different perspective, indicating that R&D capital demonstrates greater robustness than R&D labor in explaining productivity growth.[10] Although the exclusion of China and India represents a substantial change to the EME sample, the results continue to exhibit significant scale effects, underscoring the robustness of our findings.

Likewise, to control for the potential influence of international knowledge spillovers on our scale effects estimates, we incorporate international spillover pools into the estimating equation. These are sourced from Luintel and Khan (2017), who construct them using bilateral capital (machinery) import shares from 20 OECD countries as weights, and provide detailed documentation on their construction. Spillover data are available for 23 of the 26 emerging economies in our sample (excluding Cuba, Hong Kong, and South Africa) for the period 1988–2013. Consistent with the lagged nature of knowledge diffusion, a two-year lag is applied, following Mansfield (1985) and Caballero and Jaffe (1993). Despite a modest reduction in country coverage and data span, incorporating spillover pools continues to yield strong evidence in support of scale effects. The coefficient on ln *E* remains highly significant and positive for both real per capita output and TFP growth. Notably, the inclusion of spillovers reinforces the results obtained by excluding China and India, as the qualitative nature of the results remain while the magnitude of the parameter of ln *E* increases noticeably.[11] In sharp contrast, as evidenced by the results in the last two columns of Table 1, the corresponding estimates of scale effects for the DE panel are entirely insignificant.[12]

This difference in scale effects results between the EME and DE country panels is consistent with the view that scale effects are unlikely to exist amongst mature economies that are on or close to their BGP but may exist when economies are progressing through growth transitions, as may be the case for the emerging countries. Our findings vis-à-vis the industrialized countries are consistent with those of Jones (1995a, b), despite our different empirical approach.[13]

Interestingly, the estimated scale parameters show large scale effects of R&D on economic and productivity growth rates of emerging countries. Dependent variables are measured as proportions (i.e., 5% as 0.05) and covariates are measured in log levels, hence the reported parameters

are semi-elasticities. To provide some perspective on the magnitudes of the scale effects, using the estimated scale parameters for the full EME sample, a 1% increase in Chinese R&D labor would lead to a 0.55% increase in the growth rate of per capita real GDP in China (point elasticity).[14] Likewise, a 1% increase in real R&D capital expenditure would increase the Chinese growth rate of per capita real GDP by 0.20%. Chinese average annual productivity growth has been 1.91% during the sample period. The scale effect parameter estimates imply point elasticity of Chinese TFP of above unity with respect to both Z and E. Likewise, other emerging countries appear to benefit by increasing the scale of their R&D activities: countries experiencing lower growth rates are set to benefit more by expanding their R&D sectors.

There is also a useful analogy to Schumpeterian models, which imply positive relationships between R&D intensity, technological change, and the growth rate of per capita output. Specifically, these models suggest that higher R&D intensity—here measured by increases in the quantities of inputs used in the R&D sector—leads to improved technologies and higher economic growth. Although the theoretical model we develop in the following section is semi-endogenous—in the sense that new ideas or innovations are exogenous—it nevertheless incorporates Schumpeterian elements, as these innovations take shape and become usable through R&D efforts.

3. Endogenous growth model

Our empirical results from the emerging countries panel show significant scale effects of R&D, using both R&D labor and R&D capital as measures of scale. In contrast, our results across developed countries suggest the absence of scale effects for both measures of scale. Existing growth models typically assume that R&D labor is the sole input into the R&D process (e.g., Romer, 1990b; Jones, 1995a; Ha and Howitt, 2007). However, we argue that this assumption may be overly simplistic, as our results highlight the importance of R&D capital expenditure alongside labor. We propose a new semi-endogenous growth model, building on Acemoglu's (1998) framework and incorporating both R&D labor and R&D capital as key drivers of technological progress in analyzing the dynamics of scale effects. The structure of our model is similar to that of Acemoglu (1998), while the evolution of technology is modeled along the lines of Jones (1995a, b). Analyzing both R&D inputs—R&D labor and R&D capital— we offer a clearer understanding of how technological advancement affects growth and productivity in both developed and emerging economies. Specifically, our model reconciles the significant scale effects across emerging country panels and the insignificant scale effects across developed country panels. There is a continuum of infinitely-lived individuals, with identical intertemporally additive preferences defined over consumption. The marginal utility of consumption is assumed to be constant, which implies that the rate of time preference $r > 0$ is also the interest rate.

3.1. Production of final goods

Aggregate output, Y, is produced by perfectly competitive firms, defined on the unit interval such that $Y = \int_0^1 y(i)di$, where $y(i)$ denotes the output produced by firm i. The price of the final output is the numeraire. Output for firm i is produced using neutral technology $A(i)$, labor $n(i)$, and, general capital $k(i)$, such that $y(i) = A(i)n(i)^\beta k(i)^{1-\beta}$, where $0 < \beta < 1$. The general capital is the physical capital owned by consumers who rent it out to firms. The aggregate supply of workers and general capital are given by $N \equiv \int_0^1 n(i)di$ and $K \equiv \int_0^1 k(i)di$, respectively. The evolution of neutral technology is driven by R&D-induced intangible technology, Q, which takes on a tangible form through the use of R&D capital that enables the technology to be used in the production process. Put differently, Q acquires material form once it is embedded into a device using firm-specific R&D capital, $e(i)$. The function that maps Q into tangible technology devices is $F(i) = Qe(i)^\lambda$, where $0 < \lambda < 1$. The firm that utilizes technology Q must incur the firm-invariant cost of R&D capital, denoted by χ, per unit of R&D capital. It follows that the change of firm-specific neutral

technology $A(i)$ is given by $\dot{A}(i) = \lambda^{-(1-\phi_A)} A(i)^{\phi_A} F(i)$, where $0 < \phi_A < 1$ and $\dot{A}(i)$ is the derivative of $A(i)$ with respect to time. For the sake of notational simplicity, we omit time as an argument unless it is necessary. The profit function for firm i is, $\pi(i) = A(i)n(i)^\beta k(i)^{1-\beta} - \chi e(i) - wn(i) - r_K k(i)$, where w and r_K are the wage rate and the rental price for general capital, respectively, while the level of neutral technology is given by:

$$A(i) = \frac{1}{\lambda} \left[(1 - \phi_A) \int_0^t Q(\tau)e(i, \tau)^\lambda d\tau \right]^{\frac{1}{1-\phi_A}}. \tag{2}$$

The firm chooses quantities of $n(i)$, $k(i)$ and $e(i)$ in order to maximize profit. Firms are identical, therefore, in equilibrium, they end up making the same choices, hence the optimality conditions reduce to: $w = \beta A(K/N)^{1-\beta}$; $\chi = \lambda^{\phi_A} A^{\phi_A} Q E^{-(1-\lambda)} N^\beta K^{1-\beta}$; and, $r_K = (1 - \beta)A(K/N)^{-\beta}$; where $N \equiv n$, $E \equiv \int_0^1 e(i) = e(i)$, and $Y \equiv y(i)$. Due to the risk-neutrality of consumers, who are also the owners of physical capital, $r_K = r + \delta$, where δ is the depreciation rate of general capital.[15]

3.2. Research & development sector

We assume a research sector with free entry that is populated by perfectly competitive firms. A research firm j contributes $q(j)$ to the development of technology $Q = \int_0^1 q(j)dj$ by carrying out R&D using researchers, $z(j)$.[16] The profit from R&D activity is given by $\pi^D(j) = \rho(j)V(j) - Bq(j)z(j)$, where $\rho(j) \equiv \varphi(Q, Z)Cz(j)$ is the flow rate of innovation with $\varphi(Q, Z) = v^{-(1-\phi_Q)}Q^{\phi_Q}Z^{v-1}$, $Z = \int_0^1 z(j)dj$, C is a productivity factor defined below, $0 < \phi_Q < 1$, $0 < v < 1$, $V(j)$ is the value of innovation, and $Bq(j)$ is the firm's cost per researcher with $B > 0$. Research firms take $\varphi(Q, Z)$ as given, i.e., they perceive themselves to be too small to affect the aggregate invention probability.[17] The productivity factor is given by $C(t) = \sigma \exp(g_C t)$, where σ is a scale factor that coincides with the Poisson rate at which new ideas arrive at a research firm, and $g_C > 0$ is the sectoral growth rate of productivity due to new ideas.[18] In other words, while new ideas at the sectoral level grow over time, at every instant a new idea is randomly allocated to a single research firm and although access to C is not exclusive to the allocated firm, the latter acquires a monopoly right over the particular vintage by receiving a patent. Once the technology is upgraded, the patent holding research firm is the only firm that can sell the upgraded technology to the final goods firms, charging a profit maximizing rent for the R&D capital needed to translate technology into tangible form.

While the patent prevents the rest of the research firms from accessing the market of final goods, they can sell their research output to the firm that possesses the patent for $\rho(i)V(i)$. The firm that owns the patent has an incentive to purchase the research output of other firms in order to motivate them to work on improving technology further since the latter improves its own productivity via $\varphi(Q, Z)$. The ownership value of the leading vintage of technology input is given by:

$$rV(j) = \pi^m(j) - \int_0^1 \rho(i)V(i)di + \dot{V}(j), \tag{3}$$

where $\pi^m(j)$ is the instantaneous profit of the monopolist that owns the leading vintage and $\dot{V}$ is the derivative of V with respect to time, which captures changes in the valuation of the leading vintage. The profit function $\pi^m(j)$ is written as $\pi^m = \chi(E)E - QE$; where $\chi(E)$ is the inverse demand for R&D capital derived from the problem of the firm producing final output. As in Acemoglu (1998), the profit-maximizing price, χ, turns out to be a constant mark-up over marginal cost, that is, $\chi = Q/\lambda$.[19]

3.3. The balanced growth path and transition dynamics

In this section, we characterize the Balanced Growth Path (BGP) and the transition dynamics to it. Using the production function of final output, the growth rate of per capita output, $x = Y/N$, can be written as $g_x = g_A + (1 - \beta)g_{\bar{K}}$, where $g_{\bar{K}}$ denotes the growth rate of per capita general capital. The optimal condition for general capital implies that $g_{\bar{K}} = (1/\beta)g_A$, which means that $g_x = (1/\beta)g_A$. Since C is common across all research firms, in equilibrium, $z(j) \equiv Z$, $V(j) = V$, $q(j) \equiv Q$ and thus $\rho(j) = \rho$, while the optimal condition for a research firm becomes $\varphi(Q, Z)CV = BQ$ for all j. It follows that Q evolves according to $\dot{Q} \equiv \rho$ which implies that,

$$Q = \frac{1}{\nu} \left[(1 - \phi_Q) \int_0^t C(\tau)Z(\tau)^\nu d\tau \right]^{\frac{1}{1-\phi_Q}}. \tag{4}$$

Along the BGP all variables grow at a constant rate, i.e., $\dot{g}_A = \dot{g}_Q = \dot{g}_x = \dot{g}_N = \dot{g}_E = \dot{g}_Z = \dot{g}_V = 0$. As shown in Online Appendix B, at the BGP the growth rates of all endogenous variables are driven by the growth rates of the exogenous arrival of new ideas, g_C, and aggregate employment, g_N, that is,

$$g_J = \gamma_{J,C}g_C + \gamma_{J,N}g_N, \tag{5}$$

for $J = Z, Q, E, A, x$, where $\gamma_{J,C}$ and $\gamma_{J,N}$ are functions of structural parameters.[20] For $g_C > 0$ and $g_N > 0$, the existence of a BGP requires that,

$$0 < \nu < \bar{\nu} \equiv \frac{(1 - \phi_Q)[\beta(1 - \phi_A) - \lambda]}{1 - \beta(1 - \phi_A)},$$

with the necessary condition $\lambda < \beta(1 - \phi_A)$. If parameter values do not satisfy these inequalities, the non-negativity conditions for Q, A, E as well as g_C and g_N, along the BGP, are violated.[21] Thus, our model's solution shows that long-run growth is driven not only by labor growth, as a typical semi-endogenous model would suggest, but also by the rate of new discoveries (or the flow of new ideas), which enable innovators to earn monopoly rents through patents. In other words, our model includes an additional exogenous growth component—the flow of new ideas—that exists beyond the growth rate of labor.

Figure 1 displays the transition dynamics towards the BGP, which are summarized by two lines in the (g_A, g_Q) space: a vertical line for g_Q, as per the BGP equation (B3) (Online Appendix B), and an upward sloping line for g_A, as per the equation (B2) (Online Appendix B). The economy tends to converge to the unique BGP where the two lines intersect and the growth rates of A and Q are driven only by g_C and g_N.[22]

To examine the dynamics of scale effects, we consider the semi-elasticities: $\partial g_J / \partial \ln E$ and $\partial g_J / \partial \ln Z$ for $J = Q, A, x$. These semi-elasticities, denoted by $\varepsilon_{J,E}$ and $\varepsilon_{J,Z}$, can be written as:

$$\varepsilon_{A,E}(t) = \frac{\lambda \theta_A(t)\,[1 - \theta_A(t)]}{1 - \phi_A}, \quad \varepsilon_{A,Z}(t) = \frac{\nu \theta_Q(t)\theta_A(t)\,[1 - \theta_A(t)]}{(1 - \phi_A)(1 - \phi_Q)}, \quad \varepsilon_{Q,Z}(t) = \frac{\nu \theta_Q(t)\,[1 - \theta_Q(t)]}{1 - \phi_Q}$$

$$\varepsilon_{x,E}(t) = \frac{\varepsilon_{A,E}(t)}{\beta}, \quad \varepsilon_{x,Z}(t) = \frac{\varepsilon_{A,Z}(t)}{\beta},$$

where $\theta_A(t)$ and $\theta_Q(t)$ are the following technology shares:

$$\theta_A(t) = \frac{Q(t)E(t)^\lambda}{\int_0^t Q(\tau)E(\tau)^\lambda d\tau} \equiv (1 - \phi_A)g_A(t), \text{ and } \theta_Q(t) = \frac{C(t)Z(t)^\nu}{\int_0^t C(\tau)Z(\tau)^\nu d\tau} \equiv (1 - \phi_Q)g_Q(t).$$

As shown above, the growth rates of technology are proportional to the technology share parameters, θ_Q and θ_A, which reflect the economy's state of development. In this model, scale effects are clearly supply-side in nature, as decreasing returns reduce the impact of R&D capital and labor on the growth rates of technology and per capita output. We argue that emerging
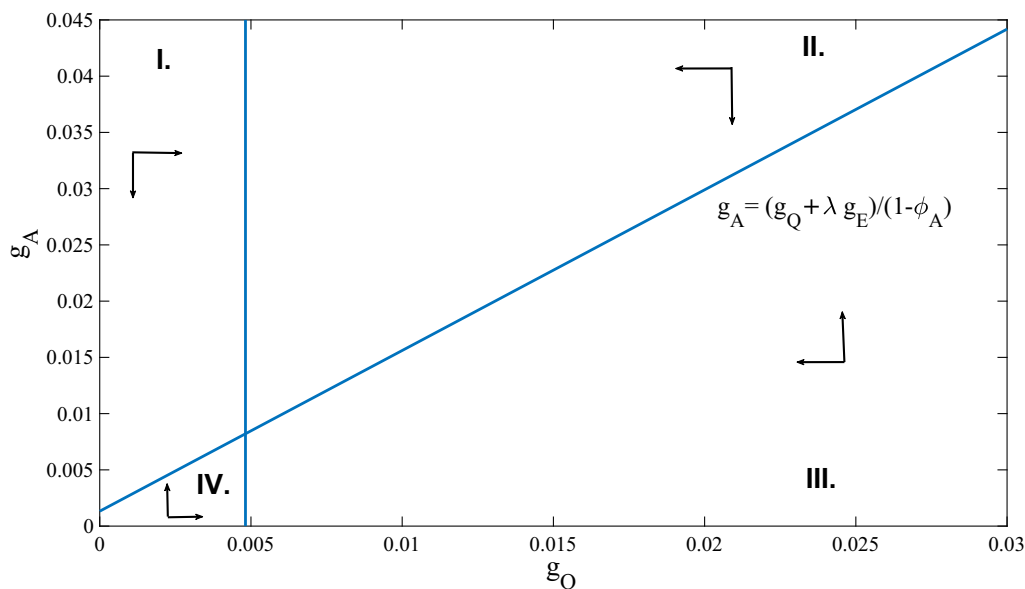
Figure 1. Transition dynamics toward the BGP.

economies, which are in the early stages of development, exhibit high technology shares, while developed economies, which operate close to their BGP, exhibit low technology shares. This is because emerging (developed) economies have small (large) accumulated stocks of technology, therefore incremental technology forms a large (small) share of their accumulated stocks. Since g_A and g_Q are proportional to technology shares, the technological growth rates of emerging economies are relatively high while those of the developed economies are relatively low. It follows that emerging economies at the initial stages of development exhibit high growth rates of technology which lie at the top corners of Figure 1, denoted by areas I and II or at a point along the 45° line. However, when EMEs gradually develop by accumulating technology, their rates of growth of technology slow down as they converge towards their BGP values at the intersection of the two lines of Figure 1. This is a classical feature of semi-endogenous growth models, where economies are characterized by diminishing marginal returns to innovation capacity. This implies that scale effects are weakened as the level of accumulated technology increases, consistent with the theory that innovation capacity becomes less responsive to further technological accumulation over time.

Figure 2 is a visual display that relates the transition dynamics of scale effects with the technology shares. Specifically, it highlights the BGP semi-elasticities versus the technology shares obtained from a calibration exercise where $\lambda = \phi_A = \phi_Q = \nu = 0.1$, $\beta = 0.75$ and $g_C = g_N = 1\%$. The implied BGP scale effect from this calibration is $\varepsilon_{x,E} = 0.34\%$, as measured by R&D capital, and $\varepsilon_{x,Z} = 3.9079e - 05$, as measured by R&D labor.[23] To make the BGP values of θ_Q and θ_A visually distinct we denote them by θ_Q^* and θ_A^*.

As is evident from Figure 2, economies that are either at the very early stages of development where technology growth rates are high (on the lower right corners of Figure 2a–c and on the lower right corners of Figure 2d and e) or operating close to their BGP, where technology growth rates are low (on the lower left corners of Figure 2a–c and on the lower front corners of Figure 2d and e), exhibit small scale effects, measured both by R&D capital and researchers. That is, technology growth rates at the very early stages of economic development are large since the accumulated levels of technology are small, and so logarithmic increments of Z and E have negligible effects on the former. On the other hand, for economies operating very close to their BGP, scale effects
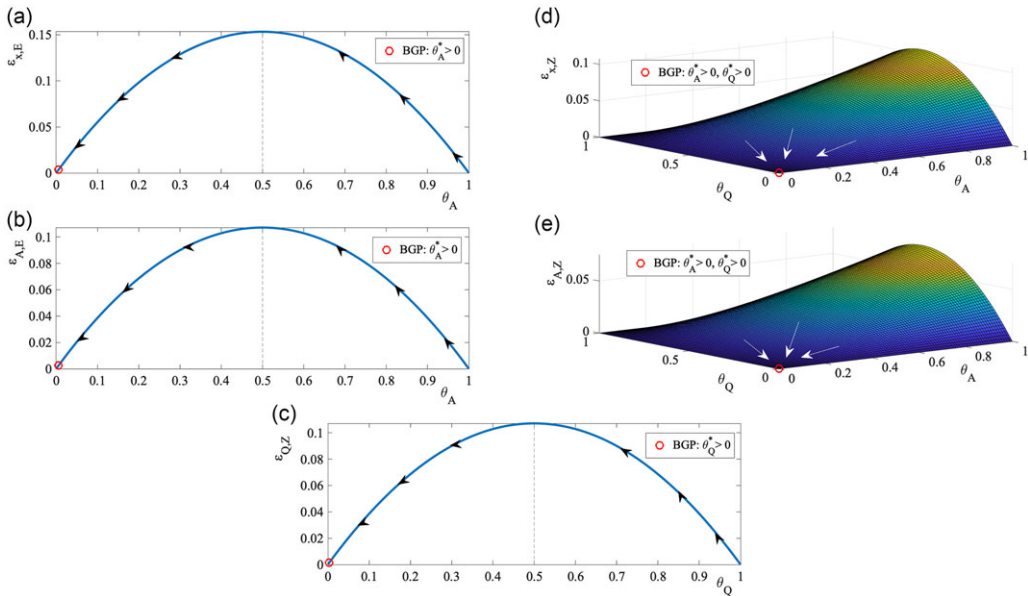
Figure 2. Transition of scale effects toward the BGP.

cease to exist due to decreasing returns to technology. Intuitively, economies which are about to converge to their BGP have large accumulated stocks of knowledge, hence any new incremental knowledge induced from either Z or E exerts a trivial effect on overall knowledge stocks.[24]

Thus, our model shows that once the emerging countries pass through their initial stages of development and begin their transition towards their long-run equilibrium, they initially experience amplified scale effects. As they approach closer and closer to their BGPs, the scale effects gradually subside. Hence, scale effects are seen during growth transitions but not at the BGP or at its vicinity, which reconciles our empirical results of significant scale effects across EME countries but their insignificance across DE countries, unless EME countries are at the very early stages of their development.

4. Testing the model's balanced growth path

A typical semi-endogenous growth model implies that along the BGP, economic growth is primarily driven by the growth rate of research effort (Jones, 2021). Our model predicts that, along the BGP, economic growth is driven by both the growth rate of labor and the growth rate of new discoveries, as shown by equation (5). The latter implies cointegrating relationships between lnJ_t, lnC_t, and lnN_t, where $J = Z, E, A, x$, with the cointegrating vector of $(1, -\gamma_{J,C}, -\gamma_{J,N})$, which can be shown as:

$$lnJ_t = \gamma_{J,C}lnC_t + \gamma_{J,N}lnN_t, \tag{6}$$

where A is approximated with *TFP*.[25] A direct way to evaluate the model's prediction along the BGP is by testing the cointegrating relationships in equation (6).

In our model, there is a distinction between new ideas and the process of transforming these ideas or discoveries into intangible technologies, Q, which are later used in production by translating them into tangible form through R&D capital. New ideas are exogenous and correspond to variable C in our model. These ideas require R&D labor to refine and develop them. Since C is directly unobservable, we consider the flow of domestic patent filings of sample countries to be the closest available proxy—though not a perfect one—for these new ideas. Patents are a widely used

measure of new-to-the-world ideas, which also serve as legal protections for these new ideas.[26] Note that in our model, whenever a new idea is developed, it increases C, and is patented by the inventor. Hence, there is a natural link between patent flows and the variable C. The key notion is that new discoveries, as ideas, are exogenous, while their transformation into usable technology is an endogenous process. This is consistent with the fact that firms often patent discoveries before committing to large-scale R&D investment, meaning patents serve as a snapshot of the exogenous innovation landscape, distinct from the endogenous process of converting ideas into productive technology.

For a robust inference on the BGP relationships, we employ two estimators of cointegrating relationships, namely, DOLS and the FMOLS (Fully Modified OLS; Phillips and Hansen, 1990, and Phillips and Moon, 1999), supplemented by the Levin, Lin, and Chu (ibid.) t-test (t_{llc}), on the estimated error correction term. Unlike the scale effects specifications in equation (1), which include variables in both logarithmic first differences and levels, equation (6) is specified entirely in log levels, which are $I(1)$, hence the use of FMOLS is valid.[27]

The cointegration estimates, which proxy the BGP relationships, as predicted by our model, are reported in Table 2. Panel A reports the results for the EME panel, and Panel B for the DE panel. It is evident that the levels of per capita real GDP, R&D employment, and R&D capital expenditure are cointegrated with the flow of new ideas (innovations) and the level of total employment across emerging countries. Their cointegrating parameters are positive and significant at very high levels of precision (1% or better), and the t_{llc} tests reject the null of non-cointegration (i.e., the non-stationarity of the error correction term). However, TFP only appears cointegrated with the flow of innovations, as the cointegrating parameter of lnN appears statistically insignificant. The reported t_{llc} test for TFP only captures the significant parameret of lnC. These results are robust across both estimators: DOLS and FMOLS. They imply that, in the long run, growth rates of per capita real GDP (g_x), R&D employment (g_Z), and R&D capital expenditure (g_E) are driven by both the growth rates of exogenous technology (g_C) and total employment (g_N) across emerging countries. These results are consistent with the long-run predictions of our model. However, the growth rate of TFP is driven by (g_C) alone, a purely Schumpeterian outcome.

Although our theory suggests that both $\ln C_t$ and $\ln N_t$ should drive *TFP*, this does not seem to hold in our estimates for the EME panel. This discrepancy may reflect that emerging economies are not operating near their balanced growth paths (BGP) and may exhibit significant deviations from optimal resource allocation, including in labor markets. This may be due to several factors: emerging economies often face capital constraints, which limit the extent to which increased employment translates into technological progress; a large share of employment occurs in the informal sector, outside regulated labor markets; and institutional rigidities and policy inefficiencies may delay adjustments, prolonging the time to equilibrium and dampening the positive impact of employment on TFP. In contrast, results of Panel B show significant cointegrating relationships of the levels of per capita real GDP, total factor productivity, R&D employment, and R&D capital expenditure with the flow of innovations and the level of total employment across developed countries. All estimated cointegrating parameters are positive and highly significant, and the t_{llc} test unequivocally rejects the null of non-cointegration across all estimates. It can be argued that developed economies are more likely to operate near the technological frontier, where changes in employment are closely associated with the adoption of new technologies. Given that all cointegrating vectors—except for the one describing the long-run dynamics of TFP in emerging markets—are consistent with theoretical expectations suggests that, overall, our model accurately captures the underlying long-run relationships. Specifically, the results of Table 2 suggest that a 1% increase in the flow of innovation, C, induces an increase of per capita real GDP by 0.09% (0.21%) for emerging countries in the long run, and by 0.20% (0.19%) for OECD countries based on the DOLS (FMOLS) estimates. Likewise, a 1% increase in the aggregate employment, N, induces an increase of per capita real GDP by 0.78% (0.62%) for emerging countries and 1.14% (1.32%) for OECD countries.

Table 2. Estimates of long-run (BGP) relationships

	Panel A: 26 emerging countries							
	DOLS				FMOLS			
Regressors	Dependent variables							
	lnGDP	*lnZ*	*lnE*	*lnTFP*	*lnGDP*	*lnZ*	*lnE*	*lnTFP*
ln *C*	0.092	0.303	0.443	0.023	0.211	0.308	0.483	0.066
	(0.011)	(0.022)	(0.028)	(0.011)	(0.022)	(0.033)	(0.036)	(0.012)
	{0.000}	{0.000}	{0.000}	{0.042}	{0.000}	{0.000}	{0.000]}	{0.000}
ln *N*	0.784	2.164	2.042	−0.011	0.618	2.108	1.986	−0.094
	(0.034)	(0.145)	(0.163)	(0.039)	(0.121)	(0.189)	(0.212)	(0.072)
	{0.000}	{0.000}	{0.000}	{0.775}	{0.000}	{0.000}	{0.000}	{0.188}
	Panel cointegration tests							
t_{llc}	{0.000}	{0.046}	{0.023}	{0.001}	{0.002}	{0.047}	{0.049}	{0.000}
N [OBS]	26 [584]	26 [575]	26 [561]	23 [522]	26 [679]	26 [659]	26 [632]	23 [568]
	Panel B: 19 OECD countries							
	lnGDP	*lnZ*	*lnE*	*lnTFP*	*lnGDP*	*lnZ*	*lnE*	*lnTFP*
ln *C*	0.200	0.289	0.273	0.076	0.194	0.462	0.313	0.052
	(0.019)	(0.055)	(0.046)	(0.011)	(0.027)	(0.065)	(0.068)	(0.018)
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.004}
ln *N*	1.144	2.887	2.377	0.383	1.324	2.826	2.443	0.462
	(0.064)	(0.134)	(0.117)	(0.033)	(0.084)	(0.196)	(0.208)	(0.054)
	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
	Panel cointegration tests							
t_{llcf}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
N [OBS]	19 [883]	19 [885]	19 [885]	19 [885]	19 [957]	19 [961]	19 [961]	19 [923]

Variable mnemonics are: *lnGDP* = real GDP per capita; *lnZ* = scientists and engineers employed in the R&D sector; *lnE* = capital expenditure in the R&D sector; *lnTFP* = total factor productivity; *lnC* = flow of exogenous technological innovations proxied by the patent filings of sample countries; and, *lnN* = total employment. All variables are measured in natural logarithms. Numbers in parentheses are standard errors and those within curly braces are *p*-values of Wald tests under the null that the estimated coefficient is zero, which are $\chi^2(1)$. Country fixed effects are maintained in all estimations. Belarus, Cuba, and Pakistan do not have data on *TFP*. *N* [OBS] denotes the number countries [data points] of each estimation. t_{llc} denotes the Levin, Lin and Chu (ibid.) test of the null of non-cointegration (i.e., the non-stationarity of the error correction tem), *p*-values reported.

5. Conclusion

This paper provides new theoretical and empirical insights on the long-debated issue of R&D scale effects on the growth rates of technology and output. By conducting separate but parallel estimates of scale effects across both the developed and the emerging country panels, we report significant scale effects across emerging countries, and their complete insignificance across developed countries. Importantly, our results are based on a more realistic measure of the scale of R&D activities than has been applied hitherto, as well as on an empirical method that addresses the issues of unbalanced and spurious regressions. Specifically, R&D activities are captured by the joint use of R&D labor and R&D capital expenditure of each sample country, which gives the true scale of R&D. We elucidate that the theoretical prediction of scale effects implies a test equation that is statistically unbalanced, relating stationary regressands to non-stationary covariates and hence the potential problem of spurious regressions. We apply an appropriate estimator which addresses these estimation issues.

To analyze the dynamics of scale effects as a function of a country's position on its transition path to a long run equilibrium, we propose an extension of Acemoglu's (1998) endogenous growth model with a production technology in the lines of Jones (1995a, b). We show that during the course of transition to long-run equilibrium, scale effects of both R&D labor and R&D capital tend to be large and prominent, but as a country approaches its long-run equilibrium, scale effects deplete.

Our model predicts, among other things, that the long-run growth rates are driven by both the rate of growth of technological innovations and the rate of growth of aggregate employment. Empirical scrutiny of the long-run implications of our model reveals that all long-run relationships implied by the model are validated, except for the long-run relationship between TFP and labor in emerging economies. We argue that the latter is likely to be driven by transitional factors and structural rigidities, as these economies operate farther from their BGPs. To sum up, this paper shows, both empirically and theoretically, that scale effects are present in emerging economies undergoing growth transitions, but not in mature economies that operate near their long-run equilibrium.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1365100525100801.

Competing interests. The author declares none.

Notes

1 A robust conclusion of the literature, as noted by Young (1998), is that the measurement of the effect of the scale on growth requires production factors, which are used intensively in the innovation process. The reason the scale of the economy is directly linked to innovation activity is the non-rivalrous nature of the latter, which allows technologies to be widespread in the economy (see Romer, 1990a, b), as opposed to changes in other (general) factors of production which mainly affect certain sectors.

2 E.g., see among others, Zachariadis (2003), Laincz and Peretto (2006), Ha and Howitt (2007).

3 Our augmented technology production function differs from the one used by Ha and Howitt (2007) for comparing and testing second generation growth models.

4 Nadiri and Prucha (1996) provide evidence that R&D capital differs from other physical capital by estimating significantly different depreciation rates for the two types of capital using U.S. data. This is consistent with the assumptions in our paper as general capital may have a non-zero depreciation rate while R&D capital fully depreciates.

5 Rivera-Batiz and Romer (1991) and Bloom et al. (2020) embrace the inclusion of 'lab equipment' in their R&D output specifications.

6 For instance, consider a technological discovery which facilitates the fast and effective processing of speech recognition and another technological discovery that identifies whether a Wi-Fi network is shared with neighbors. All sector research firms have access to these technologies and can make further improvements to them, i.e., to maximize their effectiveness/quality. However, only the firm that initially discovered the technology and was granted a patent, can build and sell a device that incorporates the technology (e.g., to internet service providers in the case of the Wi-Fi identifier). All the other research firms can only sell their output to the firm that owns the patent. Note that while the number of Wi-Fi identifier devices produced enable the firm to control accessibility, and thus increase the speed of connection and operational efficiency, without the technology embedded in each device, the devices on their own are of no use.

7 While we acknowledge that the PWT measure is derived from a production function that differs from the one used in our theoretical model of the following section, we consider it a reasonable proxy for long-run technological progress. This is especially the case in our context, where the emphasis is on long-term trends rather than short-run fluctuations, which are likely to average out over time.

8 We implement a Fisher-type Phillip-Perron (PP) panel unit root test (Maddala and Wu, 1999), by setting a fixed lag length of 3, that allows for heterogenous unit root processes across panel units. The tests confirm that $lnZ_{i,t}$ and $lnE_{i,t}$ are $I(1)$, whereas, $g_{x,i,t}$, and $g_{TFP,i,t}$, are $I(0)$. Specifically, under the null of unit root, the p-values for $lnZ_{i,t}$ and $lnE_{i,t}$, are 0.590 and 0.903 respectively, across the EME panel and 0.823 and 0.267 across the DE panel. Likewise, the respective p-values for $g_{x,i,t}$, and $g_{TFP,i,t}$, are 0.000 for both panels.

9 Kao's Engle-Granger residual based panel test of cointegration decisively rejects the null of non-cointegration between lnZ and lnE across both panels. Under the null of no cointegration, the test statistics are -2.70 (0.004) and -1.625 (0.052), respectively for the EME and the DE panels; figures within parentheses are p-values. We allow for individual intercept and

set a fixed lag length of 3 while calculating these test statistics; allowing for the automatic lag length selection improves the precision of both test statics (*p*-values of 0.000).

10 Whereas Growiec et al. (2023) estimate the Ideas Production Function and "Idea TFP" using R&D labor and R&D capital stock, we focus on modeling scale effects using R&D labor and real R&D capital expenditure as measures of R&D scale.

11 The estimated equations inclusive of spillover pools are:

$$g_x = 0.041 \{0.110\} \ln Z_{i,t} + 0.132 \{0.002\} \ln E_{i,t} + 0.065 \{0.000\} \ln S_{i,t-2}^{fm},$$

$$g_{\text{TFP}} = -0.048 \{0.155\} \ln Z_{i,t} + 0.088 \{0.021\} \ln E_{i,t} + 0.030 \{0.241\} \ln S_{i,t-2}^{fm},$$

where $S_{i,t-2}^{fm}$ denotes the lagged foreign spillover pool weighted by bilateral machinery imports from OECD countries, and the values in curly braces {.} represent *p*-values from the Wald test under the null hypothesis that the corresponding parameter equals zero. We thank an anonymous reviewer for suggesting the inclusion of the international knowledge spillover pool in our estimating equation.

12 Truncating the DE sample to 1988–2016 and making it exactly match the sample period of the EME panel also shows no support for the scale effects across the DE panel. Furthermore, the significant scale effects observed for the EME panel remain robust when China and India are excluded individually. Specifically, excluding China from the EME panel maintains the significant scale effects of lnE on $g_{x,i,t}$ and $g_{TFP,i,t}$ while the parameter estimates of lnZ turns imprecise (insignificant at 10%). The exclusion of India does not alter the significant scale effects on $g_{x,i,t}$ but the parameter estimates of lnZ turns imprecise for $g_{TFP,i,t}$.

13 One of the anonymous referees noted that some advanced countries may not have remained on their balanced growth paths (BGPs) over the entire period and suggested measuring their distances from the BGP to assess the potential effect on the R&D coefficient. We acknowledge this important point. Measuring each OECD country's time-varying scale effect requires time-varying parameter estimation within a time-series framework. However, our focus is on estimating average scale effects across countries and time, which our approach adequately captures. Although countries may move on and off their BGPs at different times (short run effects), we believe the average magnitude of the scale parameters is unlikely to differ markedly from our estimates. Moreover, even if advanced economies are found to occasionally deviate from their BGPs, calculating their distances from the BGP would merely complement the scale-effect estimates for OECD panel reported in the last two columns of Table 1. It would neither affect (i) our robust scale-effect results for the EME panel nor (ii) our model and its predictions that characterize the dynamics of scale effects in the transition toward the steady state. The advantage of estimating an average scale effect is that it downplays short-run deviations of individual countries within the pool of countries. Nevertheless, we consider the referee's suggestion an insightful and intend to explore it in future research.

14 Note that $\frac{\partial g_{x,i,t}}{\partial lnZ_{i,t}} = \frac{\partial g_{x,i,t}}{\partial Z_{i,t}/Z_{i,t}} = \varepsilon_{x,Z}$, which can be rewritten as $\frac{\partial g_{x,i,t}}{\partial Z_{i,t}} = \frac{\varepsilon_{x,Z}}{Z_{i,t}}$. Using the latter, the elasticity of $g_{x,i,t}$ with respect to $Z_{i,t}$ is written as $\frac{\partial g_{x,i,t}}{\partial Z_{i,t}} \frac{Z_{i,t}}{g_{x,i,t}} = \frac{\varepsilon_{x,Z}}{Z_{i,t}} \frac{Z_{i,t}}{g_{x,i,t}} = \frac{\varepsilon_{x,Z}}{g_{x,i,t}}$. Then, using the cross-time average growth rate for country i, the elasticity for country i is written as $\frac{\varepsilon_{x,Z}}{\bar{g}_{x,i}}$. Thus, the growth rate elasticity with respect to R&D labor for China is $0.045/0.082 \approx 0.55$.

15 This condition further implies that in equilibrium general capital per capita depends on the level of technology that is, $K/N = (1-\beta)A^{1/\beta}/(r+\delta)$.

16 Researchers may also be included in aggregate labor N and receive a premium for their research work, in addition to general salary, w. None of the results will be affected by this assumption.

17 This assumption is equivalent to Acemoglu's (1998)assumption that small firms ignore their impact on the invention probability of other firms working to improve the same machine.

18 Without compromising our findings, we have assumed that the growth rate of productivity $C(t)$ is time invariant. Alternatively, we could have assumed that $C(t) = \sigma e^{\int_0^t g_C(s)ds}$, where $g_C(t)$ captures the fact that the magnitude of the impact of each new idea on productivity varies over time. Whether g_C is time invariant or not however does not have any effect on our results. Hence, we simplify the model by assuming that g_C is constant.

19 The fact that $\rho(i)V(i)$ is subtracted from the right hand side of (3) is also consistent with Acemoglu's (1998) assumption that this term captures the rate that firm j loses its monopoly position.

20 The solutions for $\gamma_{J,C}$ and $\gamma_{J,N}$ by way of structural parameters are provided in Online Appendix B.

21 If $\nu > \bar{\nu}$ then, $\gamma_{A,C} < 0$ and $\gamma_{A,N} < 0$ and as a consequence $g_A < 0$, which is infeasible since $Q > 0$, $A > 0$ and $E > 0$ imply that $g_A = \lambda^{-(1-\phi_A)}A^{-(1-\phi_A)}QE^\lambda > 0$. If $\nu = \bar{\nu}$ then, $g_N = -\frac{1-\beta(1-\phi_A)}{\beta(1-\phi_A)(1-\phi_Q)}g_C$ which contradicts the fact that both g_C and g_N are strictly positive. If $\nu < \bar{\nu}$ then, $\gamma_{A,C} > 0$ and $\gamma_{A,N} > 0$ and thus $g_A > 0$.

22 Differentiating the growth rate of A with respect to time, using equation (2), allows us to obtain an equation that describes the evolution of g_A that is, $\dot{g}_A/g_A = g_Q + \lambda g_E - (1-\phi_A)g_A$. It follows that when $g_A < (g_Q + \lambda g_E)/(1-\phi_A)$ then $\dot{g}_A/g_A > 0$, whereas when $g_A > (g_Q + \lambda g_E)/(1-\phi_A)$ then $\dot{g}_A/g_A < 0$. Likewise, differentiating the growth rate of Q with respect to time, using equation (4), allows us to obtain an equation that describes the evolution of g_Q that is, $\dot{g}_Q/g_Q = g_C + \nu g_Z - (1-\phi_Q)g_Q$. It follows that when $g_Q < (g_C + \nu g_Z)/(1-\phi_Q)$ then $\dot{g}_Q/g_Q > 0$, whereas when $g_Q > (g_C + \nu g_Z)/(1-\phi_Q)$ then $\dot{g}_Q/g_Q < 0$. Therefore, g_A and g_Q always tend to converge towards their BGP, which explains the arrows of Figure 1.

23 This calibration is chosen for expositional purposes only. It further implies that $\varepsilon_{A,E} = 0.18\%$, $\varepsilon_{A,Z} = 2.9309e - 05$, $\varepsilon_{Q,E} = 0.25\%$, and $\varepsilon_{Q,Z} = 0.18\%$.

24 In terms of the mathematical expressions of the elasticities, at early stages of economic development since θ_Q and θ_A are close to unity, $\theta_Q(1 - \theta_Q)$ and $\theta_A(1 - \theta_A)$ are close to zero. Likewise, for economies operating in the neighborhood of their BGP, $\theta_Q(1 - \theta_Q)$ and $\theta_A(1 - \theta_A)$ are close to zero because θ_Q and θ_A are close to zero.

25 To show that (5) is reduced to (6), let $g_{J,t} \in G_J(g_J)$, $g_{C,t} \in G_C(g_C)$, $g_{N,t} \in G_N(g_N)$ where G_J, G_C and G_N are the neighborhoods of the corresponding BGP's. Since the logarithmic first differences approximate growth rates, the BGP relationships of equation (5) can be expressed in terms of the variables in log levels as $lnJ_t - lnJ_{t-1} = \gamma_{J,C}[lnC_t - lnC_{t-1}] + \gamma_{J,N}[lnN_t - lnN_{t-1}]$, which can be re-arranged as $lnJ_t - \gamma_{J,C}lnC_t - \gamma_{J,N}lnN_t = lnJ_{t-1} - \gamma_{J,C}lnC_{t-1} - \gamma_{J,N}lnN_{t-1}$. At the BGP, the latter can be extended further to $lnJ_t - \gamma_{J,C}lnC_t - \gamma_{J,N}lnN_t = lnJ_{t-1} - \gamma_{J,C}lnC_{t-1} - \gamma_{J,N}lnN_{t-1} = lnJ_{t+1} - \gamma_{J,C}lnC_{t+1} - \gamma_{J,N}lnN_{t+1} = lnJ_{t+2} - \gamma_{J,C}lnC_{t+2} - \gamma_{J,N}lnN_{t+2} = \ldots = lnJ_{t+n} - \gamma_{J,C}lnC_{t+n} - \gamma_{J,N}lnN_{t+n} = \ldots$.

26 We recognize that patents, despite their wide usage, are a noisy measure of innovations, as they greatly differ in their "universality" and "size" (Eaton et al., 1998), as well as in values (Battke et al., 2016). Nevertheless, patent flows are the only consistent proxy of innovations that are available and in wide usage.

27 The Fisher-type PP panel unit root test (Maddala and Wu (ibid.)) confirm that $lnGDP$, lnZ, lnE, $lnTFP$, and lnN are all I(1) across both panels but the null of unit root is rejected for lnC at the conventional significance levels of 10% or better. However, t_{llc} test decisively shows it to be $I(1)$; hence, on balance, we treat lnC as I(1). The latter result is also widely reported in the literature.

References

Acemoglu, D. (1998). Why do new technologies complement skills? Directed technical change and wage inequality. *Quarterly Journal of Economics* 113(4), 1055–1089.

Aghion, P. and P. Howitt. (1992). A model of growth through creative destruction. *Econometrica* 60(2), 323–351.

Basu, S., J.G. Fernald and M.S. Kimball. (2006). Are technology improvements contractionary? *American Economic Review* 96(5), 1418–1448.

Battke, B., T.S. Schmidt, S. Stollenwerk and V.H. Hoffmann. (2016). Internal or external spillovers—Which kind of knowledge is more likely to flow within or across technologies? *Research Policy* 45, 27–41.

Bloom, N., C.I. Jones, J. Van Reenen and M. Webb. (2020). Are ideas getting harder to find? *American Economic Review* 110(4), 1104–1144.

Caballero, R.J. and A.B. Jaffe. (1993). How high are the giants' shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. In Caballero, R.J. and A.B. Jaffe. (eds.), *NBER Macroeconomics Annual*. Cambridge, MA: MIT Press, pp. 15–86.

Dinopoulos, E. and P. Thompson. (1998). Schumpeterian growth without scale effects. *Journal of Economic Growth* 3(4), 313–335.

Eaton, J., E. Gutierrez and S. Kortum. (1998). European technology policy. *Economic Policy* 13, 403–438.

Growiec, J., P. McAdam and J. Mućk. (2023). R&D capital and the idea production function. Federal Reserve Bank of Kansas City Working Paper No. 23-05.

Ha, J. and P. Howitt. (2007). Accounting for trends in productivity and R&D: A schumpeterian critique of semi-endogenous growth theory. *Journal of Money, Credit and Banking* 39(4), 733–774.

Howitt, P. (1999). Steady endogenous growth with population and R&D inputs growing. *Journal of Political Economy* 107, 715–730.

Jones, C.I. (1995a). R&D-based models of economic growth. *Journal of Political Economy* 103(4), 759–784.

Jones, C.I. (1995b). Time series tests of endogenous growth models. *Quarterly Journal of Economics* 110(2), 495–525.

Jones, C.I. (2021). The past and future of economic growth: A semi-endogenous perspective. *Annual Review of Economics* 14, 125–152.

Kortum, S. (1997). Research, patenting, and technological change. *Econometrica* 65, 1389–1419.

Laincz, C.A. and P.F. Peretto. (2006). Scale effects in endogenous growth theory: An error of aggregation, not specification. *Journal of Economic Growth* 11, 263–288.

Levin, A., C.-F. Lin and C.-S. Chu. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics* 108, 1–24.

Luintel, K.B. and M. Khan. (2009). Heterogeneous ideas production and endogenous growth: An empirical investigation. *Canadian Journal of Economics* 42(3), 1176–1205.

Luintel, K.B. and M. Khan. (2017). Ideas production and international knowledge spillovers: Digging deeper into emerging countries. *Research Policy* 46(10), 1738–1754.

Maddala, G.S. and S. Wu. (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics (Special Issue)* 61(S1), 631–652.

Mansfield, E. (1985). How rapidly does new industrial technology leak out? *Journal of Industrial Economics* 34, 217–223.

Nadiri, M.I. and I.R. Prucha. (1996). Estimation of the depreciation rate of physical and R&D capital in the U.S. total manufacturing sector. *Economic Inquiry* 34(1), 43–56.

Phillips, P.C.B. and B.E. Hansen. (1990). Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57, 99–125.

Phillips, P.C.B. and H.R. Moon. (1999). Linear regression limit theory for non-stationary panel data. *Econometrica* 67, 1057–1111.

Rivera-Batiz, L.A. and P.M. Romer. (1991). Economic integration and endogenous growth. *Quarterly Journal of Economics* 106(2), 531–555.

Romer, P.M. (1990a). Are nonconvexities important for understanding growth? *American Economic Review Papers and Proceedings* 80, 97–103.

Romer, P.M. (1990b). Endogenous technological change. *Journal of Political Economy* 98(5), 71–102.

Segerstrom, P.S. (1998). Endogenous growth without scale effects. *American Economic Review* 88(5), 1290–1310.

Stock, J.H. and M.W. Watson. (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61, 783–820.

Young, A. (1998). Growth without scale effects. *Journal of Political Economy* 106(1), 41–63.

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.

Appendix A. Data sources and descriptive statistics

Table A1. Data sources and sample countries

Resident patent applications at national patent office—OECD and EME countries: Word Intellectual Property Organization (WIPO), Statistics Database: https://www3.wipo.int/ipstats/ https://www.wipo.int/ipstats/en/#resources

Employment at the R&D sector (Research scientists and engineers employed at the R&D Sector)—OECD countries: Main Science and Technology Indicators (MSTI database): https://www.oecd.org/sti/msti.htm EME countries: UNESCO—http://data.uis.unesco.org/

Total R&D expenditure—OECD countries: Total R&D expenditure on research and development in constant 2010 PPP$ in million. Source OECD MSTI. MSTI: https://www.oecd.org/sti/msti.htm EME countries: Total R&D expenditure on research and development (GERD series) in national currency, UNESCO—http://data.uis.unesco.org/ R&D expenditure in 2011 PPP$—our own calculation by using GDP deflator and PPP series from the World Bank

R&D capital expenditure—OECD countries: Total R&D capital expenditure in constant PPP$ (2010) in million. Gross domestic expenditure on R&D by sector of performance and type of expenditure (OECD R&D database): https://www.oecd.org/sti/inno/researchand developmentstatisticsrds.htm EME countries: UNESCO (http://data.uis.unesco.org/) by "type of cost series". R&D Capital expenditure in 2011 PPP$, million, is our own calculation by using GDP deflator and PPP data series from the World Bank

Total factor productivity (TFP)—OECD and EME countries: Penn world table version 9. TFP index 2011 = 1. Series code rtfpna. https://www.rug.nl/ggdc/productivity/pwt/pwt-releases/pwt9.0

GDP and GDP deflator—OECD countries: Gross domestic products in constant 2010 PPP$, million. From World Bank WDI dataset for all countries except Switzerland, for the latter is from OECD Economic Outlook 105. World Bank WDI—https://databank.worldbank.org/source/world-development-indicators OECD Economic Outlook: OECD (2019), "OECD Economic Outlook No. 105 (Edition 2019/1)", OECD Economic Outlook: Statistics and Projections (database), https://doi.org/10.1787/b8fe9e35-en

GDP deflator series are also from the same sources as GDP except for Canada, New Zealand and Switzerland. Deflator series for these three countries are obtained from OECD Economic Outlook No. 105 (Edition 2019/1). EME countries: GDP in 2011 PPP$, millions. Our own calculation based on: GDP in national currency (series code: NY.GDP.MKTP.CN); GDP deflator (series code: NY.GDP.DEFL.ZS); PPP$ (Series code PA.NUS.PPP) World Bank WDI—https://databank.worldbank.org/source/world-development-indicators

Purchasing power parity, national currency per USD—OECD and EME countries: World Bank WDI—https://databank.worldbank.org/source/world-development-indicators

Employment—OECD countries: Total employment in million from OECD Economic Outlook Nos. 79 and 105. OECD Economic Outlook: OECD (2019), "OECD Economic Outlook No. 105 (Edition 2019/1)", OECD Economic Outlook: Statistics and Projections (database), https://doi.org/10.1787/b8fe9e35-en EME countries: World Bank data series, "Employment to population ratio, 15+, total (%)" (series code:SL.EMP.TOTL.SP.ZS). We employ total population data to convert ratios to absolute numbers. World Bank WDI—https://databank.worldbank.org/source/world-development-indicators

Table A1. Continued.

Purchasing power parity, national currency per USD—OECD and EME countries: World Bank WDI—https://databank.worldbank.org/source/world-development-indicators
Employment—OECD countries: Total employment in million from OECD Economic Outlook Nos. 79 and 105. OECD Economic Outlook: OECD (2019), "OECD Economic Outlook No. 105 (Edition 2019/1)", OECD Economic Outlook: Statistics and Projections (database), https://doi.org/10.1787/b8fe9e35-en EME countries: World Bank data series, "Employment to population ratio, 15+, total (%)" (series code:SL.EMP.TOTL.SP.ZS). We employ total population data to convert ratios to absolute numbers. World Bank WDI—https://databank.worldbank.org/source/world-development-indicators
Population: – OECD and EME countries: World Bank WDI—https://databank.worldbank.org/source/world-development-indicators
Sample Countries:
Twenty-six countries of emerging (EME) Panel: Argentina, Belarus, Brazil, China, Colombia, Croatia, Cuba, Czechoslovakia, Hong Kong, Hungary, India, Latvia, Malaysia, Mexico, Morocco, Pakistan, Poland, Russian Federation, Singapore, Slovakia, Slovenia, South Africa, Thailand, Tunisia, Turkey, and Uruguay
Nineteen countries of developed (DE or OECD) Panel: Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and United States

Notes: Data on GDP, GDP per capita, total R&D expenditure, and R&D capital expenditure across sample countries are converted into constant (2010) PPP US dollars. TFP is measured as index: 2011 = 1 for emerging countries and 2010 = 1 for developed countries at the source. R&D labor is the full-time equivalent of research scientists and engineers employed in the R&D sector.
Total employment is the labor force employed (in millions). All data accessed in March 2019.

Table A2. Descriptive statistics (sample mean)*

	g_x	Patents	R&D exp.	R&D int.	Z	Res. int.	Res. prod.	R&D–KEXP%
				EME panel				
Max	8.192^{E1}	208.556^{E1}	129505.5^{E1}	1.898^{E4}	976.787^{E1}	1.037^{E4}	16.672^{E1}	21.98^{E7}
Min	0.006^{E2}	0.031^{E3}	154.204^{E3}	0.207^{E5}	1.193^{E3}	0.034^{E5}	0.420^{E6}	5.050^{E2}
Mean	2.694	10.232	10517.04	0.750	83.185	0.348	3.982	14.113
Obs.	693	714	636	627	676	676	674	636
				OECD panel				
Max	3.798^{1}	129.229^{3}	248977.00^{3}	2.541^{5}	823.183^{3}	0.868^{6}	47.657^{7}	17.923^{4}
Min	1.161^{2}	0.223^{4}	1440.287^{4}	0.609^{4}	7.310^{1}	0.296^{4}	2.963^{4}	0.221^{3}
Mean	2.054	26.594	28607.35	1.688	117.912	0.558	17.662	10.518
Obs.	963	988	988	982	988	980	988	988

*Reported sample means are calculated over the available data length for each country and each variable. g_x is the average annual growth rate of real per capita GDP (%); Patents refer to annual national filings in "000" R&D Exp. (expenditures) are in millions 2010 PPP$ (2011 PPP$) for OECD (EME) countries; R&D int. (intensity) refers to total R&D expenditure as percentage of GDP; Z refers to researchers, scientists and engineers in R&D Sector in "000" Research int. (intensity) refers R&D researchers as percentage of the total employment; Research prod. (productivity) refers to resident patent applications per 100 researchers; and $R\&D - KEXP$% refers to R&D capital expenditure as a percentage of total R&D expenditure. Superscripts: $E1$ = China; $E2$ = Russain Federation; $E3$ = Uruguay; $E4$ = Singapore; $E5$ = Colombia; $E6$ = Pakistan; $E7$ = Poland; refer to the countries resuming either the maximum or the minimum values of the corresponding measures in the emerging country panel. Likewise, subscripts: 1 = Ireland; 2 = Switzerland; 3 = USA; 4 = Portugal; 5 = Sweden; 6 = Finland; 7 = Japan; denote countries in the developed country panel.