CEIS Tor Vergata Research Papers

Università di Roma Tor Vergata

ISSN 2610-931X

CEIS Tor Vergata

RESEARCH PAPER SERIES

Vol. 18, Issue 2, No. 483 – May 2020

Complexity and Growth

Alberto Bucci, Lorenzo Carbonari, Pedro Mazeda Gil and Giovanni Trovato

Complexity and Growth

Alberto Bucci[*] Lorenzo Carbonari[†] Pedro Mazeda Gil[‡]
Giovanni Trovato[§]

Abstract

Over the past decades, research effort in high income countries has substantially increased. Meanwhile, the growth rates of per capita *output have been rather stable. The first goal of this paper is to investigate the reasons for such trends. The second goal of the paper is to show that the occurrence of different phases in the economic growth dynamics traces back to the interplay between complexity and specialization in production. To do this we use data from a sample of OECD countries and estimate a Hidden Markov Model, through which we identify four distinct growth regimes.*

Keywords: Economic growth; Population growth; Complexity; Hidden Markov Model.

JEL codes: O3; O4; J1.

[*]Università di Milano (DEMM and FinGro Lab), CEIS Tor Vergata, and RCEA (Rimini Center for Economic Analysis). E-mail address: alberto.bucci@unimi.it.

[†]Università di Roma "Tor Vergata", DEF and CEIS. Corresponding author. E-mail address: lorenzo.carbonari@uniroma2.it.

[‡]Universidade do Porto and CEF.UP. E-mail address: pgil@fep.up.pt.

[§]Università di Roma "Tor Vergata", DEF and CEIS. E-mail address: giovanni.trovato@uniroma2.it.

"Rich countries specialize in complicated products".

Kremer (1993), p. 563.

1 Introduction

Figures 1 and 2 provide some aggregate evidence, from a group of advanced economies over the period 1983-2007, that deserves to be investigated. With the partial exception of UK, no positive correlation seems to exist between long-run *per capita* real GDP growth and innovation, captured by either research inputs (e.g. number of researchers employed per million inhabitants) or research outputs (e.g. number of patent applications by residents). Data from other OECD countries would produce similar figures. Of course, one may question the economic and statistical significance of this suggestive evidence, since aggregate innovation is a broad phenomenon, only partially captured by these two R&D related measures. Nonetheless, the increasing trends reported in Table 1, in terms of patenting activity and researchers employed, have, at least, no corresponding increase in the long-run growth rates of *per capita* GDP. Long-run income and productivity growth have rather declined over the last twenty years (apart from the modest recovery immediately before the 2007-2008 global financial crisis).

Table 1: Four OECD countries, 1983-2007

Country	5Y-AVG. Real GDP growth (% variation)	Patent appl. by residents[a] (% variation)	Researchers[b] (% variation)
Germany	+1.10	+51.45	+107.69
Japan	-2.03	+11.63	+46.70
United Kingdom	-1.10	-12.88	+83.10
United States	-1.50	+253.28	+52.96

Note: Japanese data are referred to the period 1983-2006. Source: [a] *World Bank,* [b] *(per mil. of inhab.) OECD.*

The first goal of this paper is to investigate the reasons for such diverging trends. In the spirit of Aghion and Howitt (1998), our explanation is based on the idea that innovation makes production activities more complex and this requires the development of appropriate skills and abilities to adapt to changing technological needs. More

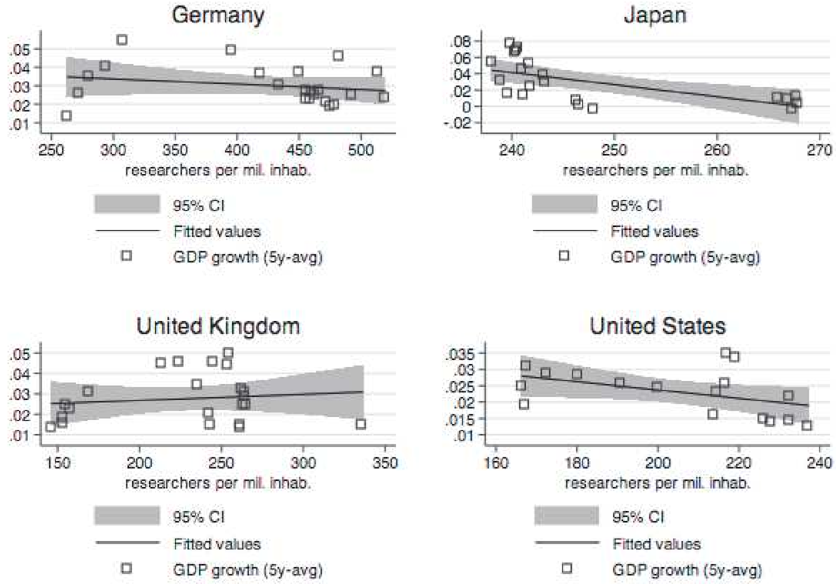

Figure 1: Real GDP growth and Research Input

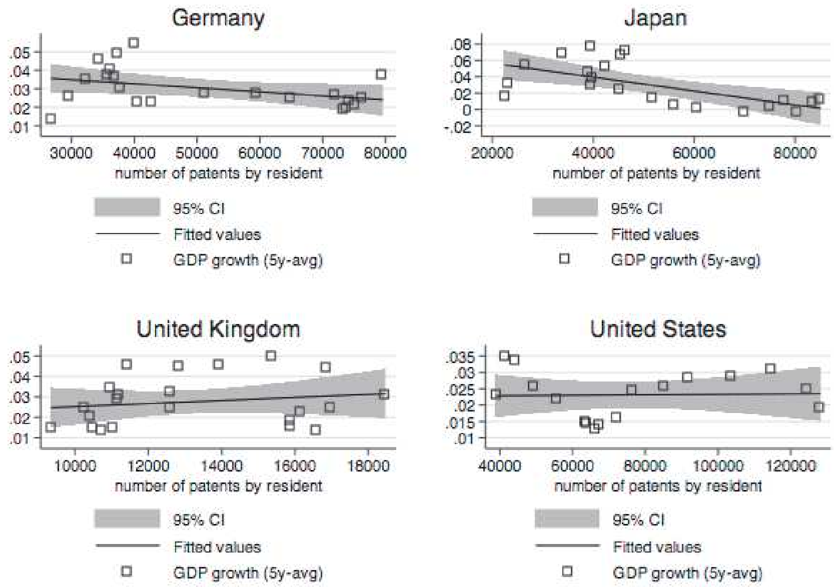

Figure 2: Real GDP growth and Research Output

formally, a larger number of intermediate-input varieties to be assembled in the same manufacturing process triggers, at the same time, an increase in the cost, due to the higher complexity, and new opportunities of gain, originating from specialization. Complementarily, we explore the idea, in the spirit of Jones (1995) and Ha and Howitt (2007), that innovation benefits from increasing research inputs but may be curtailed by complexity effects pertaining to innovation activities (under the form of duplication, difficulty and dilution effects), as the number of varieties of intermediates and the amount of research inputs rise in the economy.

The second goal of the paper is to show that the occurrence of different phases in the economic growth dynamics traces back to the interplay between complexity and specialization in production. The tension between complexity (both in production and innovation activities) and specialization relies on the mechanism through which population growth and monopolistic markups rewarding prospective innovators may simultaneously affect economic growth. To rationalize this mechanism, we propose a simple extension of the canonical *semi-endogenous* growth model (Jones, 1995). To allow for a variety of growth phases, we assume – in the econometric part of the paper – that there exist different "states of nature", in which growth behavior (captured by some key structural parameters) differs for otherwise identical countries.[1] To identify the growth regimes and the transition between them, we estimate our theoretical growth equation using a Hidden Markov Model, which allows to deal with both observed and unobserved (*hidden*) factors that affect long-run growth (Baum and Petrie, 1966; Baum and Eagon, 1967).[2] For a sample of OECD countries, we find four distinct regimes, corresponding to four growth processes. These regimes capture different

[1]In this sense, the regime-varying parameters are reduced forms of the underlying processes generating aggregate complexity. Hence, (possible) changes in these parameters may be interpreted as a change in the complexity generated by these processes.

[2]A Hidden Markov Model is specified by the following components: i) a set of states (or *regimes* or *classes*), ii) a transition probability matrix, iii) a sequence of observations, iv) a sequence of observation likelihoods and v) an initial probability distribution over states. The main advantage of this econometric technique is that it allows to make inference about an unobserved process based on the observed one. The set of the unobserved factors which may affect a country's growth path is potentially large and includes institutional setting, political uncertainty, educational system, etc.

balanced growth paths, with different long-run average growth and different growth volatility. Importantly, our results are robust to several sensitivity checks and alternative estimation techniques. In particular, our growth regimes classification survives when we employ an alternative theoretical specification in which we (partially) endogenize complexity by letting some of its determinants to be functions of the research input (i.e., researchers/population).

From the theoretical standpoint, the paper closest to ours is Bucci, Carbonari and Trovato (2019), who analyze how the degree of complexity in production may affect not only the rate of economic growth, but also the correlation between the latter, population growth and monopolistic markups. The model presented below extends their R&D technology to simultaneously deal with difficulty, duplication and dilution effects. Moreover, in the empirical part of the paper, we use the restrictions provided by the model for an econometric exercise to understand how complexity affects the transition between different phases of economic development and the membership to a specific growth regime.

Other papers have recently dealt with the link between complexity and growth. Using a radically different approach, Pintea and Thompson (2007) identify in the increasing technological complexity, experienced during the second half of 20th century, the main cause for the puzzling coexistence of secular increases in R&D expenditure and educational attainment and no corresponding increase in *per capita* income growth. Using trade data and the notion of "product space", Hausmann and Klinger (2006) develop a model offering a broad appraisal of the complexity effect. Along this line and using a sample of 89 countries over the period 1990-2009, Ferrarini and Scaramozzino (2012) support the idea that, especially for advanced countries, the productivity-gains from more specialization are smaller than the associated productivity-losses due to increased complexity in production. Unlike Ferrarini and Scaramozzino (2012), we employ a different econometric technique and analyze how the balance between complexity and specialization (induced by input proliferation) contributes to affect both the rate of economic growth and the growth regime in which a country may fall.

Finally, within the literature which reads economic growth as a sequence of transitions between distinct growth phases that countries visit with different frequencies, two contributions deserve to be mentioned, namely Jerzmanowski (2006) and Kerekes (2012). Both papers deliver purely empirical exercises: in particular, Jerzmanowski (2006) studies 89 countries over the period 1962-1994 while Kerekes (2012) studies 84 countries over the period 1961-2003. They employ an econometric technique – namely, a Markow switching model – similar to ours.[3] Differently from us, in both papers growth dynamics simply originates from an auto-regressive process and there is no room for the implications of complexity in determining the long-run pattern of the real *per capita* GDP.

The structure of the paper is the following: Section 2 develops our theoretical model, Section 3 presents the econometric analysis and Section 4 concludes.

2 The Model

We build upon the R&D-driven growth model of Romer (1990) and the extensions of this model provided, respectively, by Grossman and Helpman (1991) and Bucci, Carbonari and Trovato (2019). These are dynamic general-equilibrium endogenous growth models where a homogeneous final (consumption) good is competitively produced using labor and a continuum of varieties of intermediate inputs. The latter are, in turn, produced in a monopolistically-competitive sector using (one-to-one) solely labor. Potential entrants into the intermediate-good sector devote labor to horizontal R&D, by which they increase the measure of varieties of intermediate inputs. The economy is populated by infinitely-lived (dynastic) households who consume and inelastically supply labor to firms in the final-good, the intermediate-good, and the R&D sectors. We

[3] The Hidden Markov models are a subclass of autoregressive models with Markov regime, for which the conditional distribution of the depend variable does not depend on its lagged values but only on the regime. For a discussion on the link between Markov-switching models and Hidden Markov models, see Junag and Rabiner (1985) and Rabiner (1989).

assume that the aggregate labor force equals total population,[4] which increases at a constant positive exogenous growth rate.

2.1 Production

The final good, Y, is produced at time t using, as private and rival inputs, both labor and a continuum of intermediate goods,

$$Y_t = L_{Yt}^{1-\alpha} \left[\frac{1}{N_t^{\beta}} \int_0^{N_t} (x_{it})^{1/m} di \right]^{\alpha m}, \qquad 0 < \alpha < 1 \quad \text{and} \quad m > 1. \tag{1}$$

In the equation above, L_{Yt} is the quantity of the labor input in final-good production, x_{it} is the quantity of the *i-th* variety of differentiated intermediate inputs with $i \in [0, N_t]$, α is a parameter that controls for the the labor share in final-good production (this share is given by $1-\alpha$), and m is a parameter that controls for the elasticity of substitution between any generic pair of varieties of intermediate goods (this elasticity is given by $m/(m-1)$). Following Ethier (1982) and Benassy (1996a, 1996b, 1998), the aggregate production function (1) allows to disentangle the optimal markup on the marginal production cost in the intermediate-good sector (or, alternatively, the measure of product market concentration in that sector), m, from the factor-shares in final-good production, α and 1-α.[5] Finally, parameter β controls for the *complexity effect* on aggregate output induced by the expansion of varieties of intermediate goods, as emphasized by, e.g., Aghion and Howitt (1998). In particular when positive, β captures the detrimental effect on aggregate output of having a larger number of intermediate-input varieties to be assembled in the same manufacturing process. This

[4]Under this assumption, *per capita* and per worker variables coincide.

[5]Since final output is produced competitively under constant returns to scale to rival inputs, L_Y and x_i are rewarded according to their marginal productivities at equilibrium. Hence, for a given N, $1-\alpha$ is the share of Y going to labor and α is the share going to intermediate inputs. As we show below, a decrease in m increases the substitutability between intermediate goods and, thus, leads to tougher competition across intermediate-good firms and to lower prices. Therefore, m can be regarded as a (inverse) measure of the degree of competition in the intermediate-good market. See, e.g., Bucci (2013) for a more exhaustive and formal discussion on the relationship between the aggregate production function (1) and those employed in Ethier (1982) and Benassy (1996a, 1996b, 1998).

effect contrasts with the standard positive specialization effect which results from the increasing availability of differentiated intermediate inputs in aggregate production and which is reflected by the time-varying upper bound of the integral within the square bracket of equation (1). Under symmetry, i.e., when $x_{it} \equiv x_t > 0 \; \forall i \in [0, N_t]$, and with $L_{Yt} > 0$ and $N_t \in (0, \infty)$, equation (1) indicates that an increase in N may have either a net positive (if $\beta < 1$, i.e. the specialization effect is larger than the complexity effect), a net negative (if $\beta > 1$, i.e. the specialization effect is lower than the complexity effect) or else a null net impact on aggregate output (if $\beta = 1$, i.e. the *specialization effect* is offset by the complexity effect).[6] Given perfect competition in the final-good sector, producers take wages, w_{Yt}, and input prices, p_{it}, as given and sell their output at a price also taken as given (and which we normalize to unity).

Each intermediate good, x_i, is produced in a monopolistically-competitive sector using labor as the sole input. The sector uses the technology (see Grossman and Helpman, 1991):

$$x_{it} = l_{it}, \;\; \forall i \in [0, N_t], \;\; N_t \in [0, \infty), \tag{2}$$

where l_{it} is the amount of labor required to produce the *i-th* intermediate good. Given this technology, the marginal cost of production is the wage rate w_{xt}. For a given N_t, equation (2) implies that the total amount of labor employed in the intermediate-good sector at time t, L_{xt}, is given by $\int_0^{N_t} x_{it} di = \int_0^{N_t} l_{it} di = L_{xt}$. Maximization of the generic *i-th* firm's instantaneous profit leads to the standard markup rule $p_{it} \equiv p_t = m w_{xt}$, $\forall i \in [0; N_t]$.[7] Because of the symmetry of producers of intermediate goods, the

[6]Notice that $\beta < 0$ implies that there is no *complexity effect* (the expansion in variety of intermediate goods would, in this case, amplify the standard specialization effect referred to above) whereas $\beta = 0$ trivially implies that only the standard specialization effect exists. Therefore, in order to model explicitly a complexity effect arising from an increase in N, some positive β is needed in our model. However, we do not make any *ad hoc* assumption on the sign and the magnitude of β.

[7]More precisely, we assume that each of these firms is so small to take $\left[\frac{1}{N_t^\beta} \int_0^{N_t} (x_{it})^{1/m} di\right]^{\alpha m - 1}$ as given, hence $\frac{\partial}{\partial x_{it}} \left[\frac{1}{N_t^\beta} \int_0^{N_t} (x_{it})^{1/m} di\right]^{\alpha m - 1} = 0$

price is homogeneous across all intermediate goods i and equal to a constant markup, m, on the marginal production cost, w_{xt}. This then implies $x_{it} \equiv x_t = L_{xt}/N_t$, $\forall i \in [0; N_t]$, and, thus:

$$\pi_{it} \equiv \pi_t = \alpha \left(\frac{m-1}{m}\right) \left(\frac{L_{Yt}}{N_t}\right)^{1-\alpha} \left(\frac{L_{xt}}{N_t}\right)^{\alpha} N_t^{\alpha[m(1-\beta)-1]}, \ \forall i \in [0; N_t]. \tag{3}$$

2.2 R&D activities

R&D is performed by (potential) entrants in the intermediate-good sector. A successful innovation leads to a new blueprint pertaining to a new variety of intermediate good, which is granted a perpetual patent. There is free entry, perfect competition, and constant returns to scale in R&D activities. The R&D production function at the aggregate level is:

$$\begin{aligned} \dot{N}_t &= \frac{1}{X} \cdot \underbrace{N_t}_{\text{knowledge spillover}} \cdot \underbrace{\frac{1}{N_t^{\chi_1}}}_{\text{difficulty effect}} \cdot \underbrace{\frac{1}{L_t^{\chi_2}}}_{\text{dilution effect}} \cdot \underbrace{\frac{1}{L_{Nt}^{1-\lambda}}}_{\text{duplication effect}} \cdot L_{Nt} \\ &= \frac{1}{X} \cdot N_t^{1-\chi_1} \cdot L_t^{-\chi_2} \cdot L_{Nt}^{\lambda}, \end{aligned} \tag{4}$$

where N_t is the number of ideas already invented, L_{Nt} is the labor input allocated to R&D activities, L_t is the total labor force in the economy (and which works as a scale variable for the market dimension), and $X > 0$ is a parameter that controls for the efficiency in R&D activities.

The component $\bar{\delta} \equiv \frac{1}{X} \cdot N_t^{1-\chi_1} \cdot L_t^{-\chi_2} \cdot L_{Nt}^{\lambda-1}$ in equation (4) is external to each individual R&D firm/researcher. Within this completely external component, N_t captures (as in Romer, 1990) the presence of a positive intertemporal knowledge-spillover effect. When $\chi_1 > 0$, the term $N_t^{-\chi_1}$ accounts, instead, for the occurrence of a difficulty effect in innovation (see, e.g., Jones, 1995 and Segerstrom, 1998, among others).[8] When

[8] The difficulty effect in innovation captures the notion that ideas that are easier to discover tend to be

$\chi_2 > 0$, the term $L_t^{-\chi_2}$ denotes the canonical R&D dilution-effect due to population size (as, e.g., in Ha and Howitt, 2007)[9]. Finally, when $0 < \lambda \leq 1$, the term $L_{Nt}^{\lambda-1}$ describes (as in Jones, 1995) the possible duplication-effect related to the amount of labor input employed in R&D. Equation (4) represents the main departure from Bucci, Carbonari and Trovato (2019) who, in its place, employ an otherwise standard Jones (1995)-type aggregate ideas production function.

The parameters χ_i, $i = 1, 2$, may be interpreted as (factor-specific) complexity indices pertaining to R&D activity, as in Sequeira, Gil and Afonso (2018). Notice that these complexity factors add to the complexity effect we consider in final-good production, as captured by a positive value of β in equation (1).

When $\chi_1 > 0$, higher values of N_t imply that the same amount of R&D resources (research labor, L_{Nt}) generates a lower rate of innovation, $\dot{N}_t/N_t$, i.e., there exist diminishing technological opportunities. The presence of the latter is key to the removal of the strong scale effect on growth found in the first-generation of R&D-based growth models (e.g., Jones, 1995): since the marginal impact of an individual researcher on the growth rate of new ideas decreases with the stock of existing ideas, it is possible to sustain a constant positive rate of innovation only by increasing (at a constant rate, too) the number of researchers. This, in turn, is possible in long-run equilibrium solely if population grows at a positive rate. On the other hand, $\chi_2 = 1$ also allows for the removal of the strong scale effect, even if $\chi_1 = 0$. In this case, it is possible to maintain a constant positive rate of innovation as long as the proportion of the number of researchers in the population, L_{Nt}/L_t, is constant. In this scenario, a positive rate of innovation does not require that population grows at a positive rate.

To sum up, the R&D function (4) features a general formulation in line with Ha and Howitt (2007), and nests the following specific well-known cases: under $\chi_1 = \chi_2 = 0$

discovered first, making it harder to find new ideas subsequently.

[9] An economic interpretation of this effect is that "*... a larger population increases the number of people who can enter an industry with a new product, thus resulting in more horizontal innovations, which dilutes R&D expenditure over a larger number of separate projects*" (Ha and Howitt, 2007).

(and $\lambda = 1$), we recover the fully endogenous growth model with strong scale effect by Romer (1990); under $\chi_1 > 0$, and $\chi_2 = 0$ (and $\lambda \leq 1$), we get the *semi-endogenous* growth model without the strong scale effect by Jones (1995); under $\chi_1 = 0$, and $\chi_2 = \lambda$, we get the fully endogenous growth model without the strong scale effect, as e.g. in Dinopoulos and Thompson (1999, 2000).

Since the R&D sector is competitive, N_t is endogenously determined so that the wage rate of one unit of research labor input satisfies the free-entry condition $w_{Nt}L_{Nt} = \dot{N}_t V_{Nt}$. This condition suggests that entry into the R&D sector will cease when total revenues from the discovery of new ideas (the RHS) equal total direct costs related to ideas-production (the LHS), with V_{Nt} indicating the market value of any new idea being discovered.

2.3 Households

The number of identical infinitely lived households is constant and normalized to unity. Hence, the size of population/labor force coincides with the size of the single dynastic family, L, which supplies labor inelastically to firms. Population and, thus, the labor force, grows at a constant exogenous rate $\dot{L}_t/L_t \equiv n > 0$. In this model, as shown above, labor is employed to produce final (consumption) goods (L_Y), intermediate goods (L_x), and ideas (L_N). Since labor is assumed to be homogeneous and perfectly mobile, it will be rewarded according to a unique wage rate at equilibrium. Thus, the following equalities must hold at equilibrium:

$$L_t = L_{Yt} + L_{xt} + L_{Nt} \tag{5}$$

$$w_{Yt} = w_{xt} = w_{Nt} \equiv w_t \tag{6}$$

The objective of the household is to maximize the discounted utility of *per capita* consumption of all its members. With a constant-elasticity-of-substitution instantaneous utility function, the household solves:

$$\max_{\{c_t,a_t\}_{t=0}^{\infty}} U \equiv \int_0^{\infty} \left(\frac{c_t^{1-\theta}-1}{1-\theta} \right) e^{-(\rho-n)t} dt, \tag{7}$$

subject to the usual flow budget contraint, $\dot{a}_t = (r_t - n)a_t + w_t - c_t$, and $a_0 > 0$ given, where $a_t \equiv A_t/L_t$ and $c_t \equiv C_t/L_t$ denote *per capita* asset holdings (taking the form of ownership claims on firms) and *per capita* consumption, respectively. In equation (7), $\theta > 0$ is the inverse of the intertemporal elasticity of substitution in consumption, $\rho > 0$ is the pure subjective discount rate and we have normalized population at time 0 to one, $L_0 = 1$. The representative dynastic family chooses the optimal path of *per capita* consumption and asset holdings, $\{c_t, a_t\}_{t=0}^{\infty}$, taking the real rate of return on asset holdings (the real interest rate), r_t, and the wage rate, w_t, as given. . The solution to this problem gives the well-known Ramsey-Keynes rule.

2.4 Equilibrium and Balanced-Growth Path

In this economy, an *allocation* is a set of time paths for *per capita* consumption and asset holdings $\{c_t, a_t\}_{t=0}^{\infty}$, for the available number of intermediate-good varieties $\{N_t\}_{t=0}^{\infty}$, for prices and quantities of each intermediate good $\{p_t, x_t\}_{t=0}^{\infty}$, and for the real interest rate and wages $\{r_t, w_t\}_{t=0}^{\infty}$. We define an *equilibrium* as an allocation in which: (i) the time paths for consumption and asset holdings are consistent with the solutions of the households' problem (7); (ii) the time paths for prices and quantities of each intermediate good maximize instantaneous profits (3); (iii) the time path for the number of intermediate-good varieties is determined by the free-entry condition in the R&D sector; and (iv) the time paths for the real interest rate and wages are consistent with market clearing.

We can now characterize the *Balanced-Growth Path* (BGP) of this model. We define a BGP equilibrium as follows (see Barro and Sala-i-Martin, 2004, and Strulik, 2005):

Definition 1 *A BGP equilibrium is an equilibrium path along which:*

1. *All variables depending on time grow at constant exponential rates;*

2. *The sectoral shares of labor employment* ($s_j = \frac{L_{jt}}{L_t}$, *with* $j = Y, x, N$) *are constant.*

Let $\chi_1 \neq 0$. It can then be shown that the following results hold along a BGP equilibrium:[10]

$$\gamma_N \equiv \frac{\dot{N}_t}{N_t} = \Psi n \tag{8}$$

$$\gamma_c \equiv \frac{\dot{c}_t}{c_t} = \gamma_a \equiv \frac{\dot{a}_t}{a_t} = \gamma_y \equiv \frac{\dot{y}_t}{y_t} = \Phi\Psi n \tag{9}$$

$$r = \theta\Phi\Psi n + \rho \tag{10}$$

where:

$$\Phi \equiv \alpha \left[m(1-\beta) - 1\right] \gtreqless 0 \tag{11}$$

and

$$\Psi \equiv \left(\frac{\lambda - \chi_2}{\chi_1}\right) \gtreqless 0 \tag{12}$$

and with the transversality condition of the households' optimisation problem satisfied when $\rho > n + (1-\theta)\gamma_y$. Equation (8) gives the BGP equilibrium growth rate of the number of varieties of intermediate goods in the economy. According to equation (9), the *per capita* values of consumption, asset holdings, and income, $y \equiv Y/L$, grow at the same constant rate in the BGP equilibrium. Equation (10) gives the BGP value of the real rate of return on asset holdings.

As in any canonical R&D-based growth model, the BGP growth rate of *per capita* income is strictly related to the BGP innovation rate: $\gamma_y = \Phi\gamma_N$. In turn, as in the basic *semi-endogenous growth* model (Jones, 1995), and provided that $\chi_1 \neq 0$, the innovation rate (γ_N) depends only on the parameters of the innovation technology (λ,

[10] The derivations are available from the authors upon request.

χ_1, and χ_2) and the exogenous population growth rate (n), and is independent of the markup (m) and the parameter measuring the degree of complexity in final output production due to input proliferation (β).[11]

Notice that the BGP growth rate γ_N may display a positive, null, or negative relationship with the growth rate of the population, n, depending on the sign and magnitude of the parameters of the innovation technology. Unlike γ_N, through Φ, γ_y depends instead on m and β. This implies that the degree of market power, the overall level of complexity in the economy, and the interactions between these two variables do ultimately matter for the rate of economic growth and for the sign of the correlation between population and economic growth rates in the long-run.

From equations (8)-(10), it is evident that although γ_N is positive if $\Psi > 0$, r is positive if $\Phi\Psi$ is sufficiently large, i.e. $\Phi\Psi > -\frac{\rho}{\theta n}$. Instead, for γ_y to be positive, $\Phi\Psi$ needs to be strictly greater than zero. In what follows, we present the results of our model in their most general possible form without imposing any *ex ante* restriction on the magnitude of the upper bound of β.[12]

The following proposition summarizes our results.

Proposition 1 *Along the BGP:*

1. $Sign(\gamma_y) = Sign(\Phi\Psi)$*;*
2. $Sign\left(\frac{\partial \gamma_y}{\partial n}\right) = Sign(\Phi\Psi)$*;*
3. *Real* per capita *income growth is equal to zero in the absence of any population change (i.e., when* $n = 0$*);*
4. $Sign\left(\frac{\partial \gamma_y}{\partial m}\right)$ *depends simultaneously on* $Sign(\Psi)$ *and on whether* β *is greater,*

[11] Although m and β do not affect γ_N in the BGP equilibrium, they do have an impact, respectively, on the market value of any generic idea (through π_i), and on the wage accruing to one unit of research labor w_N. As a consequence, m and β are ultimately able, by means of wage equalization, to influence the allocation of the available labor input across sectors (see equations (5) and (6)).

[12] Notice that in our regressions we deal with negative GDP growth rates, as well (see Table A9). In terms of our model, this is explained by the fact that in certain circumstances/countries, following the proliferation of varieties of the same employed intermediate input, the resulting productivity-losses due to increased complexity in production are sufficiently large.

smaller, or else equal to one.

Proof. *The proof of the first part of the proposition is immediate when one takes into account that, in the model, $n > 0$. The proof of the second part is also immediate by observing that: $\frac{\partial \gamma_y}{\partial n} = \Psi\Phi$. To prove the third result, we notice that, in equation (9), $\gamma_y = 0$ if $n = 0$. Regarding the proof of the fourth result, notice that equation (9) also implies that:*

$$\frac{\partial \gamma_y}{\partial m} = \begin{cases} \Psi n\left[\alpha(1-\beta)\right] > 0 & if \quad (\Psi > 0 \, and \, \beta < 1) \, or \, (\Psi < 0 \, and \, \beta > 1) \\ \Psi n\left[\alpha(1-\beta)\right] = 0 & if \quad \Psi = 0 \, or \, \beta = 1 \\ \Psi n\left[\alpha(1-\beta)\right] < 0 & if \quad (\Psi < 0 \, and \, \beta < 1) \, or \, (\Psi > 0 \, and \, \beta > 1) \end{cases}$$

■

Using the definition of Φ and equation (9), we finally observe that several combinations of the signs of γ_y, $\frac{\partial \gamma_y}{\partial n}$, and $\frac{\partial \gamma_y}{\partial m}$ may occur when $\Psi > 0$ (say because $\chi_1 > 0$ and $\lambda > \chi_2$ in equation (4)):

- if $0 < \beta < \frac{m-1}{m}$, then: $\Phi > 0$, $\gamma_y > 0$, $\frac{\partial \gamma_y}{\partial n} > 0$ and $\frac{\partial \gamma_y}{\partial m} > 0$;
- if $\beta = \frac{m-1}{m}$, then: $\Phi = 0$, $\gamma_y = 0$, $\frac{\partial \gamma_y}{\partial n} = 0$ and $\frac{\partial \gamma_y}{\partial m} = 0$;
- if $\frac{m-1}{m} < \beta < 1$, then: $\Phi < 0$, $\gamma_y < 0$, $\frac{\partial \gamma_y}{\partial n} < 0$ and $\frac{\partial \gamma_y}{\partial m} > 0$;
- if $\beta = 1$, then: $\Phi < 0$, $\gamma_y < 0$, $\frac{\partial \gamma_y}{\partial n} < 0$ and $\frac{\partial \gamma_y}{\partial m} = 0$;
- if $\beta > 1$, then: $\Phi < 0$, $\gamma_y < 0$, $\frac{\partial \gamma_y}{\partial n} < 0$ and $\frac{\partial \gamma_y}{\partial m} < 0$.

Proposition 1 suggests that the sign of $(1-\beta)$, by determining the sign of $\partial\Phi/\partial m$, is a key determinant of the impact of the intermediate sector's markup on real *per capita* growth. In more detail, if $\beta < 1$, the specialization gains obtained from an expansion in intermediate-input variety are larger than the possible losses due to more complexity in production and, thus, $\frac{\partial \Phi}{\partial m} > 0$. If $\beta > 1$, the specialization gains are smaller than the possible losses due to more complexity in production and $\frac{\partial \Phi}{\partial m} < 0$. Finally, when $\beta = 1$, the specialization gains are offset by the possible losses due to more complexity in production and $\frac{\partial\Phi}{\partial m} = 0$. These results, combined with the sign of Ψ, determine the sign of $\frac{\partial\gamma_y}{\partial m}$, as shown in Proposition 1. The sign of Ψ depends, in turn, on the sign of $(\lambda - \chi_2)$, that is, on the balance between dilution and duplication effects in R&D activities, as shown by equation (4). Therefore, conditional on the sign of Ψ, a decrease in m, by increasing the elasticity of substitution across intermediate inputs and, hence, the toughness of competition in this industry, can imply a lower, or a higher, or else no effect at all on *per capita* income growth (i.e., the degree of competition in the product market and the economic growth are ambiguously correlated).

The result that, in the absence of demographic change, there is no growth in real *per capita* income is a distinctive characteristic of any basic *semi-endogenous growth* model (see Jones, 1995). Our setting, however, includes additional features that cannot be found in canonical *semi-endogenous growth* theory. In particular, the relation between γ_N and γ_y is mediated, in our case, by the term $\Phi \equiv \alpha\,[m(1-\beta)-1]$, unlike Jones (1995), where $\gamma_N = \gamma_y = \gamma$, with γ being a definitely positive function of n. Thus, while in Jones (1995), $\beta = 0$ and, hence, $\Phi \equiv \alpha(m-1) > 0$, in our model Φ can also be negative. This occurs, for any given $m > 1$, when β is sufficiently large, $\beta > (m-1)/m \in (0,1)$, i.e. when the complexity effect (in final-good production) is strong enough. When this is true ($\Phi < 0$), innovation output and long-run growth may exhibit diverging patterns, as in the suggestive evidence provided at the beginning of the paper. In this case, it is also possible that an increase in the rate of population growth would yield (unlike Jones, 1995) a negative impact on *per capita* income growth. This happens whenever $\Psi > 0$, as n affects positively γ_N in such a case.[13] Finally, unlike Jones (1995), in our model $\Psi \equiv \left(\frac{\lambda-\chi_2}{\chi_1}\right) \gtreqless 0$, because of our explicit and simultaneous consideration of a dilution, a difficulty, and a duplication effect in R&D activity. Depending on

[13] There is another trait that makes our model different from Jones (1995): in equation (9) the growth rate of the economy depends not only on n and, among others, the parameters χ_1, χ_1, λ and β, but also on m. This specific difference with respect to Jones (1995) can be ascribed to the fact that in our model we have: (i) postulated that intermediate firms produce with labor (rather than forgone consumption) and, more importantly, (ii) disentangled the intermediate firms' gross markup of price over the marginal production cost from the factor-input shares in GDP. It can be easily demonstrated that, using exactly the Jones' assumptions, our model is able to reproduce the same BGP growth rate of the Jones' economy.

the combined sign of Ψ and Φ, innovation (physical) inputs and economic growth may exhibit diverging patterns, also as in the suggestive evidence presented at the beginning of the paper.

3 Quantitative analysis

The theory developed above rests on a precise notion of complexity, which emerges as a consequence of the interaction between the entire labor force and the aggregate research effort, i.e. the number of new ideas and of researchers at work. The combination between "complexity parameters"(χ_1, χ_2, λ and β), "technological parameters"(α and m) and the exogenous population growth rate (n) determines the long-run growth rate of a country's *per capita* income. We now confront the theoretical predictions with OECD data, by allowing that these fundamental parameters may vary over time and across countries. To do this, we employ a Hidden Markov Model. Using our theoretical model as a guidance, the parameter estimates for population growth and the interaction between "complexity parameters" and "technological parameters" let us infer which effect is predominant, between complexity and specialization. Notably, this econometric approach provides a classification of our countries' long-run growth rates, in which the growth-regime membership depends on changes in our theoretical growth determinants.

3.1 Econometric strategy

Our econometric strategy is articulated in two steps. In the first step, in order to obtain country specific estimates for χ_1, χ_2 and λ, we take logs on both sides of equation (4) and regress the growth rate of patents (by residents) on the number of patents, population and number of researchers, for each country i, with $i = 1, \ldots, I$:

$$\left(\frac{\dot{N}}{N}\right)_t = \alpha_0 + \alpha_1 N_t + \alpha_2 L_t + \lambda L_{Nt} + \epsilon_t \tag{13}$$

where $\widehat{\chi}_j = -\widehat{\alpha}_j$, with $j = 1, 2$, and $\epsilon_{it} \sim$ i.i.d $N(0, \sigma_t)$. For each country i, then, we use the OLS estimates $\widehat{\lambda}$, $\widehat{\chi}_1$ and $\widehat{\chi}_2$ to compute the theoretical variable $\gamma_{N,i}$, according to equations (8) and (12).

In the second step, we exploit the longitudinal nature of our data set and explore the transitions from phases of low (and even negative) growth to phases of high growth and vice versa, by taking into account the unobserved heterogeneity. In this perspective, the real GDP (RGDP, hereafter) growth is interpreted as a result of countries switching between distinct growth regimes (as in Kerekes, 2012). Let $\{\mathbf{\Gamma}_{it}; i = 1, ..., I, t = 1, ..., T\}$ denote sequences of multivariate longitudinal observations for the real GDP growth rate recorded on I countries and T year, where $\mathbf{\Gamma}_{it} = (\gamma_{it1}, ..., \gamma_{itP})^{\mathsf{T}} \in \mathbb{R}^P$, and let $\{\mathcal{S}_{it}\}$ be a first-order Markov chain defined on the state space $\{1, ..., k, ..., K\}$. A Hidden Markov Model (HMM, hereafter) is a stochastic process consisting of two parts: the underlying unobserved process $\{\mathcal{S}_{it}\}$, fulfilling the Markov property, i.e.

$$Pr(\mathcal{S}_{it} = s_{it}|\mathcal{S}_{i1} = si1, \mathcal{S}_{i2} = s_{i2}, \ldots, \mathcal{S}_{it-1} = s_{it-1}) = Pr(\mathcal{S}_{it} = s_{it}|\mathcal{S}_{it-1} = s_{it-1})$$

and the state-dependent observation process $\{\mathbf{\Gamma}_{it}\}$ for which the conditional independence property holds:

$$f\left(\mathbf{\Gamma}_{it} = \gamma_{it}|\mathbf{\Gamma}_{i1} = \gamma_{i1}, \ldots, \mathbf{\Gamma}_{it-1} = \gamma_{it-1}, \mathcal{S}_{i1} = s_{i1}, \ldots, \mathcal{S}_{it} = s_{it}\right) = f\left(\mathbf{\Gamma_{it}} = \gamma_{it}|\mathcal{S}_{it} = s_{it}\right)$$

where $f(\cdot)$ is a generic probability density function (Maruotti and Punzo, 2017). The distribution of $\mathbf{\Gamma}_{it}$ depends only on s_{it}, i.e. $\mathbf{\Gamma}_{it}$ is conditionally independent given the s_{it}. In our baseline specification, we assume that the state-dependent distributions come from a parametric family of continuous or discrete distributions. Thus, the unknown parameters in the HMM involve both the hidden Markov chain and the state-dependent distributions of the random variable $\mathbf{\Gamma}_{it}$. The hidden Markov chain has K states with initial probabilities $\pi_{ik} = Pr(\mathcal{S}_{i1} = k)$, $k = 1, \ldots, K$, and transition probabilities

$$\pi_{i,k|j} = Pr(\mathcal{S}_{it} = k|\mathcal{S}_{t-1} = j)$$

with $t = 2, \ldots, T$ and $j, k = 1, \ldots, K$. For sake of simplicity, we will consider a homogeneous HMM in which common transition and initial probabilities are assumed, i.e. $\pi_{i,k|j} = \pi_{k|j}$ and $\pi_{ik} = \pi_k$, for $i = 1, \ldots, I$. The transition probabilities are constant over time and among individuals.[14]. The initial probabilities are collected in the K-dimensional vector π, while the time-homogeneous transition probabilities are collected in the $K \times K$ transition matrix $\mathbf{\Pi}$.

The empirical counterpart of the equation for the BGP growth rate of real *per capita* GDP (9) can be written as:

$$E(\gamma_{it}|\mathbf{\Gamma}_{i,t-1}, s_{it}; \mathbf{m_{it}}, \mathbf{n_{it}}) = \omega_{0,s_i} + \mathbf{m_i}^{\mathsf{T}}\omega_{1,s_i} + \mathbf{n_i}^{\mathsf{T}}\omega_{2,s_i} \tag{14}$$

where, for each country i and time t, $\gamma_{i,t}$ is the RGDP growth rate, $\mathbf{m_i}^{\mathsf{T}}$ is the vector of the annual products of intermediate sector's markup and γ_N and $\mathbf{n_i}^{\mathsf{T}}$ is the vector of the annual observations of γ_N.[15] The parameters ω_{2,s_i} and ω_{1,s_i} capture the state-specific unobserved factors that affect RGDP growth, through γ_N and its interaction with the intermediate sector's markup, respectively.

3.2 The data

Our final dataset merges information drawn from four different sources. First, we get data on real GDP growth, population growth, human capital and exchange rate (to convert all the monetary values in constant 2005 US$) from the Penn World Table

[14]This assumption can be easily relaxed to include covariates and/or unit-specific random effects (see, e.g., Maruotti and Rocci, 2012).

[15]In this section, for notational simplicity, we omit the subscript y on GDP growth rate γ.

database (PWT, hereafter).[16]

Second, in order to construct a measure for the markup in the intermediate sector[17], we use the EUKLEMS database, which collects data on output, productivity, employment (skilled and unskilled), physical capital at industry level, for all European Union member states and for five of the high developed countries (US, Japan, Korea, Canada and Australia) from 1970 to 2007.[18] At the lowest level of aggregation, data are collected for 72 industries according to the European NACE revision 1 classification. We proxy the intermediate sector with the sum of the following industries: *basic metals and fabricated metal; electrical and optical equipment; electricity; gas and water supply; machinery; other non-metallic mineral; rubber and plastics; textile, leather and footwear; transport and storage; transport equipment; wood and cork.* Following Griffith, Harrison and Macartney (2006), we compute the markup index for the intermediate sector, of country i at time t, as follows:[19]

$$m_{it} = \frac{Value\ Added_{it}}{Total\ Labor\ Costs_{it} + Total\ Capital\ Costs_{it}} \tag{15}$$

where all variables are in nominal prices and are provided by EUKLEMS.[20] In our baseline regressions, m is an index set equal to 1 in the base year 1995.[21] Third, the "number of ideas already invented", i.e., a country's stock of knowledge, has been measured through the patent applications by residents, collected by the World Bank.[22] Finally, OECD provides data on the number of researchers at work.[23]

Our final dataset consists of a sample that includes 18 OECD countries, with a time span ranging from 1983 to 2007. Table A8 presents summary statistics. On average,

[16] Time span: 1983-2007, 2005 as reference year. For more information on the PWT see: http://www.rug.nl/ggdc/productivity/pwt/.

[17] There exist several methods to measure markups. See Nekarda and Ramey (2013) for a review.

[18] For more information on the EUKLEMS database see: http://www.euklems.net/.

[19] Griffith, Harrison and Macartney (2006) point out that this approach is equivalent to that proposed by Roeger (1995). See Klette (1999) for a discussion.

[20] For details see Timmer et al. (2007).

[21] Results with alternative measures for sector profitability are discussed in section 3.5.

[22] Data are available at the web page: https://data.worldbank.org/indicator/ip.pat.resd.

[23] Data are available at the web page: https://data.oecd.org/rd/researchers.htm.

our data seem reject the hypothesis of strong scale effect, i.e. $\widehat{\chi}_1 > 0$.[24] Table A9 shows that all countries in the sample experienced a positive average RGDP growth rate along the period under observation. Since we are dealing with OECD countries, this is not surprising. Notice finally that the sample also contains observations in which the 5-year average *per capita* RGDP growth rate is negative.

3.3 Results

Our initial state distribution assigns equal probability to all regimes and we let the number of regimes vary between zero (that is a homogenous time dependent process) and five. To determine the number of regimes we use the Bayesian Information Criterion (BIC), which rejects a model without clustering in favor of a model containing four regimes, as in Jerzmanowski (2006) and Kerekes (2012) (see Table A10). On the basis of the 5-years average *per capita* RGDP growth rate, we label the regimes as follows: *slow growth*, *steady growth*, *sustained growth* and *miracle growth*.

Table 2: Growth regression, equation (14)

	Constant	γ_N	$\mathbf{m} \times \gamma_N$	**Cluster standard deviation**	**5-year avg. *per capita* RGDP growth rate (%)**
OLS FE	0.028***	0.023	-0.011		
HMM:					
1–Slow growth	0.821***	0.032***	0.031***	0.712***	1.134
2–Steady growth	2.531***	0.029***	0.007***	0.414***	2.676
3–Sustained growth	3.936***	0.027***	0.006***	0.561***	4.163
4–Miracle growth	5.794***	-0.001***	0.092***	0.997***	6.615

Significance levels: : 10% **: 5% ***: 1%.*

Table 2 presents our main results. The first line presents the OLS fixed effects estimates of the growth equation while the rest of the table reports the regime-specific parameter estimates; the last column of the table provides the implied long-run growth rates of each regime. The *slow growth* regime is the cluster with the lowest growth rate of RGDP (1.134%, standard deviation of 0.712). This regime captures the long lasting

[24] Hungary presents an extremely high value for χ_2 (160.182). When we exclude Hungary from the sample, the sample mean becomes -2.594 while the standard deviation declines to 2.496. Our estimates are robust to the exclusion of this outlier.

stagnation of the Japanese economy, but also the severe downturns in economic activity experienced in France over the period 1998-2006 (see Cette, Fernald and Mojon, 2016), Spain at the end of the 80s (see Blanchard and Jimeno, 1995) and Sweden in the mid-90s as a consequence of the second OPEC crisis (see Jonung and Hagberg, 2005). The *steady growth* regime is characterized by a moderate growth rate of RGDP (2.676%) and variability (with a standard deviation of 0.414, this cluster presents the lowest level of variability). This regime captures, for instance, the growth experience of countries like Germany, Australia, Belgium, Canada, Denmark and US. Interestingly, these countries spent most of the time within this regime: this means that, despite cyclical fluctuations, no persistent changes has been found – along the period under observation – in their growth trajectories. The *sustained growth* regime is characterized by higher long-run growth rate (4.163%) and variability (0.561): Austria (continuously from 1989 to 2000) and Spain (continuously from 1993 to 2002) are are the most frequent countries within this regime. Finally, the *miracle growth* regime features the highest long-run growth rate (6.615%) and standard deviation (0.997). This regime takes account of the spectacular growth performance of the Irish economy (which continuously stays within this cluster, from 1990 to 2005) but also of the growth successes of Japan (from 1988-1993), Portugal (in the mid-90s) and Spain (in the first decade of 2000s). Table 3 presents our classification.

The annual growth rate of population is a key ingredient of our theory. In our data, its impact on *per capita* RGDP growth is found to be generally positive. In the OLS fixed effects model, $d\gamma/dn$=0.115 (p-value=0.081), while in the HMM we obtain: $d\gamma/dn$=0.389 (p-value=0.000) in the *slow growth* regime, $d\gamma/dn=-0.269$ (p-value=0.000) in the *steady growth* regime, $d\gamma/dn$=0.485 (p-value=0.000) in the *sustained growth* regime and $d\gamma/dn$=0.083 (p-value=0.000) in the *miracle growth* regime.

We pointed out in Section 2 that, depending on the sign of Φ, the losses due to complexity may be larger or smaller than the gains due to specialization.[25] We also

[25]See the discussion at the end of Paragraph 2.2.

Table 3: Classification

Country	**Regime 1** *Slow growth*	**Regime 2** *Steady growth*	**Regime 3** *Sustained growth*	**Regime 4** *Miracle growth*
Australia	4	14	4	0
Austria	4	6	12	0
Belgium	3	14	4	0
Canada	7	12	0	0
Denmark	7	12	3	0
France	12	7	3	0
Germany	1	16	5	0
Hungary	2	6	5	0
Ireland	0	0	4	16
Italy	7	8	7	0
Japan	10	3	2	6
Poland	1	6	0	0
Portugal	3	4	8	4
Slovenia	0	7	2	0
Spain	2	6	10	4
Sweden	5	8	9	0
United Kingdom	9	8	5	0
United States	5	12	0	0

showed that, depending on the sign and magnitude of Φ, innovation and long-run growth may exhibit a weak or even negative correlation, in line with the suggestive evidence provided in Section 1. Notice that the theoretical growth equation (9) and the empirical equation (14) differ because of the presence in the latter of a random intercept, to account for omitted covariates or country-specific heterogeneity, which are not captured by the observed covariates. Thus, in order to obtain an estimate of Φ more directly linkable to the theoretical model, we use the estimates of (14) with $\omega_{0,s_i} = 0$ to get the regime-specific $\widehat{\Phi}_{s_i} = \widehat{\omega}_{1,s_i}\overline{m}_{s_i} + \widehat{\omega}_{2,s_i}$ with $i = 1, \ldots, 4$, where $\overline{m}_{s_i}$ is the regime-mean of the intermediate sector markup.[26] We find that $\widehat{\Phi}$ is positively correlated with the long-run RGDP growth rate, being 0.072 (p-value=0.000) in the *slow growth* regime, 0.090 (p-value=0.000) in the *steady growth* regime, 0.278 (p-value=0.000) in the *sustained growth* regime and 0.495 (p-value=0.000) in the *miracle growth* regime. Our theory predicts that a higher Φ implies a stronger positive impact of the proliferation of "new ideas"on the long-run *per capita* real GDP dynamics. The switch between different growth regimes, therefore, implies at the same time the

[26]To consider the model without intercept in a logitudinal setting, we estimate ω_{1,s_i} and ω_{2,s_i} by Maximum Likelihood.

transition from phases in which complexity hampers more the GDP growth (lower Φ) to phases in which complexity hampers it less (higher Φ), and vice-versa.[27]

Table 4: Transition matrix

To Regime / *From Regime*	*1–Slow growth*	*2–Steady growth*	*3–Sustained growth*	*4–Miracle growth*
1–Slow growth	0.488	0.402	0.085	0.024
2–Steady growth	0.245	0.565	0.163	0.027
3–Sustained growth	0.060	0.386	0.530	0.024
4–Miracle growth	0.000	0.000	0.267	0.733

Note: the entry in row j, column i should be interpreted as $p_{ij} = P(st = j|s_{t-1} = i)$.

Table 4 presents the cluster-specific transition probability matrix, which indicates the probability of moving from the column regime to the row regime. The highest persistency is found in the *miracle growth*: when a country (namely, Ireland, Japan, Portugal and Spain) spends a year in this regime, it continues to growth at a so terrific rate in the subsequent year with a probability equal to 0.733; otherwise it moves into the *sustained growth* regime. The *steady growth* and the *sustained growth* regimes also show a high persistency: when a country finds itself in one of the two regimes, the probabilities of remaining there in the subsequent year are 0.565 and 0.530, respectively. At the same time, it is more likely that a country in the *sustained growth* regime slowdowns, moving towards the *steady growth* one (0.386), rather than the opposite occurs (0.163). Finally, a country in the *slow growth* regime will improve its conditions, moving into the *steady growth* regime with a probability of 0.402, the probability of moving into regimes characterized by faster growth being 0.109.

For a better comprehension of our results, Table 5 reports the conditional means for our theoretical growth determinants, whereas Table 6 reports conditional means for a list of other structural factors which have been not explicitly taken into account in our theoretical framework. A rich relation emerges between growth and the theoretical growth determinants across the four growth regimes. Not surprisingly, since we are dealing with a sample of OECD countries, we do not observe significant differences in

[27] With exclusive reference to the countries in Figures 1 and 2, we find a $\widehat{\Phi}$=0.186 (p-value=0.000).

growth fundamentals across clusters. According to our theory, a higher growth rate of patents is found associated with a faster long-run RGDP growth, which is also positively correlated with the saving rate and gross capital formation. Interestingly, five non-monotonic relationships with RGDP growth rate emerge: the annual population growth rate shows an U-shaped effect on it while the effects of intermediate sector's markup and also λ, χ_1 and χ_2 present an inverted U-shaped pattern. Thus, our evidence suggests that the duplication, difficulty and dilution effects, defined in equation (4), also relate non-linearly with long-run growth. These non-linearities may explain why the underlying positive relationship between patents or researchers and the long-run RGDP growth is not clearly found in the aggregate data, although it is predicted by the theory.

Table 5: Key growth-affecting factors, conditional means

Regime	**m**	$\widehat{\lambda}$	$\widehat{\chi_1}$	$\widehat{\chi_2}$	$\dot{\mathbf{N}}/\mathbf{N}$	**n**
1–Slow growth	1.061	0.011	0.250	-0.520	6.073	0.531
2–Steady growth	1.193	0.011	0.273	4.413	10.442	0.517
3–Sustained growth	1.127	0.184	0.264	-2.917	11.868	0.398
4–Miracle growth	1.051	-0.157	0.248	-4.193	12.496	0.659

Table 6: Other growth-affecting factors, conditional means

Regime	**saving rate**	**investment share**	**public spending share**
1–Slow growth	0.424	0.253	0.170
2–Steady growth	0.423	0.259	0.168
3–Sustained growth	0.423	0.277	0.169
4–Miracle growth	0.484	0.280	0.143

3.4 Other growth fundamentals

Finally, we estimate several specifications of the Multinomial Logit Model (MNL) to assess the role of the other potential long-run growth fundamentals in affecting a country cluster's membership. In this exercise, we take the *slow growth* regime as reference.

Therefore, the multinomial logistic regression evaluates the relative probability of being in one of the remaining growth regimes against the reference, using a linear combination of predictors. The obtained MLE-estimated coefficients represent the effects of every predictor variable in the log-odds of being in any other regime versus the reference regime. As predictor variable we employ the human capital index (`hc`) provided by PWT. Results are reported in Table 7. *Ceteris paribus*, a unit increase in the human capital index increases the probability of being in the *steady growth* regime, relative to the *slow growth* regime, by a multiplicative factor of exp(0.191)=1.210, i.e. increasing it by 21% (p-value=0.000). This confirms that, even for high income countries, human capital is one of the underlying factors that lead to the transition from stagnation to growth. The increase in educational level measured by the human capital index, however, does not explain the take-off toward regimes characterized by higher growth rates, being this ultimately related to R&D intensity and specialization.

Table 7: MNL Model for regime membership

From *slow growth* to...	**Coef.**	**Std.Err.**
... steady growth		
hc index	0.191***	0.014
constant	0.132	1.658
... sustained growth		
hc index	-1.741***	0.111
constant	5.424***	1.661
... miracle growth		
hc index	-2.879***	0.143
constant	7.647***	1.961

3.5 Robustness checks

In this paragraph, we briefly discuss how estimates and classification behave in response to changes in the econometric specification and/or in the way we measure some explanatory variables.[28]

$\widehat{\Psi}$. First, we check the robustness of our results using an alternative estimate of

[28] As our results survive various robustness checks, for the sake of brevity we do not present and discuss in detail all the parameter estimates, which are available upon request.

Ψ. For each country i, with $i = 1, \ldots, I$, we regress the growth rate of patents (by residents) on the demographic growth rate, according to (8):

$$\left(\frac{\dot{N}}{N}\right)_t = \Psi n_t + \epsilon_t \tag{16}$$

with $\epsilon_{it} \sim$ i.i.d $N(0, \sigma_t)$. For each country, then, we use the OLS estimates $\widehat{\Psi}$ to compute the theoretical variable $\gamma_{N,i}$. Despite some modifications occur in the composition of the clusters, we still identify four growth regime. The regressions, however, result in less significant estimates and lower explanatory power.

$\widehat{\chi}_1$ and $\widehat{\chi}_2$. The HMM rests on the fact that the growth regime transition probabilities are time dependent. A theoretical explanation for having (at least some) time dependent parameter in our model, can be found in the possible interplay between N and L_N. To take this into account, we modify equation (4) as follows:

$$\dot{N}_t = \frac{1}{X} \cdot N_t^{1-\chi_1(N_t)} \cdot L_t^{-\chi_2(N_t)} \cdot L_{Nt}^{\lambda}, \tag{5'}$$

The exponents $\chi_i(N)$, $i = 1, 2$, are now intended as (factor-specific) complexity indices pertaining to R&D activity and depend positively on N, in line with Sequeira, Gil and Afonso (2018). As the number of varieties of specialized (intermediate) goods rises, the economy becomes more complex because there is an increase in the diversity (and maybe redundancy) of the different components (i.e., intermediate goods) that need to be assembled for producing the final output. This, in turn has two effects. Concerning $\chi_1(N_t)$, we postulate that a proliferation of N_t amplifies the difficulty with which new ideas are discovered starting from the oldest ones. Similarly, as for $\chi_2(N_t)$, we maintain that the same increase in N_t strengthens the effect by which a rise in population size ultimately leads to a dilution of the total amount of R&D expenditures over a larger number of (more) dispersed research projects. However, we also assume that the indices will eventually level off, i.e., $\chi_i(N_t) \to \chi_i$, $i = 1, 2$, despite the continuous increase in the number of varieties, reflecting the fact that part of the modern innovations (leading to new varieties of goods) have a stabilizing role in the complexity of the economies.[29] The empirical counterpart of equation (5') is given by the following reduced form equation:

$$\left(\frac{\dot{N}}{N}\right)_t = \alpha_0 + \alpha_1 N_t + \alpha_2 N_t L_t + \lambda L_{Nt} + \epsilon_t \tag{17}$$

where $\widehat{\chi}_j = -\widehat{\alpha}_j$, with $j = 1, 2$, and $\epsilon_{it} \sim$ i.i.d $N(0, \sigma_t)$.The results of the HMM model obtained using these alternative $\widehat{\chi}_1$, $\widehat{\chi}_2$ and $\widehat{\lambda}$ are in line with those presented above. We still identify four growth regimes: *slow growth* (with 5-years average growth rate of 0.994%), *steady growth* (2.654%), *sustained growth* (3.826%) and *miracle growth* (6.371%). Clusters' standard deviations follow the same pattern of those in the baseline model and only slight modifications occurs in clusters' composition. Estimates, however, are less accurate (see Table A11).

m. Several robustness checks are carried out for the intermediate sector's markup. As usual, when dealing with index numbers, results can be dependent on the base year. We use an alternative base year, the 1985, and we don't find any significant change in our estimates. We also apply the procedure proposed by Roeger (1995) to tackle the issue of the empirical measurement of the markups. In particular, assuming the markup constant over the period under observation, for each country, we run the following regression:

$$SR_t - SRP_t = \left(1 - \frac{1}{\tilde{m}}\right)[(\Delta p_t + \Delta Q_t - u_t) + (\Delta r_t + \Delta K_t - v_t)] \tag{18}$$

where: $(SR_t - SRP_t)$ is the difference between the Solow residual and the price-based Solow residual, $\tilde{m}$ is the intermediate sector's markup, $(\Delta p_t + \Delta Q_t)$ is the nominal

[29] Sequeira, Gil and Afonso (2019, p. 107) point out that, at least in modern economies, "*some inventions have reduced or attenuated the effect of complexity either directly or indirectly. Computers have allowed calculation of things and analysis of data and models in ways impossible before, while, as an indirect effect, one could argue that the development of machinery that replaces humans (e.g., earth-moving equipment) have allowed more humans to spend their time managing complexity and its effects*". Their empirical results indeed suggest a stabilization of complexity over the long-run.

output growth, $(\Delta r_t + \Delta K_t)$ is the growth of capital cost. Both capital costs and nominal output are measured with error, with $v_t \sim i.i.d.(0, \sigma_v)$ and $u_t \sim i.i.d.(0, \sigma_u)$.[30] Estimates for $\tilde{m}$ are presented in Table A12. Because of the lack of time variability of $\tilde{m}$, using this estimated markup, rather than the index computed using equation (15), implies less accurate estimates.

Human capital. Following Bucci (2015), we estimate a version of the model in which the growth rate of the human capital (measured by the `hc` index provided by PWT) replaces the growth rate of population. Let $\tilde{n}$ denote the rate of change of the human capital index. In the OLS fixed effects model we get that $d\gamma/d\tilde{n}$=0.089 (p-value=0.001) while no significative changes are observed in the HHM.

4 Concluding remarks

In this paper, we advance an explanation for a cross-country evidence which is inconsistent with most endogenous growth models: the coexistence of increasing trends in aggregate research effort and the no corresponding increases in long-run *per capita* income growth. Building on Jones (1995), we develop a model in which long-run growth is determined by the interplay between "complexity parameters", "technological parameters"and the (exogenous) population growth rate. In this setting, the coexistence of stagnant *per capita* income growth and the proliferation of "new ideas"is shown to be the equilibrium response to greater complexity. Using the theoretical model as a guidance, we estimate a regime-switching model of growth (namely, a Hidden Markow Model), which allows for an endogenous classification mechanism, using a sample of OECD countries. We find four different regime for the long-run *per capita* real GDP growth. Each country's growth pattern is the result of transitions between distinct growth regimes. The transitions are determined by regime-specific transition probabilities. In light of our classification, we further establish several facts about the transition

[30] Christopoulou and Vermeulen (2012) apply the same empirical strategy to estimate price-marginal cost ratios for 50 sectors in eight Euro area countries and the US over the period 1981-2004.

between different growth regimes. We find that growth accelerations, in high income countries, are strongly associated with patenting activity, i.e. the annual growth rate of patent applications (by residents); growth failures meanwhile are characterized by a weaker degree of *specialization*, i.e. a lower average estimated value for Φ. In this situation, in which complexity in production is more harmful for long-run growth, our theory suggests that increases in the population growth rate allow to achieve faster growth. An alternative way-out, which our model can easily be extended to account for, is to increase the investment in education: using a Multinomial Logit Model, we find that the level of human capital – here proxied by the `hc` index provided by PWT – positively affects the transition from a regime with modest growth towards a regime of faster growth.

References

[1] Aghion, Philippe and Peter Howitt (1998), *Endogenous Growth Theory*, Cambridge, MA: The MIT Press.

[2] Barro, Robert.J., and Xavier Sala-i-Martin (2004). *Economic Growth*, MA: MIT Press.

[3] Baum, Leonard. E. and J.A. Eagon (1967). An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. *Bulletin of the American Mathematical Society*, 73(3), 360-363.

[4] Baum, Leonard. E. and and Ted Petrie (1966). Statistical inference for probabilistic functions of finite-state Markov chains. *Annals of Mathematical Statistics*, 37(6), 1554-1563.

[5] Benassy, Jean-Pascal (1996a) Taste for Variety and Optimum Production Patterns in Monopolistic Competition. *Economics Letters* 52(1), 41-47.

[6] Benassy, Jean-Pascal (1996b) Monopolistic Competition, Increasing Returns to Specialization and Output Persistence. *Economics Letters* 52(2), 187-191.

[7] Benassy, Jean-Pascal (1998) Is there always too little research in endogenous growth with expanding product variety? *European Economic Review* 42(1), 61-69.

[8] Blanchard, Olivier and Juan F. Jimeno (1995) Structural Unemployment: Spain versus Portugal. *The American Economic Review, Papers & Proceedings* 85(2), pp. 212-218.

[9] Bucci, Alberto (2013) Returns to Specialization, Competition, Population, and Growth. *Journal of Economic Dynamics and Control* 37(10), 2023-2040.

[10] Bucci, Alberto (2015) Product Proliferation, Population, and Economic Growth. *Journal of Human Capital* 9(2), 170-197.

[11] Bucci, Alberto, Lorenzo Carbonari and Giovanni Trovato (2019) Variety, Competition, and Population in Economic Growth: Theory and Empirics. *Macroeconomic Dynamics*, forthcoming.

[12] Cette, Gilbert, John Fernald and Benoît Mojon (2016) The pre-Great Recession slowdown in productivity. *European Economic Review* 88, pp. 3-20.

[13] Dinopoulos, Elias and Peter Thompson (1999) Scale effects in Schumpeterian models of economic growth. *Journal of Evolutionary Economics* 9(2), 57-185.

[14] Dinopoulos, Elias and Peter Thompson (2000) Endogenous growth in a cross-section of countries. *Journal of International Economics*, 51(2), 335-362.

[15] Ethier, Wilfred J. (1982) National and International Returns to Scale in the Modern Theory of International Trade. *American Economic Review* 72(3), 389-405.

[16] Ferrarini, Benno and Pasquale Scaramozzino (2016) Complexity, Specialization and Growth. *Structural Change and Economic Dynamics* 37, 52-61.

[17] Griffith, Rachel, Rupert Harrison and Gareth Macartney (2006) Product Market Reforms, Labour Market Institutions and Unemployment. *Economic Journal* 117, 142-166.

[18] Grossman, Gene and Elhanan Helpman (1991) *Innovation and Growth in the Global Economy*, MA: MIT Press.

[19] Hausmann, Ricardo and Bailey Klinger (2006) Structural Transformation and Patterns of Comparative Advantage in the Product Space. *KSG Working Paper No. RWP06-041*, Harvard University: Center for International Development.

[20] Jerzmanowski, Michay (2006) Empirics of hills, plateaus, mountains and plains: A Markov-switching approach to growth. *Journal of Development Economics* 81, 357- 385.

[21] Jones, Charles I. (1995) R&D-based Models of Economic Growth. *Journal of Political Economy* 103, 759-84.

[22] Kerekes, Monika (2012) Growth Sustaineds and failures in a Markov switching classification model of growth. *Journal of Development Economcs* 98, 167-177.

[23] Klette, Tor J. (1999) Market Power, Scale Economies and Productivity: Estimates from a Panel of Establishment Data. *Journal of Industrial Economics* 47(4),451-476.

[24] Kremer, Michael (1993) The O-Ring Theory of Economic Development. *Quarterly Journal of Economics* 108(3), 551-575.

[25] Maruotti, Antonello, and Roberto Rocci (2012) A mixed non-homogeneous hidden Markov model for categorical data, with application to alcohol consumption. *Statistics in Medicine* 31 (9), 871-886.

[26] Maruotti, Antonello, and Antonio Punzo (2017) Model-based time-varying clustering of multivariate longitudinal data with covariates and outliers *Computational Statistics & Data Analysis*, 113, 475-496.

[27] Nekarda, Christopher J., and Valerie A. Ramey (2013) The Cyclical Behavior of the Price-Cost Markup. *NBER Working Paper*, No. 19099.

[28] Pintea, Mihaela and Peter Thomphson (2007) Technological complexity and economic growth. *Review of Economic Dynamics* 10, 276-293.

[29] Rabiner, Lawrence R. (1989) A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, 77, 257-286.

[30] Roeger, Werner (1995) Can Imperfect Competition Explain the Difference Between Primal and Dual Productivity Measures? Estimates for US Manufacturing. *Journal of Political Economy* 103, 316-330.

[31] Romer, Paul M. (1990) Endogenous Technological Change. *Journal of Political Economy*, 98(5), 71-102.

[32] Sala-i-Martin, Xavier, Gernot Doppelhofer and Ronald I. Miller (2004) Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *American Economic Review* 94(4), 813-835.

[33] Sequeira, Tiago Neves, Pedro Mazeda Gil and Óscar Afonso (2018) Endogenous growth and entropy. *Journal of Economic Behavior & Organization*, 154, 100-120.

[34] Strulik, Holger (2005) The Role of Human Capital and Population Growth in R&D-Based Models of Economic Growth. *Review of International Economics* 13(1), 129-145.

[35] Timmer, Marcel, Ton van Moergastel, Edwin Stuivenwold, and Gerard Ypma, M. (2007) *EUKLEMS Growth and productivity accounts. Part I: methodology*, The Netherlands: University of Groningen & University of Birmigham.

Appendix

Table A8: Summary statistics

Variable	Description	Obs.	Mean	Std. dev.	Min	Max
γ_y	5-year avg. *per capita* RGDP growth rate	344	3.011	1.695	-0.938	9.045
m	intermediate sector's markup index (1995=1)	344	1.133	0.503	0.719	4.794
n	annual population growth rate	344	0.504	0.488	-0.290	2.205
N	number of patents by residents	344	14,177.910	23,380.950	6.000	12,8152.200
L_N	number of researchers	344	12,745.67	14,257.49	251	61800
$\widehat{\lambda}$	country-by-country OLS estimate for λ	344	0.038	0.845	-0.979	3.865
$\widehat{\chi}_1$	country-by-country OLS estimate for χ_1	344	0.263	0.237	0.061	1.332
$\widehat{\chi}_2$	country-by-country OLS estimate for χ_2	344	0.718	23.148	-9.271	160.182

Table A9: Descriptive statistics on 5-year average *per capita* RGDP growth rate

Country	Min	Mean	Max
Australia	0.942	2.526	3.560
Austria	1.579	3.250	5.114
Belgium	-0.938	2.617	5.633
Canada	-0.228	1.943	3.395
Denmark	1.485	2.590	3.888
France	-0.275	1.987	4.242
Germany	1.358	3.022	5.414
Hungary	1.084	2.949	5.015
Ireland	3.569	6.534	9.045
Italy	-0.338	2.401	4.938
Japan	-0.419	2.968	7.747
Poland	1.590	3.354	6.361
Portugal	0.580	4.015	7.325
Slovenia	2.046	2.719	3.894
Spain	-0.828	3.762	6.349
Sweden	-0.445	2.643	4.410
United Kingdom	1.311	2.767	4.955
United States	1.248	2.310	3.474

Table A10: Information criteria

	2 regimes	3 regimes	4 regimes	5 regimes
Log-likelihood	-566.69	-501.56	-463.24	-439.53
AIC	1155.4	1043.1	988.48	967.07
BIC	1197.6	1119.9	1107.5	1136.1

Table A11: Robustness check – growth regression, using equation (17) to get $\widehat{\chi}_1, \widehat{\chi}_2$ and $\widehat{\lambda}$

	Constant	$\gamma_\mathbf{N}$	$\mathbf{m} \times \gamma_\mathbf{N}$	Cluster standard deviation	5-year avg. *per capita* RGDP growth rate (%)
OLS FE	0.031***	1.025***	-0.561***		
HMM:					
1–Slow growth	1.060***	0.925	-0.886	0.741	0.994
2–Steady growth	2.608***	0.256	-0.257	0.493	2.654
3–Sustained growth	4.095***	0.942	-0.454	0.590	3.826
4–Miracle growth	5.651***	1.467**	5.106**	1.091	6.371

Significance levels: : 10% **: 5% ***: 1%.*

Table A12: Estimated intermediate sector's markup

Country	$\tilde{m}$
Australia	1.665
Austria	1.510
Belgium	1.650
Canada	1.755
Denmark	1.572
France	1.569
Germany	1.665
Hungary	1620
Ireland	1.518
Italy	1.599
Japan	1.540
Poland	1.596
Portugal	1.683
Slovenia	1.906
Spain	1.608
Sweden	1.538
United Kingdom	1.686
United States	1.713
mean	1.635

Note: all parameters are significant at 1%.

RECENT PUBLICATIONS BY *CEIS Tor Vergata*

Nowcasting Monthly GDP with Big Data: a Model Averaging Approach
Tommaso Proietti and Alessandro Giovannelli
CEIS Research Paper, 482, May 2020

Who did it? A European Detective Story Was it Real, Financial, Monetary and/or Institutional: Tracking Growth in the Euro Area with an Atheoretical Tool
Mariarosaria Comunale and Francesco Paolo Mongelli
CEIS Research Paper, 481, May 2020

On the Effects of Taxation on Growth: an Empirical Assessment
Marco Alfò, Lorenzo Carbonari and Giovanni Trovato
CEIS Research Paper, 480, May 2020

Optimal Correction of the Public Debt and Fiscal Resilience Measures
Barbara Annicchiarico, Fabio Di Dio and Stefano Patrì
CEIS Research Paper, 479, April 2020

The Legacy of Historical Emigration: Evidence from Italian Municipalities
Erminia Florio
CEIS Research Paper, 478, December 2019

Likelihood Induced by Moment Functions Using Particle Filter: a Comparison of Particle GMM and Standard MCMC Methods
Fabio Franco
CEIS Research Paper, 477, December 2019

How Does the Public Spending Affect Technical Efficiency? Some Evidence from 15 European Countries
Sabrina Auci, Laura Castellucci and Manuela Coromaldi
CEIS Research Paper, 476, December 2019

The Relative Price of Housing and Subsequent GDP Growth in the USA
Robert J. Waldmann
CEIS Research Paper, 475, November 2019

Identify More, Observe Less: Mediation Analysis Synthetic Control
Giovanni Mellace and Alessandra Pasquini
CEIS Research Paper, 474, November 2019

Efficient Particle MCMC with GMM likelihood representation
Fabio Franco
CEIS Research Paper, 473, November 2019