Electronic Journal of Statistics
Vol. 10 (2016) 2179–2203
ISSN: 1935-7524
DOI: 10.1214/16-EJS1165

Innovation, growth and aggregate volatility from a Bayesian nonparametric perspective

Antonio Lijoi[*†]

Department of Economics and Management, University of Pavia
via San Felice 5, 27100 Pavia, Italy
e-mail: lijoi@unipv.it

Pietro Muliere

Department of Decision Sciences, University "L. Bocconi"
Via Roentgen 1, 20136 Milano, Italy
e-mail: pietro.muliere@unibocconi.it

Igor Prünster[†]

Department of Decision Sciences, BIDSA and IGIER, University "L. Bocconi"
Via Roentgen 1, 20136 Milano, Italy
e-mail: igor@unibocconi.it

and

Filippo Taddei[*]

School of Advanced International Studies (SAIS), The Johns Hopkins University
Via Belmeloro 11, 40126 Bologna, Italy
e-mail: ftaddei@jhu.edu

Abstract: In this paper we consider the problem of uncertainty related to growth through innovations. We study a stylized, although rich, growth model, in which the stochastic innovations follow a Bayesian nonparametric model, and provide the full taxonomy of the asymptotic equilibria. In most cases the variability around the average aggregate behaviour does not vanish asymptotically: this requires to accompany usual macroeconomic mean predictions with some measure of uncertainty, which is readily yielded by the adopted Bayesian nonparametric approach. Moreover, we discover that the extent of the asymptotic variability is the result of the interaction between the rate at which the economy creates new sectors and the concavity of returns in sector specific technologies.

MSC 2010 subject classifications: Primary 62F15, 60G57, 91B62.
Keywords and phrases: Bayesian nonparametrics, aggregate volatility, asymptotics, economic growth, Poisson-Dirichlet process.

Received November 2015.

[*]Also affiliated to Collegio Carlo Alberto, via Real Collegio 30, 10024 Moncalieri, Italy.
[†]Supported by the European Research Council (ERC) through StG "N-BNP" 306406.

1. Introduction

The massive literature regarding economic growth spurred in modern times from the Solow model [38] where one of the basic assumptions is that the economy is characterized by homogenous individuals (or heterogeneous individuals who behave on average as the representative agent) and features an exogenous technological structure. On the other hand, the seminal contributions of [2] and [36] provided a novel focus on the interaction between capital accumulation and innovation in delivering economic growth. An extensive review of the numerous contributions that followed can be found in [1, Chapters 12 & 13]. Most of the subsequent research in the area has tried to overcome the assumption of agents' homogeneity and relies on models that endogenize the development of technological progress. These assume that, when the number of innovations grows asymptotically large, one could expect that the aggregate output variability vanishes. Hence, the dynamics of the economy can be summarized by the average of some endogenous variable such as, e.g., number of sectors, variety of intermediate goods or capital accumulation. An additional stimulating account on these models can be found in [15]. The present contribution complements this literature and introduces a model whereby, as the economy develops, innovations are distributed within and between existing and new productive sectors and the output volatility heavily depends on such a distribution. This is therefore entirely novel and interesting in its own right as it leads one to identify situations where the averages do not effectively account for the dynamics of the economy.

In this paper we study economic growth by adopting a Bayesian nonparametric viewpoint, which naturally allows for a rich probabilistic modeling and for flexible estimation procedures via conditional (or posterior) distributions. For our purposes, one of the main advantages of this approach is represented by the fact that it yields intuitive and coherent prediction mechanisms for modeling unseen species, categories, genes but also types of economic agents, sectors, products, technologies and, more generally, for the analysis of the impact of heterogeneity in economic models. An example, which is coherent with the general framework of the paper, can be described in terms of an economy in which, at a certain point in time, there exist k distinct sectors and $n \geq k$ "innovations" have occurred. Of these innovations, n_1 have taken place in the first sector, n_2 in the second etc., so that $\sum_{i=1}^{k} n_i = n$. Then, conditional on the configuration of the economy as determined by the first n innovations, the prediction rule (associated to a discrete nonparametric prior) is such that with positive probability the $(n+1)$–th innovation might happen not only in any of the already existing k sectors, but crucially also in a new sector. The latter event, of positive probability, corresponds to creating a new sector. This possibility of naturally incorporating the unseen is certainly one of the main reasons for the recent success of Bayesian Nonparametrics. See [22] for a review of the discipline and for references on applications to, e.g., biology, computer science, engineering, language models, machine learning, medicine, physics. In contrast, Bayesian nonparametric methods have not yet been extensively exploited for economic

applications. Among the contributions to date we mention [31, 24, 20] for financial time series, [19, 17, 29, 30] for volatility estimation, [23] for option pricing, [5, 9] for discrete choice models, [18] for stochastic frontier models, [35, 21] for market share and income dynamics. Emphasis in most of these papers is on estimation, whereas we focus on modeling and perform a theoretical analysis of the implications of realistic assumptions on the distribution of the innovations on economic growth. In general, we will try to highlight how Bayesian nonparametric models allow to treat heterogeneity in macroeconomic models delivering accurate assessments of the uncertainty related to prediction.

We present a relatively simple model of endogenous growth, which follows the intuition of [4]: the economy is made of different sectors, each endowed with its own production technology. All sectors produce an homogenous output good and the economy grows by innovations. These are stochastic events of two types: the first type is a productivity rise in an existing sector, whereas the second type is the creation of a new sector of production. Given the economic problem at hand, we assume that the pace of innovations is regulated by a two-parameter Poisson-Dirichlet process, one of the most popular nonparametric priors introduced in [32]. For the sake of realism we only assume that innovations in a given sector are more productive, the larger is the number of innovations that took place in that sector and the smaller is the set of innovations in the economy as a whole. With this setup in mind, we study the full taxonomy of the equilibria of this stylized, although rich, endogenous growth model.

The economic implications of the analytical results of the paper can be essentially summarized as follows. If the distribution of the innovations is such that:

(i) at any point in time a new sector is created with a positive probability, whose value depends on the number of existing sectors;

(ii) given any two existing sectors having experienced n_i and n_j innovations, with $n_i > n_j$, the probability of an innovation occurring in sector i is more than proportional to n_i/n_j;

then the overall output of the economy cannot be studied and therefore predicted by looking solely at the average behaviour. The variability around the mean needs to be taken into account and the Bayesian approach readily yields a tool for its quantification. Moreover, it turns out that the order of magnitude of the uncertainty about the mean depends on the structure of sector specific technologies precisely in the way one would expect: if returns in sector specific technologies are sufficiently concave, then innovations are more beneficial when they create new sectors, which in turn implies a more variable output given the contribution to the aggregate output of new sectors is more difficult to predict.

From a technical point of view, we will resort to the self–averaging condition, a popular concept in Physics and other disciplines first used in Economics by [3], to identify situations in which asymptotically the mean resembles the random phenomenon at issue. Loosely speaking, a random variable—innovations in our setup—is self–averaging if it tends to cluster around its mean as the number of observations grows. Therefore, variables that are non–self–averaging can be hardly summarized by their mean and models that involve them should consider the variability as much as the mean to predict dynamics. Assumptions (i) and

(ii) above necessarily imply that the number of sectors is non–self–averaging. The same can be said for other quantities related to the innovation distribution and such a behaviour is shown to carry over to the aggregate output under the hypothesized growth model. The extent of the phenomenon is shown to explicitly depend on the relation between two fundamental parameters of the model, the first controlling the rate at which new sectors are created and the second tuning the concavity of returns in sector technologies. If the latter "dominates", the aggregate output is highly non–self–averaging. It is worth noting that our work, as far as its starting point is considered, is in the spirit of [4]. However, our model, methodology and conclusions considerably differ. The perspective of this paper not only does allow to properly identify the determinants of non–self–averaging behaviours within endogenous growth models but also, and most importantly, it provides the appropriate countermeasures to face such phenomena.

An outline of the paper is as follows. Section 2 provides the necessary background and setup. Section 3 introduces the models of innovations and endogenous growth, provides the asymptotic results and describes how to obtain uncertainty estimates and their concrete implementation. Section 4 contains two major concluding remarks on the extensions of the present study: the first, of theoretical nature, concerns the generalization of the results to innovation distributions belonging to an extremely wide class of priors, whereas the second, of applied nature, discusses extensions to more general growth models, which being more complex cannot be faced analytically but need to be tackled relying on simulation studies. The Appendix contains some additional technical material on Bayesian nonparametric models necessary for the understanding of the proofs, which are also deferred to the Appendix.

2. The underlying framework

In this section we recall the notions of exchangeability and nonparametric prior. We, then, introduce the two–parameter Poisson–Dirichlet process and illustrate a few of its properties that are used in the sequel.

2.1. Quick overview on Bayesian nonparametric modelling

To sum up the main theoretical framework, suppose $X^{(\infty)} = (X_n)_{n\geq 1}$ is a sequence of observations, defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with each X_i taking values in a complete and separable metric space $\mathbb{X}$ endowed with the Borel σ–algebra $\mathscr{X}$. Later, X_i will be interpreted as the label identifying the sector where the i–th innovation occurs, for any $i \geq 1$. The customary assumption in a Bayesian framework is *exchangeability* of the sequence $X^{(\infty)}$, which means that for any $n \geq 1$ and any permutation π of the indices $1, \ldots, n$, the probability distribution of the random vector $(X_1, \ldots, X_n)$ coincides with the distribution of $(X_{\pi(1)}, \ldots, X_{\pi(n)})$. The fundamental result concerning exchangeable sequences is known as de Finetti's representation theorem: it states that a sequence $X^{(\infty)}$ is exchangeable if and only if there exists a probability measure

Q on the space $\mathcal{P}_{\mathbb{X}}$ of all probability measures on $\mathbb{X}$ such that, for any $n \geq 1$ and $A = A_1 \times \cdots \times A_n \times \mathbb{X}^{\infty}$, one has

$$\mathbb{P}\left[X^{(\infty)} \in A\right] = \int_{\mathcal{P}_{\mathbb{X}}} \prod_{i=1}^{n} P(A_i)\, Q(\mathrm{d}P) \tag{1}$$

where $A_i \in \mathscr{X}$ for any $i = 1, \ldots, n$ and $\mathbb{X}^{\infty} = \mathbb{X} \times \mathbb{X} \times \cdots$. The probability Q is termed the *de Finetti measure* of the sequence $X^{(\infty)}$. In other terms, (1) states that, conditional on a random probability measure $\tilde{P}$ from Q, $X^{(\infty)}$ is a sequence of independent and identically distributed random elements with common probability distribution $\tilde{P}$, namely

$$\begin{aligned} X_i|\tilde{P} &\overset{\text{iid}}{\sim} \tilde{P} \qquad\qquad i \geq 1 \\ \tilde{P} &\sim Q. \end{aligned} \tag{2}$$

In a Bayesian setup Q represents the prior distribution and the model is termed *parametric*, whenever Q degenerates on a finite dimensional subspace of $\mathcal{P}_{\mathbb{X}}$, otherwise one has a *nonparametric* model. If $Q(\,\cdot\,|X_1, \ldots, X_n)$ denotes the posterior distribution of $\tilde{P}$ given the sample $X_1, \ldots, X_n$, prediction is then achieved by deriving explicit expressions for

$$\mathbb{P}[X_{n+1} \in \,\cdot\,|X_1, \ldots, X_n] = \int_{\mathcal{P}_{\mathbb{X}}} P(\,\cdot\,)\, Q(\mathrm{d}P|X_1, \ldots, X_n). \tag{3}$$

If $\mathscr{D} \subset \mathcal{P}_{\mathbb{X}}$ is the set of discrete probability distributions on $\mathbb{X}$, we will focus on nonparametric priors Q such that $Q(\mathscr{D}) = 1$: when this happens one typically refers to Q as a "discrete prior", although one has to keep in mind that the (weak) topological support of most such nonparametric Q is still the whole $\mathcal{P}_{\mathbb{X}}$. For an exchangeable sequence $X^{(\infty)}$ whose de Finetti measure Q is discrete, the sample $X_1, \ldots, X_n$ contains ties with positive probability, *i.e.* $\mathbb{P}[X_i = X_j] > 0$ for $i \neq j$. Hence, Q partitions the n data into a random number K_n of clusters with respective (random) frequencies $N_{1,n}, \ldots, N_{K_n,n}$ and such a partition turns out to be exchangeable in a sense that will be made clear in the Appendix. Even if the Bayesian nonparametric framework, as described through (1), had been laid out by de Finetti during the 30's, the definition of tractable priors Q on $\mathcal{P}_{\mathbb{X}}$ represented a challenging task, completed only 40 years later with the introduction of the Dirichlet process prior by T.S. Ferguson [14]. In the following section we introduce the two–parameter Poisson–Dirichlet process, due to J. Pitman [32], which represents one of the most popular priors and includes the Dirichlet process as a special case.

2.2. The two–parameter Poisson–Dirichlet process

The two–parameter Poisson–Dirichlet process is a random probability measure $\tilde{P}$, whose realizations are discrete probability distributions and therefore it can be represented as

$$\tilde{P}(\,\cdot\,) = \sum_{j \geq 1} \tilde{p}_j\, \delta_{Y_j}(\,\cdot\,) \tag{4}$$

where δ_a is the point mass at a, the $\tilde{p}_j$'s are random weights such that $\sum_{j\geq 1}\tilde{p}_j = 1$ (almost surely) and the Y_j's are random locations in $\mathbb{X}$. A simple and intuitive procedure for assigning the random weights $\tilde{p}_j$'s is given by the so–called *stick–breaking* construction. The rationale of the procedure is to obtain the random probability masses $\tilde{p}_j$ through a random partition of the interval $[0,1]$. To be more precise, let $(V_i)_{i\geq 1}$ be a sequence of independent random variables taking values in $[0,1]$. Starting from a unit length stick, break it into two bits of length V_1 and $1-V_1$. The first bit represents $\tilde{p}_1$ and in order to obtain $\tilde{p}_2$ it is enough to split the remaining part, of length $1-V_1$, into two parts having respective lengths $V_2(1-V_1)$ and $(1-V_2)(1-V_1)$. The former will coincide with $\tilde{p}_2$ and the latter will be split to generate $\tilde{p}_3$, and so on.

We are now in a position to state the definition of two–parameter Poisson–Dirichlet process.

Definition 1. *Let (α,θ) be parameters such that $\alpha \in [0,1]$ and $\theta > -\alpha$. Moreover, $(V_i)_{i\geq 1}$ is a sequence of independent random variables with $V_i \sim Beta(1-\alpha, \theta+i\alpha)$ and $(\tilde{p}_i)_{i\geq 1}$ are random weights defined as*

$$\tilde{p}_1 = V_1, \qquad \tilde{p}_i = V_i \prod_{j=1}^{i-1}(1-V_j) \quad i \geq 2.$$

If $(Y_i)_{i\geq 1}$ is a sequence of i.i.d. random variables with non–atomic probability distribution P_0, the random probability measure in (4) is a two-parameter Poisson–Dirichlet process, in symbols $PD(\alpha,\theta)$, and its law Q is two–parameter Poisson–Dirichlet process prior.

The Dirichlet process is then recovered as a particular case by setting $\alpha = 0$. The above definition presents only one of the possible constructions of this stochastic process, probably the most intuitive but, from an analytical point of view, not necessarily the most useful. See [34] for alternative constructions.

2.3. Self–averaging phenomena

As mentioned in the Introduction, we will investigate whether, in endogenous growth models, the deterministic mean behaviour of aggregate output asymptotically resembles its stochastic evolution, as the number of innovations increases. In this respect, it is useful to introduce the simple self–averaging condition, first used in Economics by [3], which allows to precisely identify situations in which the variability vanishes asymptotically.

Definition 2. *A sequence of size–dependent random variables $(Z_n)_{n\geq 1}$ is termed self–averaging if*

$$Var\left(\frac{Z_n}{\mathbb{E}(Z_n)}\right) \to 0 \qquad n \to \infty, \tag{5}$$

and non–self–averaging otherwise.

From (5) it becomes evident that for self–averaging macroeconomic phenomena, one can focus attention on the means of the involved variables since for sufficiently large n the residual variability of Z_n, normalized by its mean, becomes negligible. The self–averaging condition typically holds for simple economic models, where some assumption of symmetry or homogeneity of the individuals underlies the whole model. The concept is best clarified by looking at an example: consider the popular Poisson model, in which for each "individual" an event (e.g. technical progress) occurs according to a Poisson process with parameter λ. Then, in the whole economy, which is based on n individuals, the number of events Z_n follows a Poisson process with rate λn. Consequently, in a one–time period, we have $\mathbb{E}(Z_n) = \mathrm{Var}(Z_n) = \lambda n$ and one has $\mathrm{Var}(Z_n\,\mathbb{E}(Z_n)^{-1}) = (\lambda n)^{-1} \to 0$ as $n \to \infty$. Hence, the Poisson model is self–averaging. The same obviously holds for the Gaussian case. An equivalent formulation of (5) can be expressed in the terms of the coefficient of variation, C.V., as

$$\text{C.V.}(Z_n) = \frac{\sqrt{\mathrm{Var}(Z_n)}}{\mathbb{E}(Z_n)} \to 0 \quad \text{as } n \to \infty. \tag{6}$$

On the other hand, for *non–self–averaging* models, even if the number of agents diverges, the uncertainty about the "normalized" trajectories of Z_n persists. Therefore, focusing solely on the mean behaviour is not enough for describing the phenomenon at hand and some measure of the oscillations around the mean is essential for providing a clear picture. In what follows we introduce an endogenous growth model and show that it leads, under reasonable assumptions, to non–self–averaging phenomena. By deriving exact asymptotic results we show how the mean can be combined with suitable measures of uncertainty to fully describe the evolution of the economy.

3. Bayesian nonparametric analysis of economic growth

3.1. The distribution of the innovations

In line with a large share of the literature on endogenous growth (see, e.g., [1] and [15]), we assume that the economy grows by innovations, which are stochastic events of two types: the first type is represented by a productivity rise in an existing sector, whereas the second type is represented by the creation of a new sector. This scheme is reminiscent of the predictive structure yielded by a discrete random probability measure that governs an exchangeable sequence of observations (i.e. the innovations) and, hence, nicely connects with a rich body of literature in Bayesian Nonparametrics. We first describe the distribution of the innovations and then detail the complete growth model.

We will assume that the innovations, which occur in the economy, are governed by a two–parameter Poisson–Dirichlet model with parameters $\alpha \in (0, 1)$ and $\theta > 0$. Specifically, the sequence of exchangeable random variables $(X_i)_{i\geq 1}$ represent innovations and the value of each random variable identifies the sector in which the innovation takes place. Clearly, observing a new value of the X_i's,

means that a new sector has been created. Therefore, the sample $X_1, \ldots, X_n$ identifies the $K_n \leq n$ distinct sectors where the n innovations have occurred and the corresponding sector labels are denoted as $X_1^*, \ldots, X_{K_n}^*$. The j–th sector, X_j^*, will have experienced $N_{j,n}$ innovations, for $j = 1, \ldots, K_n$, and $\sum_{j=1}^{K_n} N_{j,n} = n$. Given the outcome of the first n innovations, it is natural to look at the $(n+1)$–th and provide a probabilistic assessment of the possible outcomes, namely the predictive distribution. In fact, the system of predictive distributions (3) associated with a PD(α, θ) process is of the form

$$\mathbb{P}[X_{n+1} \in \cdot \,|\, X_1, \ldots, X_n] = \frac{\theta + \alpha K_n}{\theta + n} P_0(\cdot) + \frac{1}{\theta + n} \sum_{j=1}^{K_n} (N_{j,n} - \alpha)\, \delta_{X_j^*}(\cdot), \quad (7)$$

which means that the $(n+1)$–th innovation will lead to the creation of a new sector with probability

$$\frac{\theta + \alpha K_n}{\theta + n}, \quad (8)$$

whereas it will occur in any of the pre–existing sectors with probability

$$\frac{n - \alpha K_n}{\theta + n}. \quad (9)$$

In particular, it will take place in the j–the sector with probability

$$\frac{N_{j,n} - \alpha}{\theta + n}, \quad (10)$$

for $j = 1, \ldots, K_n$. Therefore, assuming the distribution of the innovations is governed by a PD(α, θ), can be thought of as the innovations being sequentially generated by the predictive scheme in (7).

Before proceeding it is worthwhile to take a close look at the implications of the PD(α, θ) on the generation of the innovations and, thus, at the behaviour of the economy. First, from (8) it is apparent that the probability of creating a new sector is monotonically increasing in the number of sectors K_n created so far. In fact, it is natural to expect a higher probability of creating a new sector within a dynamic economy, i.e. an economy which exhibited a high sector creation rate in the past. The second implication is more subtle and concerns the probability, given its occurrence in an existing sector, that an innovation takes place in sector j rather than in another sector. In this respect, the PD(α, θ) process induces a reinforcement mechanism, which can be explained as follows. The probability of having an innovation in one of the already existing sectors is (9), but the mass is not allocated proportionally to the number of innovations already observed in each sector. The probability of observing an innovation in sector j is determined by the number of innovations $N_{j,n}$ in that sector and by the parameter α, which drives the reinforcement: one can see that the ratio of the probabilities assigned to any pair of sectors (i, j) is given by $(N_{i,n} - \alpha)/(N_{j,n} - \alpha)$. As $\alpha \to 0$, the previous quantity reduces to the ratio of the number of innovations in the sectors thus reducing to the case of homogeneity among sectors, namely the

probability of having an innovation in each sector is proportional to the number of innovations that have already occurred so far. On the other hand, if $\alpha > 0$ and $N_{i,n} > N_{j,n}$, the ratio is an increasing function of α. Hence, as α increases the probability of observing an innovation is reallocated from sector j to i. This means that the dynamics tends to reinforce, among the existing sectors, those having featured a higher number of innovations. In economic terms the reinforcement mechanism driven by α can be interpreted as a measure of the "intra-sectoral learning by doing" in innovations: if we increase the value of the α, innovations tend to appear in those sectors where many innovations already took place. Table 1 provides an idea of the magnitude of the reinforcement. Such a structural feature of the innovations distribution is clearly appealing since one would naturally expect that innovations tend to occur in the most dynamic sectors rather than in those in which only a few innovations have taken place. See [25, 26] for details and more discussion on the reinforcement connected to mixture models.

TABLE 1
Ratio of the probabilities allocated to sector i observed $N_{i,n}$ times and sector j observed only once for different choices of α.

	$n_i = 2$	$n_i = 5$	$n_i = 10$	$n_i = 100$
$\text{PD}(\alpha \to 0, \theta)$	2	5	10	100
$\text{PD}(\alpha = 0.25, \theta)$	2.33	6.33	13	133
$\text{PD}(\alpha = 0.50, \theta)$	3	9	19	199
$\text{PD}(\alpha = 0.75, \theta)$	5	17	37	397
$\text{PD}(\alpha \to 1, \theta)$	$\to \infty$	$\to \infty$	$\to \infty$	$\to \infty$

Summing up, the combined effect of the two above described features, underlying the $\text{PD}(\alpha, \theta)$ assumption, imply that in a dynamic economy a lot of new sectors will be created and the innovations within existing sectors will concentrate on a few of them, while the others will become less and less likely to experience an innovation. This seems quite realistic: as the economy grows, we observe it shifting its focus away from the sectors that become progressively less innovative and into those sectors where innovations come to stage more and more often. However, the probability of an innovation occurring in a neglected sector will still be positive and it is enough that a few innovations happen in it, for starting the reinforcement mechanism, which then could completely "revive" it. On the other extreme an uninventive economy will create very few new sectors and the distribution of the innovations will typically be much more balanced. As explained above it is the parameter α, which tunes the extent of such phenomena and in section 3.4 it is shown how to endogenously estimate it.

An additional feature of the model worth highlighting is the asymptotic behaviour of the number of distinct sectors K_n as the number of innovations n grows. To this end, it is useful to first introduce a class of random variables, which will appear throughout the following developments. This class of random variables, termed generalized Mittag–Leffler random variables, is defined as follows. Let f_α be the density function of a positive α–stable random variable and define S_q to be, for any $q \geq 0$, a positive random variable with density function

one the positive real line

$$h_{S_q}(s) = \frac{\Gamma(q\alpha+1)}{\alpha\,\Gamma(q+1)}\, s^{q-1-1/\alpha}\, f_\alpha\left(s^{-1/\alpha}\right). \tag{11}$$

Then, [34, Theorem 3.8] states that

$$\frac{K_n}{n^\alpha} \xrightarrow{\text{a.s.}} S_{\theta/\alpha}. \tag{12}$$

Therefore, one has that K_n increases at a rate of n^α and, moreover, the normalized version of K_n converges to a strictly positive random variable. In other words, the sequence $(K_n)_{n\geq 1}$ is non–self–averaging, a fact first pointed out in [3].

3.2. The endogenous growth model

We now introduce the mechanism according to which the sectors are aggregated in order to produce the overall output, thus completing the description of our growth model. Recall that, by the time the n–th innovation occured, the economy will consist of a random number K_n of sectors, the i–th sector will have experienced $N_{i,n}$ innovations and obviously $\sum_{i=1}^{K_n} N_{i,n} = n$. For the sake of realism, we will postulate that the more productive are the innovations in a given sector, the larger is the number of innovations that took place in that sector and the smaller is the set of innovations in the economy as a whole. This is reflected in the assumption that the output of sector i can be represented as

$$Y_{i,n} = \gamma^{N_{i,n}/n^{1-\tau}} \tag{13}$$

where $\gamma > 1$ and $\tau \in (0,1)$. This is a richer functional form than what is typically employed in endogenous growth models (see [2] and [36]). It actually has the distinctive advantage to allow both the analysis of the effect of an innovation in an existing sector and the assessment of the importance of the sector itself in the overall economy and its dynamics. This is an important innovation capable of expanding the scope of analysis in comparison with the existing benchmarks. It is worth noting that the parameter τ tunes the concavity of the returns in sector specific technologies and, in conjunction with the parameter α of the PD(α,θ) process, will play a crucial role in quantifying the contribution of the existing sectors w.r.t. new sectors to the aggregate output of the economy. τ can be interpreted as the severity of diminishing returns in the total number of innovations in the economy n. The larger is τ, the more beneficial is an additional innovation in the existing sector i. This feature will become apparent in the presentation of the asymptotic results that follow in Section 3.3. Moreover, we will concentrate our attention on the case of γ close to 1, which is realistic in many situations. Therefore we can approximate (13) with

$$Y_{i,n} \approx 1 + \beta\,\frac{N_{i,n}}{n^{1-\tau}}$$

where $\beta = \log(\gamma) > 0$. Hence, the aggregate output of the economy after n innovations, which is the sum of the outputs of the K_n sectors, is

$$Z_n = \sum_{i=1}^{K_n} Y_{i,n} \approx K_n + \beta\, n^{\tau} \tag{14}$$

which shows that K_n is the contribution to the aggregate output of the number of sectors and that n^{τ} is the contribution of the innovations within sectors.

It is worth remarking that (14) allows us to develop a sound and intuitive analysis of the the asymptotic behaviour of the aggregate output Z_n, as the number of innovations, n, grows. This is detailed in the next section and is in the spirit of [4]. However, in that paper the authors assume a model for which the aggregate output turns out to be $Z_n = K_n + \beta n$. Unfortunately, such a formulation is misleading since it does not generate the non self–averaging behaviour of Z_n that the authors in [4] were willing to point out. In contrast, the additional flexibility gained by the single sector output specification, (13), allows us to cover both self–averaging and non self–averaging behaviours of the aggregate output.

3.3. Asymptotic analysis of the growth model

3.3.1. The unconditional case

We start our asymptotic analysis by considering a simplistic scenario, which is however useful for the understanding of the following main result. Here the economy is assumed to start from scratch and it is investigated how aggregate output evolves over time. Specific features of the economy and, in particular its history up to present, which one would obviously like to incorporate into the model, are ignored for the moment. Since in the following we will study the model from a Bayesian nonparametric viewpoint, i.e. conditional on the data, we will refer to this situation as the "unconditional" case.

The following result provides a complete taxonomy of the different asymptotic regimes arising in this unconditional setup. The various possible cases correspond to different choices of the parameters τ and α of the model. In particular, $\alpha > \tau$ means that the contribution to aggregate output from innovations represented by the creation of new sectors are more beneficial than those within an existing sector. This can either be due to the fact that the returns of sector specific technologies are highly concave, which would imply a low value for τ, or due to a particularly high rate of creation of new sectors, which would lead to a large value for α. If $\alpha < \tau$ one has exactly the opposite situation. Finally, the case in which $\alpha = \tau$ defines an economy where concavity of productivity rises in existing sectors and ability to create new sectors balance each other out.

The next result provides the mean of the aggregate output and quantifies the oscillations around this trend. It states that, when contributions to the economy given by the introduction of new sectors are at least as relevant as those given

by the existing sectors, the oscillations around the trend do not vanish asymptotically. This means that economy presents a non–self–averaging behaviour. On the other hand, when new sectors are less relevant in terms of contributions to the aggregate output, the economy can be exhaustively described by its mean $\mathbb{E}[Z_n]$, which is in agreement with the usual macroeconomic attitude to consider aggregate average quantities.

Proposition 1. *Under the growth model* (14) *with innovations following a two parameter Poisson Dirichlet process, we have*

$$\mathbb{E}[Z_n] = \frac{(\theta+\alpha)_n}{\alpha\,(\theta+1)_{n-1}} - \frac{\theta}{\alpha} + \beta n^{\tau}, \tag{15}$$

where $(a)_n = a(a+1)\dots(a+n-1)$ *is the* n*–th ascending factorial of* a*, with* $(a)_0 \equiv 1$*. Moreover,*

(i) If $\alpha = \tau = \upsilon$*,*

$$\frac{Z_n}{n^{\upsilon}} \to S_{\theta/\alpha} + \beta \quad a.s.$$

where S_q *is a generalized Mittag–Leffler random variable defined in* (11)*, and* Z_n *is non–self-averaging.*

(ii) If $\alpha = \upsilon > \tau$*,*

$$\frac{Z_n}{n^{\upsilon}} \to S_{\theta/\alpha} \quad a.s.$$

where S_q *is a generalized Mittag–Leffler random variable defined in* (11)*, and* Z_n *is non–self-averaging.*

(iii) If $\tau = \upsilon > \alpha$*,*

$$\frac{Z_n}{n^{\upsilon}} \to \beta \quad a.s.$$

and Z_n *is self-averaging.*

The main question arising from Proposition 1 is, what one should do in non–self–averaging cases, which have been shown to arise systematically in presence of highly dynamic economies. A drastic answer could be that of rejecting completely the typical macroeconomic approach which relies on the analysis of aggregate average quantities. However, we do not believe this to be the correct way of proceeding. Instead, we propose to combine the study of the mean behaviour with a measure of uncertainty and the natural tool in this framework is represented by the asymptotic highest posterior density (HPD) intervals of the limiting random variable, which represent the Bayesian counterpart to frequentist confidence intervals. In this way one can accompany the mean behaviour with an appropriate uncertainty estimate. In Section 3.4 we show how to derive such approximate HPD intervals.

Finally we note that a model leading to $Z_n = K_n + \beta n$, as in [4], is such that Z_n/n^{α} diverges and C.V.$(Z_n/n^{\alpha}) \to 0$, as $n \to \infty$. Hence, the process would actually be self–averaging. On the contrary, Proposition 1(i)–(ii) shows that our model (13) is well–grounded and flexible since it is able to capture also the important case of non–self–averaging behaviours of the aggregate output.

3.3.2. The conditional case

Let us now move from the simplistic model towards a more realistic one, which incorporates the status quo of the economy and studies the asymptotic behaviour of the aggregate output conditional on the state of the economy at any specific point in time. By "state of the economy" here we mean the number of sectors and the partition of the innovations among them. Given these data, we will be able to focus on the contribution to the aggregate output generated by sectors that will emerge only in the future. This is clearly the most interesting, and hardly predictable, scenario. From a mathematical point of view this corresponds to predicting the future behaviour conditionally on a given state of the world $X_1, \ldots, X_n$. Therefore, we now assume the *status quo* as given (i.e. $X_1, \ldots, X_n$ have generated $K_n = j$ sectors with respective frequencies of innovations $(N_{1,n}, \ldots, N_{j,n}) = (n_1, \ldots, n_j)$) and investigate the aggregate output of new sectors that will be determined by future innovations. By the time the m–th innovation occurs, there will be a random number $K_m^{(n)} = K_m - K_n$ of new sectors in the economy, where the i–th will have experienced S_i innovations. In this model, not all innovations will occur within new sectors: indeed, $\sum_{i=1}^{K_m^{(n)}} S_i = L_m^{(n)}$ represents the number of innovations, which take place in the new sectors and $m - L_m^{(n)}$ are the innovations taking place the "old" sectors, i.e. the sectors existing already at present. Therefore, the output of the i–th new sector is of the form

$$Y_{i,m}^* = 1 + \beta \frac{S_i}{m^{1-\tau}} \qquad i = 1, \ldots, K_m^{(n)}$$

where $\beta > 0$ and $\tau \in (0, 1)$. The conditioning sample enters the previous definition through $K_m^{(n)}$. The aggregate output of the $K_m^{(n)}$ new sectors is then given by

$$Z_m^* = K_m^{(n)} + \beta \frac{L_m^{(n)}}{m^{1-\tau}} \tag{16}$$

Proceeding along the same lines as in Section 3.3.1, the stochastic innovations are governed by a $\mathrm{PD}(\alpha, \theta)$ process with parameters $\alpha \in (0, 1)$ and $\theta > 0$. With this setup in mind, we study the full taxonomy of the equilibria of this stylized, although rich, endogenous growth model. In particular, the following main result provides the mean of the conditional aggregate output of the new sectors and shows that non–self–averaging appears under any assumption on the innovation parameter α and on the structural parameter τ. Recall, further, that $X_1, \ldots, X_n$ describes the state of the world after n innovations, which generated $K_n = j$ sectors and innovation frequencies $(N_{1,n}, \ldots, N_{j,n}) = (n_1, \ldots, n_j)$.

Proposition 2. *Under the growth model* (16) *with innovations following a two parameter Poisson Dirichlet process, we have*

$$\mathbb{E}[Z_m^* \mid X_1, \ldots, X_n] = \left(j + \frac{\theta}{\alpha}\right) \left\{ \frac{(\theta + n + \alpha)_m}{(\theta + n)_m} - 1 \right\} + \beta \frac{\theta + j\alpha}{\theta + n} m^{\tau}. \tag{17}$$

Moreover, as $m \to \infty$ the following holds true:

(i) *If* $\alpha = \tau = \upsilon$,

$$\frac{(Z_m^* | X_1, \ldots, X_n)}{m^\upsilon} \to U_{n,j} + \beta\, B_{\theta+\alpha j, n-\alpha j} \quad a.s.$$

where $U_{n,j} \stackrel{d}{=} B_{j+\theta/\alpha,\, n/\alpha-j}\, S_{(\theta+n)/\alpha}$, S_q *is a generalized Mittag–Leffler random variable with density* (11), $B_{a,b}$ *is a beta random variable with parameters* (a, b) *and the random variables* $B_{j+\theta/\alpha,\, n/\alpha-j}$ *and* $S_{(\theta+n)/\alpha}$ *are independent. Hence, the model is non–self-averaging.*

(ii) *If* $\alpha = \upsilon > \tau$,

$$\frac{(Z_m^* | X_1, \ldots, X_n)}{m^\upsilon} \to U_{n,j} \quad a.s.$$

and the model is non–self-averaging.

(iii) *If* $\tau = \upsilon > \alpha$,

$$\frac{(Z_m^* | X_1, \ldots, X_n)}{m^\upsilon} \to \beta\, B_{\theta+\alpha j, n-\alpha j} \quad a.s.$$

and the model is non–self-averaging.

Some comments are in order at this point. The previous result shows that, by complicating the model so to adhere more closely to realistic assumptions, non–self–averaging behaviours appear even more frequently. The various cases are not anymore distinguished by the fact of being or not self–averaging phenomena. All of them are non–self–averaging and they differ only in the order of magnitude of the non–self–averaging behaviour. Specifically, when the creation of new sectors is at least as relevant ($\alpha \geq \tau$) to the economy than rises of productivity within already created sectors, asymptotically the oscillations around the aggregate trend are dictated by a random variable taking values in $\mathbb{R}^+$. In the opposite case ($\alpha < \tau$), the limiting random variable takes on values in $(0, 1)$, which means that the oscillations are remarkably more mitigated but nonetheless present. If we keep the economic interpretation in mind, we conclude that the asymptotic degree of aggregate volatility depends on the speed at which new sectors are created in comparison with the degree of the diminishing returns present in existing sectors. A "more" dynamic economy – in the sense of new sectors creation – is also more likely to display higher volatility. This represents a clear sign that one cannot confine herself to studying mean behaviours but has to take the associated variability into account. The natural solution to this issue is to associate asymptotic HPD intervals, as measure of uncertainty, to the mean quantities. Therefore, the indication which clearly emerges from our analysis is that the usual way of proceeding in macroeconomics is legitimate as long as it is combined with suitable measures of uncertainty.

3.4. Asymptotic HPD intervals and parameter estimation

In this section we show how one can associate HPD intervals to the mean predictions to measure uncertainty. This is quite straightforward and makes use of

the rich probabilistic structure underlying Bayesian nonparametric models. The method is illustrated on this specific growth model but the idea is in principle applicable to any type of macroeconomic aggregation procedure with well–defined stochastic components and concretely applicable in cases, where one is able to work out the asymptotic regime.

Suppose to have an asymptotic result of the type $W_n/r(n) \to V$ as n diverges, where r is a suitable function of n. The determination of the asymptotic HPD intervals is as follows: take the x–order HPD interval (v_1, v_2) of V i.e. (v_1, v_2) such that $v_2 - v_1$ is minimal under the condition $\mathbb{P}(v_1 < V < v_2) \geq x$; then the x–order asymptotic HPD interval for W_n, for any finite sample size n, is given by $(v_1 r(n), v_2 r(n))$.

Turning back to our framework, we show how to concretely determine the asymptotic HPD intervals for Z_n in our growth models. By applying the above procedure, their derivation is straightforward except for the determination of the quantiles of a generalized Mittag–Leffler random variable S_q or of some transformation of it. From an analytical point of view the task seems overwhelming and therefore we show how one can generate random variates from S_q by adapting arguments in [12]: the output can then be used for evaluating quantiles of S_q as well as of its transformation like those appearing in Proposition 2. The basic idea consists in setting $W_q = S_q^{-1/\alpha}$ so that W_q has density function given by

$$f(w) = \frac{\alpha\Gamma(q\alpha)}{\Gamma(q)} w^{-q\alpha} f_\alpha(w) = \frac{\alpha}{\Gamma(q)} f_\alpha(w) \int_0^\infty u^{q\alpha-1} \mathrm{e}^{-uw} \, \mathrm{d}u$$

Via augmentation, one then has

$$f(u, w) = \frac{\alpha}{\Gamma(q)} f_\alpha(w) \, u^{q\alpha-1} \mathrm{e}^{-uw} = f(u) f_\alpha(w|u)$$

where $f(u)$ is the density function of a r.v. U_q such that $U_q^\alpha \sim$Gamma$(q, 1)$, and

$$f_\alpha(w|u) = f_\alpha(w) \, \mathrm{e}^{-uw+u^\alpha}.$$

This means that, conditional on U_q, W_q is a positive tempered–stable random variable, according to the terminology adopted in [37]. In order to draw samples from it, a convenient strategy is to resort to the series representation derived in [37], which, in our case, yields

$$W_q|U_q \stackrel{d}{=} \sum_{i=1}^{\infty} \min\left\{ (a_i \Gamma(1-\alpha))^{-1/\alpha}, e_i \, v_i^{1/\alpha} \right\} \tag{18}$$

where $e_i \stackrel{\text{iid}}{\sim} \text{Exp}(U_q)$, $v_i \stackrel{\text{iid}}{\sim} \text{U}(0, 1)$ and $a_1 > a_2 > \cdots$ are the arrival times of a Poisson process with unit intensity. An efficient and clever alternative for generating W_q, conditionally on U_q, is an exact sampler for exponentially tilted positive stable random variables derived in [10]. Other possibilities are the inverse Lévy measure method as described in [13] and a compound Poisson approximation scheme proposed in [7].

With such algorithm at hand, it is straightforward to describe the growth model via $\mathbb{E}(Z_n)$ combined with the corresponding HPD intervals, which account for the persisting uncertainty due to the non–self–averaging nature of the phenomenon at issue.

The theoretical asymptotic study clearly represents the main contribution of the paper. However, it may be some interest also to derive concrete predictions based on the studied growth model. To this end only one ingredient is missing, namely a method for specifying the endogenous parameters (α, θ), since the parameter τ has to be provided exogenously by an expert. To this end we first recall that to a (single) sample $(X_1, \ldots, X_n)$ there corresponds a certain partition (into $K_n = j$ sectors with frequencies of innovations $(N_{1,n}, \ldots, N_{K_n,n}) = (n_1, \ldots, n_j)$) such that its probability equals

$$\frac{\prod_{i=1}^{j-1}(\theta + i\alpha)}{(\theta+1)_{n-1}} \prod_{r=1}^{j}(1-\alpha)_{n_r-1} \tag{19}$$

under a PD(α, θ) process. See [32]. This partition distribution represents one of the most important instances of exchangeable partition probability distribution. See the Appendix for details. Having (19) at hand, the most natural way to specify the parameters (α, θ) consists in adopting an empirical Bayes procedure, which suggests to fix (α, θ) so to maximize (19) corresponding to the observed sample $(j, n_1, \ldots, n_j)$. In other words (α, θ) result from

$$(\hat{\alpha}, \hat{\theta}) = \arg\max_{(\alpha,\theta)} \frac{\prod_{i=1}^{j-1}(\theta + i\alpha)}{(\theta+1)_{n-1}} \prod_{r=1}^{j}(1-\alpha)_{n_r-1}. \tag{20}$$

To fix ideas consider a dynamic economy in which $n = 100$ innovations featuring $k = 50$ sectors with innovation frequencies $n_1 = 20, n_2 = 16, n_3 = n_4 = n_5 = n_6 = 5, n_7 = \ldots = n_{50} = 1$. The empirical Bayes estimate resulting from (20) would then be $(\hat{\alpha}_1, \hat{\theta}_1) = (0.86, 0.1)$. It is worth noting that the more dynamic the economy, i.e. the more new sectors it is creating, the larger $\hat{\alpha}$ will be: this is apparent from (8), which is monotonically increasing in α.

4. Concluding remarks

In this paper we have analyzed in some details the uncertainty related to growth models with stochastic innovations. We have shown that the variability associated to the average behaviour of the economy cannot be neglected and, in general, it depends on the rate at which innovations take place in new productive sectors. Technically, it can be effectively quantified by means of highest posterior density intervals. The envisaged future developments of the present contributions are both theoretical and applied.

From a theoretical point of view it is important to extend the findings to innovations distributions regulated by other nonparametric priors. For most applications the two parameter Poisson–Dirichlet process is already rich and

flexible enough in terms of both topological support and prediction structure. Nonetheless, from a theoretical viewpoint, one would like to confirm that non–self–averaging behaviours are actually even more wide–spread phenomena. In fact, it would be very interesting to ascertain whether Propositions 1 and 2 also hold, up to suitable transformations of the limiting random variable, for the wider class of Gibbs–type priors introduced in [16], which includes the two parameter Poisson–Dirichlet process as a special case. See [8] for a review of Gibbs–type priors from a Bayesian perspective.

From the applied economic point of view, we set the stage to allow the endogenous growth model to include additional variables of interest and this enables prediction of more complex growth behaviour of the economy and it reduces the distance between the model and reality. Such a task will necessarily have to be simulation based since an analytical asymptotic study beyond the present growth model setting appears to be overwhelming, especially in the conditional case. However, taking roots in the analytical results of the present paper, it is worth continuing this research line also along a more applied path: given any growth model depending on K_n productive sectors and other variables, it should actually be quite straightforward to devise a simulation algorithm which allows to investigate the asymptotic regime inherited by the overall economy in terms of both aggregate and sector–specific growth and volatility. Consistently with the proposed perspective, we also believe that the non–self–averaging nature of K_n should carry over to the dynamic behavior of the economy as a whole. Therefore, the indication would be once again to combine the analysis of the average aggregate behaviour with uncertainty quantification by means of highest posterior density intervals.

Appendix A

A.1 Exchangeable random partitions

Here we briefly recall some basic concepts on exchangeable random partitions and the corresponding notations to be used in the following Proofs. As mentioned in Section 2.1, if Q in (1) puts positive mass on a set of elements $\mathscr{D}$ in $\mathcal{P}_{\mathbb{X}}$ that are discrete, ties appear among $X_1, \ldots, X_n$ with positive probability, *i.e.* $\mathbb{P}[X_i = X_j] > 0$ for $i \neq j$. Correspondingly, define Ψ_n to be a random partition of $\{1, \ldots, n\}$ such that any two integers i and j belong to the same set in Ψ_n if and only if $X_i = X_j$. Let $k \in \{1, \ldots, n\}$ and suppose $\{C_1, \ldots, C_k\}$ is a partition of $\{1, \ldots, n\}$ into k sets C_i. Hence, $\{C_1, \ldots, C_k\}$ is a possible realization of Ψ_n. A common and sensible specification for the probability distribution of Ψ_n consists in assuming that it depends on the frequencies of each set in the partition but not on the actual values of $X_1, \ldots, X_n$. In terms of the de Finetti measure in (2) this is implied by assuming $\tilde{P}$ to belong to the class of *species sampling models* [34] i.e. $\tilde{P} \stackrel{d}{=} \sum_{i \geq 1} \tilde{p}_i \delta_{Y_i}$ such that the random weights $\tilde{p}_i$'s are independent from the locations Y_i, which are i.i.d. from a non–atomic P_0. The two parameter Poisson–Dirichlet process and the general class of Gibbs–type

priors are species sampling models. Indeed, this framework perfectly fits to our purposes since we are not much interested in the labels X_i identifying sectors where innovations take place. Our focus, instead, is on the probability that an economy experiencing n innovations will end up consisting of k distinct sectors with the innovations themselves being distributed among the k sectors according to the frequencies $(n_1, \ldots, n_k)$. In particular, let $1 \leq k \leq n$ and introduce the set

$$\Delta_{n,k} := \left\{ (n_1, \ldots, n_k) : \ n_i \geq 1, \sum_{i=1}^{k} n_i = n \right\}.$$

where each element $(n_1, \ldots, n_k)$ in $\Delta_{n,k}$ is known as a *composition* of $[n]$. If $n_i = \mathrm{card}(C_i)$, then $(n_1, \ldots, n_k) \in \Delta_{n,k}$ and

$$\mathbb{P}[\Psi_n = \{C_1, \ldots, C_k\}] = \Pi_k^{(n)}(n_1, \ldots, n_k). \tag{A.1}$$

We are now in a position to recall the following important concept introduced in [32].

Definition 3. *Let $(X_n)_{n\geq 1}$ be an exchangeable species sampling sequence, namely a sequence for which (2) holds true, with $\tilde{P}$ a species sampling model. Then, $\{\Pi_k^{(n)} : 1 \leq k \leq n, \ n \geq 1\}$ with $\Pi_k^{(n)}$ defined in (A.1) is termed exchangeable partition probability function (EPPF).*

Note that an EPPF determines the distribution of a random partition of $\mathbb{N}$. From the above definition it follows that, for any $n \geq k \geq 1$ and any $(n_1, \ldots, n_k) \in \Delta_{n,k}$, $\Pi_k^{(n)}$ is a symmetric function of its arguments, namely

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \Pi_k^{(n)}(n_{\pi(1)}, \ldots, n_{\pi(k)})$$

for any permutation π of $(1, \ldots, k)$, and it satisfies the consistency property

$$\Pi_k^{(n)}(n_1, \ldots, n_k) = \Pi_{k+1}^{(n+1)}(n_1, \ldots, n_k, 1) + \sum_{j=1}^{k} \Pi_k^{(n+1)}(n_1, \ldots, n_j + 1, \ldots, n_k). \tag{A.2}$$

formalizing the fact that the partition of $X_1, \ldots, X_n$ can be recovered from the partition of $X_1, \ldots, X_{n+1}$ by dropping X_{n+1}. On the other hand, as shown in [32], every non–negative symmetric function satisfying (A.2) is the EPPF of some exchangeable sequence. See [32, 34] for a thorough and useful analysis of EPPFs.

As for the PD(α, θ) process, which is a main ingredient of our model, the EPPF $\Pi_k^{(n)}(n_1, \ldots, n_k)$ is given by (19). If one denotes by $m_j \geq 0$, $j = 1, \ldots, n$ the number of sets in the partition which contain j objects or, within our setup, the number of sectors where j innovations occurred out of the total number n. Then, by considering all possible samples generating a certain configuration $(m_1, \ldots, m_n)$, from (19) one immediately obtains *Pitman's sampling formula*, which is given by

$$\Pi_{k,n}^{*}(m_1, \ldots, m_n) = n! \, \frac{\prod_{i=1}^{k-1}(\theta + i\alpha)}{(\theta+1)_{n-1} \prod_{i=1}^{n} m_i!} \prod_{i=1}^{n} \left[\frac{(1-\alpha)_{i-1}}{i!} \right]^{m_i} \tag{A.3}$$

for any $n \geq 1$ and $m_1, \ldots, m_n$ such that $m_i \geq 0$, $\sum_{i=1}^n i\, m_i = n$ and $\sum_{i=1}^n m_i = k$. The above expression represents a two parameter generalization of the celebrated Ewens' sampling formula [11], which is a cornerstone of population genetics and can be recovered by letting $\alpha \to 0$ in (A.3). It is important to note that, given (19), the determination of the predictive distributions in (7) is straightforward since

$$\mathbb{P}[X_{n+1} = \text{ new sector label } \mid X_1, \ldots, X_n] = \frac{\Pi_{k+1}^{(n+1)}(n_1, \ldots, n_k, 1)}{\Pi_k^{(n)}(n_1, \ldots, n_k)}$$

$$\mathbb{P}[X_{n+1} = j\text{–th sector label } \mid X_1, \ldots, X_n] = \frac{\Pi_k^{(n+1)}(n_1, \ldots, n_j + 1, \ldots, n_k)}{\Pi_k^{(n)}(n_1, \ldots, n_k)}$$

From (19) one also obtains the distribution of the number of distinct sectors K_n yielded by n innovations that have occurred in the economy. In [16] it is shown that

$$\mathbb{P}[K_n = k] = \frac{\prod_{i=1}^{k-1}(\theta + i\alpha)}{\alpha^k\,(\theta+1)_{n-1}}\, \mathscr{C}(n, k; \alpha) \qquad k = 1, \ldots, n,$$

where

$$\mathscr{C}(n, k; \alpha) = \frac{1}{k!} \sum_{j=0}^{k} (-1)^j \binom{k}{j} (-j\alpha)_n$$

is a generalized factorial coefficient. See [6] for an exhaustive account on factorial coefficients.

Appendix B: Proofs

Proof of Proposition 1. The proof is straightforward. It follows by combining the moment formula (3.13) in [34], the asymptotics of K_n as recalled in (12) and standard limiting arguments. □

Proof of Proposition 2. We start by considering the limiting behaviour of $K_m^{(n)} := K_{n+m} - K_n$, which is one of the two components the aggregate output (16) is made of, as m increases. Clearly, $K_m^{(n)}$ represents the number of new sectors created as a result of innovations ranging from the $(n+1)$–th to the $(n+m)$–th. The proof strategy is as follows: we first mimic the arguments of [12] in order to establish that $K_m^{(n)}/m^\alpha$, conditional on a sample $(X_1, \ldots, X_n)$, converges a.s. as the number of innovations m grows. Then we determine the moments of the limiting random variable and show that the limiting random variable is characterized by its moments. Finally, the asymptotic behaviour of the second component of the aggregate output is studied and the two bits combined to achieved the desired result.

First note that, as shown in [27, 28], K_n is a sufficient statistics for predictions over an additional sample $(X_{n+1}, \ldots, X_{n+m})$ for any $m \geq 1$. Hence,

in particular, $(K_m^{(n)}|X_1,\ldots,X_n)$ is equal in distribution to $(K_m^{(n)}|K_n)$ and the same holds for similar quantities considered in the sequel.

Let $P_{\alpha,\theta}^{(n)}$ be the conditional probability distribution of a PD(α,θ) process given K_n. Hence, $P_{\alpha,\theta}^{(n)}$ is absolutely continuous with respect to $P_{\alpha,0}^{(n)}$ on $\mathscr{F}_m^{(n)} := \sigma(X_{n+1},\ldots,X_{n+m})$, for any $m \geq 1$, and one can compute the likelihood ratio

$$M_{\alpha,\theta,m}^{(n)} := \frac{\mathrm{d}P_{\alpha,\theta}^{(n)}}{\mathrm{d}P_{\alpha,0}^{(n)}}\bigg|_{\mathscr{F}_m^{(n)}} = \frac{q_{\alpha,\theta}^{(n)}(K_m^{(n)})}{q_{\alpha,0}^{(n)}(K_m^{(n)})}$$

where, by virtue of [28, Proposition 1],

$$q_{\alpha,\theta}^{(n)}(k) = \frac{\alpha^k(\frac{\theta}{\alpha}+K_n)_k}{(\theta+n)_m}$$

for any integer $k\in\{1,\ldots,m\}$ and $q_{\alpha,\theta}^{(n)}(0):=1/(\theta+n)_m$. Hence $(M_{\alpha,\theta,m}^{(n)},\mathscr{F}_m^{(n)})_{m\geq 1}$ is a $P_{\alpha,0}^{(n)}$–martingale and, by a martingale convergence theorem,

$$P_{\alpha,0}^{(n)}\left[\lim_{m\to\infty} M_{\alpha,\theta,m}^{(n)} = M_{\alpha,\theta}^{(n)}\right] = 1$$

where $M_{\alpha,\theta}^{(n)}$ is an integrable random variable. One further has that $\mathscr{F}_m^{(n)} \uparrow \mathscr{F}_\infty^{(n)} = \sigma(X_{n+1},X_{n+2},\ldots)$, as $m\to\infty$, and $P_{\alpha,\theta}^{(n)}$ is still absolutely continuous with respect to (w.r.t.) $P_{\alpha,0}^{(n)}$ on $\mathscr{F}_\infty^{(n)}$. This implies that $M_{\alpha,\theta}^{(n)}$ is the Radon–Nikodým derivative of $P_{\alpha,\theta}^{(n)}$ w.r.t. $P_{\alpha,0}^{(n)}$ on $\mathscr{F}_\infty^{(n)}$ and, then, $\mathbb{E}_{\alpha,0}^{(n)}[M_{\alpha,\theta}^{(n)}] = 1$ where $\mathbb{E}_{\alpha,0}^{(n)}$ denotes the expected value w.r.t. $P_{\alpha,0}^{(n)}$. Use of the Stirling approximation $\Gamma(a+n)/\Gamma(b+n)\sim n^{a-b}$, as $n\to\infty$, easily leads to show that

$$M_{\alpha,\theta,m}^{(n)} \sim \frac{\Gamma(\theta+n)\Gamma(K_n)}{\Gamma(n)\Gamma\left(\frac{\theta}{\alpha}+K_n\right)}\left(\frac{K_m^{(n)}}{m^\alpha}\right)^{\theta/\alpha}$$

almost surely, as $m\to\infty$. Hence $(K_m^{(n)}/m^\alpha)^{\theta/\alpha}$ converges $P_{\alpha,0}^{(n)}$–a.s. to a random variable, say $U_{n,j}$ such that

$$\mathbb{E}_{\alpha,0}^{(n)}\left[U_{n,j}^{\theta/\alpha}\right] = \frac{\Gamma(n)\Gamma\left(\frac{\theta}{\alpha}+K_n\right)}{\Gamma(\theta+n)\Gamma(K_n)}.$$

In order to identify the distribution of the limiting random variable $U_{n,j}$ w.r.t. $P_{\alpha,\theta}^{(n)}$, we consider the asymptotic behaviour of $\mathbb{E}[(K_m^{(n)})^r \mid K_n]$ as $m\to\infty$, for any $r\geq 1$. According to the posterior characterization given in [33, Corollary 20], the PD(α,θ) process conditional on n innovations that have created j sectors, labeled as $X_1^*,\ldots,X_j^*$ and with innovation frequencies $(n_1,\ldots,n_j)$, coincides in distribution with the random probability measure

$$\sum_{i=1}^{j} w_i\delta_{X_i^*} + w_{j+1}\,\mathrm{PD}(\alpha,\theta+j\alpha)$$

where $w_{j+1} = 1 - \sum_{i=1}^{j} w_i$ and the random vector $(w_1, \ldots, w_j)$ has a j–variate Dirichlet distribution on the unit simplex with parameters $(n_1 - \alpha, \ldots, n_j - \alpha, \theta + j\alpha)$. It then turns out that if $w \sim \text{Beta}(\theta + j\sigma,\, n - j\sigma)$, one has

$$\mathbb{E}\left[(K_m^{(n)})^r \mid K_n = j,\, w\right] = \sum_{i=0}^{m} \binom{m}{i} w^i (1-w)^{m-i}\, \mathbb{E}\left[K_i^r\right] \tag{B.1}$$

where the unconditional moment $\mathbb{E}[K_i^r]$ is evaluated w.r.t. a $\text{PD}(\alpha, \theta + j\alpha)$ process. The expression (B.1) is already available from [39] and it is given by

$$\mathbb{E}\left[K_i^r\right] = \sum_{\nu=0}^{r} (-1)^{r-\nu} \left(1 + \frac{\theta + j\alpha}{\alpha}\right)_{\nu} S\left(r, \nu; \frac{\theta + j\alpha}{\alpha}\right) \frac{(\theta + j\alpha + \nu\alpha + 1)_{i-1}}{(\theta+1)_{i-1}}$$

where S is the non–central Stirling number of the second kind. See, e.g., [6]. Hence, one has

$$\begin{aligned}
&\mathbb{E}\left[(K_m^{(n)})^r \mid K_n = j\right] \\
&= \frac{\Gamma(\theta+n)}{\Gamma(\theta + j\alpha)\Gamma(n - j\alpha)} \int_0^1 w^{\theta + j\alpha - 1} (1-w)^{n - j\alpha - 1} \mathbb{E}\left[(K_m^{(n)})^r \mid K_n = j,\, w\right] \mathrm{d}w \\
&= \frac{\Gamma(\theta+n)}{\Gamma(\theta + j\alpha)\Gamma(n - j\alpha)} \sum_{\nu=0}^{r} (-1)^{r-\nu} \left(1 + \frac{\theta + j\alpha}{\alpha}\right)_{\nu} S\left(r, \nu; \frac{\theta + j\alpha}{\alpha}\right) \times \\
&\quad \times \sum_{i=0}^{m} \binom{m}{i} \frac{(\theta + j\alpha + \nu\alpha + 1)_{i-1}}{(\theta+1)_{i-1}} \int_0^1 w^{\theta + j\alpha + i - 1} (1-w)^{n - j\alpha + m - i - 1} \mathrm{d}w \\
&= \frac{1}{(\theta+n)_m} \sum_{\nu=0}^{r} (-1)^{r-\nu} \left(1 + \frac{\theta + j\alpha}{\alpha}\right)_{\nu} S\left(r, \nu; \frac{\theta + j\alpha}{\alpha}\right) \frac{\theta + j\alpha}{\theta + j\alpha + \nu\alpha} \times \\
&\quad \times \sum_{i=0}^{m} \binom{m}{i} (\theta + j\alpha + \nu\alpha)_i (n - j\alpha)_{m-i} \\
&= \frac{1}{(\theta+n)_m} \sum_{\nu=0}^{r} (-1)^{r-\nu} \left(\frac{\theta}{\alpha} + j\right)_{\nu} S\left(r, \nu; \frac{\theta + j\alpha}{\alpha}\right) (\theta + n + \nu\alpha)_m,
\end{aligned} \tag{B.2}$$

where the last equality follows by an application of the Chu–Vandermonde formula. See, e.g., [6]. Note, that for $r = 1$, we have

$$\mathbb{E}[K_m^{(n)} \mid K_n = j] = \left(j + \frac{\theta}{\alpha}\right) \left\{\frac{(\theta + n + \alpha)_m}{(\theta+n)_m} - 1\right\}, \tag{B.3}$$

The asymptotic moments are then derived by letting $m \to \infty$ in (B.2). Resorting again to the above–mentioned Stirling approximation we have

$$\frac{1}{m^{r\alpha}}\, \mathbb{E}\left[(K_m^{(n)})^r \mid K_n\right] \;\to\; \left(K_n + \frac{\theta}{\alpha}\right)_r \frac{\Gamma(\theta+n)}{\Gamma(\theta + n + r\alpha)} =: \mu_r^{(n)}. \tag{B.4}$$

Clearly such a moment sequence arises by taking $U_{n,j} \stackrel{d}{=} B_{j+\theta/\alpha,\, n/\alpha-j}\, S_{(\theta+n)/\alpha}$, with the beta random variable $B_{j+\theta/\alpha,\, n/\alpha-j}$ independent from $S_{(\theta+n)/\alpha}$, which has density (11). Hence, we are left with showing that the distribution of $U_{n,j}$ is uniquely characterized by the moment sequence $\{\mu_r^{(n)}\}_r$. In order to establish this, one can evaluate the characteristic function of $U_{n,j}$ which, at any $t \in \mathbb{R}$, coincides with

$$
\begin{aligned}
\Phi(t) &= \mathbb{E}\left[\mathrm{e}^{\mathrm{i}tU_{n,j}}\right] \\
&= \frac{\Gamma\left(\frac{\theta+n}{\alpha}\right)}{\Gamma\left(K_n+\frac{\theta}{\alpha}\right)\Gamma\left(\frac{n}{\alpha}-K_n\right)}\frac{\Gamma(\theta+n+1)}{\Gamma\left(\frac{\theta+n}{\alpha}+1\right)} \\
&\qquad \times \int_0^\infty \mathrm{e}^{\mathrm{i}tz}\, z^{K_n+\frac{\theta}{\alpha}-1} \int_z^\infty w\,(w-z)^{\frac{n}{\alpha}-K_n-1} f_\alpha(w)\mathrm{d}w\mathrm{d}z \\
&= \frac{\alpha\,\Gamma(\theta+n)}{\Gamma\left(K_n+\frac{\theta}{\alpha}\right)\Gamma\left(\frac{n}{\alpha}-K_n\right)} \int_0^\infty w\, f_\alpha(w) \int_0^w \mathrm{e}^{\mathrm{i}tz}\, z^{K_n+\frac{\theta}{\alpha}-1}(w-z)^{\frac{n}{\alpha}-K_n-1}\mathrm{d}z\mathrm{d}w \\
&= \frac{\Gamma(\theta+n+1)}{\Gamma\left(\frac{\theta+n}{\alpha}+1\right)} \sum_{r\geq 0} \frac{(\mathrm{i}t)^r}{r!} \frac{\left(K_n+\frac{\theta}{\alpha}\right)_r}{\left(\frac{\theta+n}{\alpha}\right)_r} \int_0^\infty w^{\frac{\theta+n}{\alpha}+r} f_\alpha(w)\mathrm{d}w \\
&= \sum_{r\geq 0} \frac{(\mathrm{i}t)^r}{r!} \frac{\left(K_n+\frac{\theta}{\alpha}\right)_r}{\left(\frac{\theta+n}{\alpha}\right)_r} \frac{\Gamma(\theta+n+1)}{\Gamma\left(\frac{\theta+n}{\alpha}+1\right)} \frac{\Gamma\left(\frac{\theta+n}{\alpha}+r+1\right)}{\Gamma(\theta+n+1+r\alpha)} = \sum_{r\geq 0} \frac{(\mathrm{i}t)^r}{r!}\mu_r^{(n)}
\end{aligned}
$$

Hence, we have established that, conditionally on $K_n = j$, $K_m^{(n)}/m^\alpha$ converges a.s. to $U_{n,j}$.

As for the second component of the aggregate output (16), namely $\beta L_m^{(n)}/m^{1-\tau}$, first note that by [28, Proposition 2], we have

$$
\mathbb{E}[L_m^{(n)} | K_n = j] = m\frac{\theta+\alpha j}{\theta+n}.
$$

This, combined with (B.3), yields immediately (17). The law of $L_m^{(n)}$ is given in Equation (22) of [28], which is easily seen to coincide with a Pólya distribution

$$
\mathbb{P}[L_m^{(n)} = s | K_n = j] = \binom{m}{s} \frac{Be(m-s+n-j\alpha,\, s+\theta+j\alpha)}{Be(n-j\alpha,\, \theta+j\alpha)} \quad s = 0,\ldots,m,
$$

where $Be(a,b)$ denotes a beta function. Hence, the number of innovations within the new sectors follows a Pólya distribution. Therefore, by well–known martingale convergence arguments, it follows that $L_m^{(n)}/m$ converges a.s. to a beta random variable with parameters $\theta+j\alpha$ and $n-j\alpha$, ocnditionally on $K_n = j$. Now, combining this limit result with the previous concerning $K_m^{(n)}$ the asymptotic statements in (i), (ii) and (iii) follow immediately. □

Acknowledgements

The authors are grateful to the Editor, an Associate Editor and a Referee for their valuable comments which led to an improvement of the presentation. Igor Prünster also wishes to thank the Department of Statistics and Data Science of the University of Texas at Austin for its kind hospitality while he was completing this paper.

References

[1] Acemoglu, D. (2009). *Introduction to Modern Economic Growth*. Princeton University Press.
[2] Aghion, P. and Howitt, P. (1997). A Schumpeterian perspective on growth and competition. In *Advances in Economics and Econometrics* (Kreps, D., ed.).
[3] Aoki, M. (2008). Thermodynamic limits of macroeconomic or financial models: one- and two-parameter Poisson-Dirichlet models. *J. Econom. Dynam. Control* **32**, 66–84. MR2381689
[4] Aoki, M. and Yoshikawa, H. (2012). Non–self–averaging in macroeconomic models: a criticism of modern micro–founded macroeconomics. *J. Econ. Interact. Coord.* **7**, 1–22.
[5] Burda, M., Harding, M. and Hausman, J. (2008). A Bayesian mixed logit-probit model for multinomial choice. *J. Econometrics* **147**, 232–246. MR2478523
[6] Charalambides, C.A. (2005). *Combinatorial Methods in Discrete Distributions*. Wiley, Hoboken, NJ. MR2131068
[7] Cont, R. and Tankov, P. (2004). *Financial modelling with jump processes*. Chapman & Hall/CRC, Boca Raton, FL. MR2042661
[8] De Blasi, P., Favaro, S., Lijoi, A., Mena, R.H., Prünster, I. and Ruggiero, M. (2015). Are Gibbs-type priors the most natural generalization of the Dirichlet process? *IEEE Trans. Pattern Anal. Mach. Intell.* **37**, 212–229.
[9] De Blasi, P., James, L.F. and Lau, J.W. (2010). Bayesian nonparametric estimation and consistency of Mixed Multinomial Logit choice models. *Bernoulli* **16**, 679–704. MR2730644
[10] Devroye, L. (2009). Random variate generation for exponentially and polynomially tilted stable distributions. *ACM Transactions on Modeling and Computer Simulation* **19**, N. 18.
[11] Ewens, W.J. (1972). The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.* **3** 87–112. MR0325177
[12] Favaro, S., Lijoi A., Mena, R.H. and Prünster, I. (2009). Bayesian nonparametric inference for species variety with a two parameter Poisson-Dirichlet process prior. *J. R. Stat. Soc. Ser. B* **71**, 993–1008. MR2750254
[13] Ferguson, T.S. and Klass, M.J. (1972). A representation of independent increments processes without Gaussian components. *Ann. Math. Statist.* **43**, 1634–1643. MR0373022

[14] Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1**, 209–230. MR0350949

[15] Gancia, G. and Zilibotti, F. (2005). Horizontal Innovation in the Theory of Growth and Development. In *Handbook of Economic Growth* (P. Aghion and S. Durlauf eds), vol. 1A, pp. 111–170, North Holland, Amsterdam.

[16] Gnedin, A. and Pitman, J. (2005). Exchangeable Gibbs partitions and Stirling triangles. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **325**, 83–102 (transl. in *J. Math. Sci. (N.Y.)* **138** (2006), 5674–5685). MR2160320

[17] Griffin, J.E. (2011). The Ornstein-Uhlenbeck Dirichlet process and other time-varying processes for Bayesian nonparametric inference. *J. Statist. Plann. Inference* **141**, 3648–3664. MR2817371

[18] Griffin, J.E. and Steel, M.F.J. (2004). Semiparametric Bayesian inference for stochastic frontier models. *J. Econometrics* **123**, 121–152. MR2126161

[19] Griffin, J.E. and Steel, M.F.J. (2006). Order-based dependent Dirichlet processes. *J. Amer. Statist. Assoc.* **101**, 179–194. MR2268037

[20] Griffin, J.E. and Steel, M.F.J. (2011). Stick-Breaking Autoregressive Processes. *J. Econometrics*, **162**, 383–396. MR2795625

[21] Hirano, K. (2002). Semiparametric Bayesian inference in autoregressive panel data models. *Econometrica* **70**, 781–799. MR1913831

[22] Hjort, N.L., Holmes, C.C. Müller, P., Walker, S.G. (Eds.) (2010). *Bayesian Nonparametrics*. Cambridge University Press, Cambridge. MR2722987

[23] Lau, J.W. and Siu, T.K. (2008). On option pricing under a completely random measure via a generalized Esscher transform. *Insurance Math. Econom.* **43**, 99–107. MR2442035

[24] Lau, J.W. and Siu, T.K. (2008). Long-term investment returns via Bayesian infinite mixture time series models. *Scand. Actuar. J.* **4**, 243–282. MR2484128

[25] Lijoi, A., Mena, R.H. and Prünster, I. (2005). Hierarchical mixture modeling with normalized inverse-Gaussian priors. *J. Amer. Statist. Assoc.* **100**, 1278–1291. MR2236441

[26] Lijoi, A., Mena, R.H. and Prünster, I. (2007a). Controlling the reinforcement in Bayesian non-parametric mixture models. *J. R. Stat. Soc. Ser. B* **69**, 715–740. MR2370077

[27] Lijoi, A., Mena, R.H., and Prünster, I. (2007b). Bayesian nonparametric estimation of the probability of discovering a new species *Biometrika*. **94** 769–786. MR2416792

[28] Lijoi, A., Prünster, I., and Walker, S.G. (2008). Bayesian nonparametric estimators derived from conditional Gibbs structures. *Ann. Appl. Probab.* **18**, 1519–1547. MR2434179

[29] Mena, R.H., Ruggiero, M. and Walker, S.G. (2011). Geometric stick-breaking processes for continuous-time Bayesian nonparametric modeling. *J. Statist. Plann. Inference* **141**, 3217–3230. MR2796026

[30] Mena, R.H. and Ruggiero, M. (2016). Dynamic density estimation with diffusive Dirichlet mixtures. *Bernoulli* **22**, 901–926. MR3449803

[31] Mena, R.H. and Walker, S.G. (2005). Stationary models via a Bayesian nonparametric approach. *J. Time Ser. Anal.* **26**, 789–805. MR2203511
[32] Pitman, J. (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields* **102**, 145–158. MR1337249
[33] Pitman, J. (1996). Some developments of the Blackwell-MacQueen urn scheme. *Statistics, probability and game theory. Papers in honor of David Blackwell* (T.S. Ferguson, L.S. Shapley and J.B. MacQueen eds.), 245–267, IMS Lecture Notes Monogr. Ser., Hayward, CA. MR1481784
[34] Pitman, J. (2006). *Combinatorial stochastic processes.* Ecole d'Eté de Probabilités de Saint-Flour XXXII. Lecture Notes in Mathematics N. 1875. Springer, New York. MR2245368
[35] Prünster, I. and Ruggiero, M. (2013). A Bayesian nonparametric approach to modeling market share dynamics. *Bernoulli* **19**, 64–92. MR3019486
[36] Romer, P. (1990). Endogeneous technological change. *J. Politic. Econ.* **98**, S71–S102.
[37] Rosiński, J. (2007). Tempering stable processes. *Stochastic Process. Appl.* **117**, 677–707. MR2327834
[38] Solow, R. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics* **70**, 65–94.
[39] Yamato, H. and Sibuya, M. (2000). Moments of some statistics of Pitman sampling formula. *Bull. Inform. Cybernet.* **32**, 1–10. MR1792352