Research Policy 55 (2026) 105587

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Modeling innovation ecosystem dynamics through interacting reinforced Bernoulli processes☆,☆☆

Giacomo Aletti [a,1], Irene Crimaldi [b,1], Andrea Ghiglietti [c,1], Federico Nutarelli [b,*,1]

[a] ADAMSS Center, Università degli Studi di Milano, Via Celoria 2, Milano, 20133, MI, Italy
[b] IMT School for Advanced Studies Lucca, Piazza S. Francesco 19, Lucca, 55100, LU, Italy
[c] Department of Statistics and Quantitative Methods, Università degli Studi di Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, Milano, 20126, MI, Italy

ARTICLE INFO

Keywords:
Innovation ecosystems
Technological interdependence
Interacting Bernoulli processes
Patent interactions
Technological specialization

ABSTRACT

Innovation is cumulative and interdependent: successful inventions build on prior knowledge within technological fields and may also affect success across related ones. Yet these dimensions are often studied separately in the innovation literature. This paper asks whether patent success across technological categories can be represented within a single dynamic framework that jointly captures within-category reinforcement, cross-category spillovers, and a set of aggregate regularities observed in patent data. To address this question, we propose a model of interacting reinforced Bernoulli processes in which the probability of success in a given category depends on past successes both within that category and across other categories. The framework yields joint predictions for success probabilities, cumulative successes, relative success shares, and cross-category dependence.

We implement the model using granted US patent families from GLOBAL PATSTAT (1980–2018), defining category-specific success through a cohort-normalized forward-citation index. The empirical analysis shows that successful innovations continue to accumulate, but less than proportionally to the growth in patent opportunities, while technological categories remain interdependent without becoming homogeneous. Under a mean-field restriction, the model-based inferential exercise points to a positive but non-maximal interaction across technological categories.

1. Introduction

Innovation is cumulative and interdependent. Innovative success builds on prior knowledge within technological fields, but it is also shaped by spillovers across related fields. These two ideas run through much of the innovation literature. Early work emphasized the role of research, technological opportunity, and appropriability conditions in shaping inventive performance (Griliches, 1958; Evenson, 1968; Levin et al., 1985; Jaffe, 1986). Subsequent contributions showed that innovative activity also differs systematically across sectors because of differences in cumulativeness, demand conditions, and sectoral regimes (Malerba and Orsenigo, 1996; Malerba, 2002). A related literature documented intersectoral technological linkages and knowledge spillovers, highlighting how inventive activity in one domain may affect innovation in other technologically related domains (Rosenberg, 1979; Scherer, 1982; Jaffe, 1989). More recently, work on recombinant innovation and on the "adjacent possible" has stressed that innovation often proceeds through combinations of related knowledge elements and expansion into nearby technological opportunities (Fleming, 2001;

☆ The work is partially supported by the European Union - NextGenerationEU through the Italian Ministry of University and Research under the National Recovery and Resilience Plan (PNRR) - M4C2 - Investment 1.3, title [Public Sector Indicators for Sustainability and Wellbeing (PUWELL)] - Program [Growing Resilient, INclusive and Sustainable (GRINS)] - PE18 - CUP J33C22002910001. This research was supported by the "Resilienza Economica e Digitale" (RED) project (CUP D67G23000060001) funded by the Italian Ministry of University and Research (MUR) as "Department of Excellence" (Dipartimenti di Eccellenza 2023-2027, Ministerial Decree no. 230/2022).

☆☆ Giacomo Aletti is a member of the Italian Group "Gruppo Nazionale per il Calcolo Scientifico" of the Italian Institute "Istituto Nazionale di Alta Matematica". Irene Crimaldi is a member of the Italian Group "Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni" of the Italian Institute "Istituto Nazionale di Alta Matematica".

* Corresponding author.
E-mail addresses: giacomo.aletti@unimi.it (G. Aletti), irene.crimaldi@imtlucca.it (I. Crimaldi), andrea.ghiglietti@unimib.it (A. Ghiglietti), federico.nutarelli@imtlucca.it (F. Nutarelli).

[1] All the authors contributed equally to the present work.

https://doi.org/10.1016/j.respol.2026.105587
Received 24 September 2025; Received in revised form 29 June 2026; Accepted 1 August 2026
Available online 1 September 2026
0048-7333/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Kauffman, 2000; Monechi et al., 2017; Taalbi, 2020). At the same time, a long tradition — from early concerns about the rate of invention to recent evidence on declining disruptiveness and research productivity — suggests that producing influential innovations may become more difficult as the stock of existing knowledge expands (Stafford, 1952; Schmookler, 1954; Jones, 1995a,b; Ha and Howitt, 2007; Park et al., 2023; Madsen et al., 2024).

What is less developed in the innovation literature is a parsimonious dynamic framework that studies these elements within a single empirical setting. The literature has documented cumulativeness within fields, spillovers across fields, and broad regularities in innovative outcomes, but these dimensions are often examined separately or through reduced-form correlations, proximity measures, and network mappings rather than through a common dynamic model of innovation success (Jaffe, 1989; Pichler et al., 2020; Taalbi, 2020; Colladon et al., 2025). This paper addresses that issue by asking whether innovation success across technological categories can be represented within a single stochastic framework that jointly captures within-category reinforcement, cross-category spillovers, and a small set of aggregate patterns observed in patent data.

We propose such a framework. Specifically, we model category-specific patent successes as a system of interacting reinforced Bernoulli processes. At each patent time-step, a category-specific indicator takes value one if the patent qualifies as a success in that category and zero otherwise. The probability of success in category h depends on past successes in the same category and on cumulative successes in other categories. Cross-category dependence is summarized by a non-negative interaction matrix $\Gamma = (\gamma_{j,h})$, which should be interpreted as a reduced-form representation of technological linkages rather than as a model of how such linkages themselves emerge. In this sense, the paper does not endogenize the formation of technological relations; rather, it studies the dynamic implications of reinforcement and interdependence for the evolution of innovation success across categories.

The empirical application uses granted US patent families from GLOBAL PATSTAT, classified at the CPC-1 level. Our empirical object is not patenting per se, nor novelty in the legal sense used in patent examination. Rather, it is success in a technological category, measured ex post through technological influence on subsequent inventions. This distinction matters. In the innovation literature, novelty and radicalness are broader concepts than legal novelty and are often associated with recombination, distinctiveness, and impact (Fleming, 2001; Dahlin and Behrens, 2005). In the present paper, however, we focus on success as the observable outcome to be modeled dynamically. As baseline outcome, we define a category-specific patent success through a cohort-normalized forward-citation measure following Squicciarini et al. (2013): for a given patent and target category, we count subsequent citations from patents in that category within a fixed time window (equal for all the patents) and normalize by the maximum value observed in the patent's year-by-category cohort. This normalization is intended to improve comparability across cohorts and technological classes, but it does not mechanically impose the dynamic patterns studied later in the paper. Indeed, it is a within-cohort rescaling, whereas the theoretical results concern the sequential accumulation of successes across patents in system time. In other terms, the normalization makes patents more comparable across cohorts, but the dynamical properties of the model arise from the reinforcement mechanism rather than from the normalization itself. In addition, note that, in order to neutralize the mechanical right-truncation problem associated with unequal citation windows, we pay attention to the fact that all the patents considered for the empirical analysis share the same time-window to receive citations.

This empirical construction yields, for each patent and category, a binary success indicator that can be interpreted as the observable counterpart of the success events studied in the model. The relevant time dimension is *system time*: the model concerns the conditional probability that the next observed patent is a success in category h, given the history of previously observed patents. This point is important for interpretation. When the model implies declining success probabilities, it does not refer to a purely mechanical trend in calendar time. Rather, it means that, as the number of patent opportunities already observed increases, cumulative successes grow less than proportionally. In this sense, success becomes rarer *per patent opportunity*, not because a downward trend is imposed ex ante, but because it is implied by the reinforcement dynamics. The latter pattern is consistent with the idea that influential innovations may become harder to generate as the system expands and the knowledge burden increases (Bloom et al., 2020; Jones, 2009).

The paper contributes to innovation research by proposing a dynamic stochastic framework for category-specific innovation success that jointly captures cumulativeness within technological categories and spillovers across categories. The framework yields three aggregate implications that can be taken to the data: cumulative successes grow sublinearly in system time; relative success shares across categories stabilize in the long run; and cross-category comovement weakens over time in the empirically relevant unbalanced case. We compare these implications with patent data and complement this comparison with a parsimonious inference exercise on interaction intensity under a mean-field restriction. Appendix E provides an econometric analysis as a robustness check to confirm whether the successes in a target category are positively associated with past successes in the same category (path dependence) and with past successes in the other categories (cross-fertilization).

The remainder of the paper is organized as follows. Section 2 defines the category-level success measure and discusses its interpretation. Section 3 presents the interacting Bernoulli framework and its main theoretical implications. Section 4 examines whether the aggregate patterns predicted by the model are present in the patent data and reports the complementary inference exercise on interaction intensity. Section 5 concludes. Finally, the appendix contains an examination of the meaning of novelty in the literature and collects technical results, together with the corresponding analytical proofs, and additional data analysis.

2. Category forward citation index and definition of "success"

In this paper, patents are used as observable innovation outcomes. Our empirical object is not patenting per se, but whether a patent qualifies as a *successful* innovation in a given technological category according to an ex post citation-based criterion. This distinction is important. In patent examination, novelty is a legal requirement: a patent must be new relative to the prior art at the time of filing. In the innovation literature, however, novelty and radicalness are broader concepts, often linked to recombination, distinctiveness, and subsequent impact (Fleming, 2001; Dahlin and Behrens, 2005). Furthermore, following Dahlin and Behrens (2005), radical inventions are typically understood as combining novelty, uniqueness, and impact. In this sense, novelty is one dimension of radicalness rather than a separate concept that we aim to distinguish from it. Our empirical measure does not aim to operationalize novelty in all these senses. Rather, it focuses on *success* as an observable ex post measure of technological influence. A patent may therefore be legally novel and yet have limited technological impact, while a highly cited patent may indicate an invention that was proven especially important for later developments. Similarly, we do not interpret patents classified as successful in our framework as necessarily novel or radical. Rather, as noted above, these are patents that became influential for subsequent inventive activity within a given technological category. Importantly, the proposed methodology could, in principle, be applied to alternative patent-based measures of novelty or radicalness. However, our baseline empirical application focuses on citation-based success. For the described reasons, throughout the paper we reserve the term "success" for patents that exceed a category-specific threshold based on forward citations.

Forward citations are widely used in the innovation literature as indicators of the technological importance of patents for subsequent developments (Dahlin and Behrens, 2005; Squicciarini et al., 2013).

We therefore build on the forward-citation index proposed by Squicciarini et al. (2013), adapting it to a category-specific setting. The empirical construction is not tied to a specific classification system or to this particular index: the same framework could be implemented with alternative taxonomies or alternative measures of technological influence.[2]

We consider a finite set $\mathcal{H}$ of technological categories. For each patent n, we denote by c_n its source category in the empirical classification and by d_n its publication (full) date. Success, however, is evaluated separately for each target category $h \in \mathcal{H}$. Hence, a patent is associated with one source category c_n but may qualify as a success in none, one, or several target categories, depending on the citations it subsequently receives from patents in those categories. This distinction between source category and target-category success is central to the empirical construction used in the rest of the paper.

At the empirical level, technological categories are defined at the CPC-1 level. On the citing side, category-specific citation counts may reflect all technological categories represented in the citing patent, so the target-category dimension can incorporate multiple CPC-1 classes. By contrast, whenever a unique source category is required in the empirical analysis below, each patent is assigned to a single CPC-1 source category, defined as the modal first-level CPC letter among the CPC codes associated with that patent in the original data.

For a given patent n, with publication date d_n, and for each category $h \in \mathcal{H}$, let

$$CIT_{n,h}$$

denote the number of patents published in the period $[d_n, d_n + 365 \times T]$, belonging to category h, and citing patent n.

Now consider patent n, published in year y_n and belonging to source category c_n. We refer to the set of patents published in the same year y_n and belonging to the same source category c_n as the *cohort of patent* n. For each target category h, we define the index

$$I_{n,h} = \frac{CIT_{n,h}}{\max_{\{i \in \text{cohort of } n\}} CIT_{i,h}}.$$

Thus, $I_{n,h}$ measures the forward citations received by patent n from category h, normalized by the maximum value observed within the patent's publication year/source-category cohort.

Note that, in practical applications, if the available dataset includes patents published up to a certain year T_f, then, once the size T of the time-window for the index calculation has been chosen, the analysis should be performed on the patents with publication year less than or equal to $(T_f - T)$, so that, for each considered patent, the entire time-window $[d_n, d_n + 365 \times T]$ required for computing the index would be available. This would neutralize the mechanical right-truncation problem associated with unequal citation windows: indeed, all the patents considered for the analysis would have the same time-window of T years to receive citations. Moreover, the purpose of the normalization in $I_{n,h}$ is the following. It accounts for the fact that forward citations vary substantially across publication years and technological categories, making raw citation counts difficult to compare across cohorts. The normalization is computed within the patent's publication-year/source-category cohort, in order to mitigate comparability problems in citation counts. However, the normalization does not mechanically impose the dynamic patterns studied later in the paper: it is a within-cohort rescaling, whereas the theoretical results concern the sequential accumulation of successes across patents in system time. In other words, the normalization makes patents more comparable across cohorts, but the asymptotic properties of the model arise from the reinforcement dynamics rather than from the normalization itself.[3]

[2] A review of novelty and impact measures that could be used in our setting is provided in Appendix A.

This category-specific index is a disaggregation of the forward-citation index in Squicciarini et al. (2013). Indeed, if $I_S(n)$ denotes their index, then

$$I_S(n) = \frac{CIT_n}{\max_{\{i \in \text{cohort of } n\}} CIT_i} = \sum_{h \in \mathcal{H}} \frac{\max_{\{i \in \text{cohort of } n\}} CIT_{i,h}}{\max_{\{i \in \text{cohort of } n\}} CIT_i} I_{n,h}.$$

Remark 1 (*Invariance Property of the Index*)**.** Note that, if we restrict our attention to (or we can observe) only the patents published in a certain category h^*, then, for each patent n belonging to this category, the index value I_{n,h^*} remains well defined and its value does not change with respect to the one computed if we consider all the published patents. Indeed, its definition is based only on the observation of the patents belonging to category h^*.

By construction, $I_{n,h} \in [0, 1]$. To obtain the binary success indicators required by the model, we fix a threshold τ and classify a patent as a success in category h whenever its index exceeds that threshold.[4] The threshold therefore translates a continuous measure of technological influence into a success/failure outcome. Different threshold choices correspond to different degrees of selectivity in the definition of success; for this reason, the choice of τ and its empirical implications are discussed in the Appendix D and revisited in the data analysis.

Definition 2.1 (*Success in Category* h)**.** Given a threshold τ, patent n is said to be a success in category h if

$$I_{n,h} > \tau.$$

Accordingly, let $X_{n,h}$ be the binary variable equal to 1 if patent n is a success in category h and equal to 0 otherwise. This yields a finite system of Bernoulli processes,

$$\{X^{(h)} = (X_{n,h})_n : h \in \mathcal{H}\},$$

which provides the empirical counterpart of the success events modeled in the next section. These indicators are not assumed to be independent. Rather, they are the observable category-specific success events whose dynamics we seek to represent through within-category reinforcement and cross-category interdependence.

In the next section, we introduce the dynamic stochastic framework and derive the aggregate implications that will later be compared with the patent data. Although the empirical implementation developed here is based on CPC categories and a citation-based success measure, the same methodological approach can be adapted to alternative classifications and alternative success indicators.

3. A dynamic framework for innovation success across categories

This section develops the stochastic model which is at the core of the paper. Starting from category-specific success indicators, such as the ones introduced in Section 2, we describe how innovation

[3] An additional reason for working with success indicators rather than cumulative patent counts is that this approach mitigates the initial-stock problem arising from the finite observation window. If the empirical outcome were the cumulative stock of patents, starting the sample in a given year T_0 would exclude patents published before that date, making the initial stock arbitrary. By contrast, the success indicator is defined at the patent level, using forward citations received within a fixed time horizon and normalized within each publication-year/source-category cohort. As a result, patents published before the beginning of the sample are not required to classify post-T_0 patents as successes or failures. Earlier technological history may still influence the dynamics through the model's initial conditions, but it does not enter the empirical outcome as an omitted cumulative patent stock.

[4] The threshold does not determine which patents enter the sample. Rather, it determines whether each patent-category observation is coded as a success or a failure.

success evolves when it is both cumulative within technological categories and interdependent across related categories. The purpose of the model is to characterize the dynamic implications of reinforcement and cross-category interdependence for the evolution of category-specific innovation success over time.

In the patent application considered in this paper, each process refers to a technological category and the time-step t indexes patents sequentially according to publication date. Unless otherwise specified, however, the model is presented in a general form and may be applied in other contexts as well.

Let $\mathcal{H}$ be a finite set with $N = \text{card}(\mathcal{H})$. We are going to define a system $\{X^{(h)} = (X_{t,h})_{t\in\mathbb{N}\setminus\{0\}} : h \in \mathcal{H}\}$ of Bernoulli processes with parameters evolving along time according to a reinforcement rule with interaction: specifically, for each $t \geq 0$, let $X_{t+1,1}, \dots, X_{t+1,N}$ be N Bernoulli random variables, that are conditionally independent given the past and such that

$$P_{t,h} = P(X_{t+1,h} = 1 \mid \text{past until time-step } t) = \frac{\theta_h + \sum_{j\in\mathcal{H}} \gamma_{j,h} S_{t,j}}{c + t} \tag{1}$$

where $\theta_h > 0$, $c \geq \theta_h$ are parameters that tune the initial condition in the model dynamics (for details, see (B.1) in Appendix B) and $S_{t,j} = \sum_{n=1}^{t} X_{n,j}$ is the number of successes until time-step t for process j.

The parameters $\gamma_{j,h}$ play the fundamental role in the dynamics of the system (see again (B.1)) and summarize how past successes in category j affect future success in category h. We refer to the matrix $\Gamma = (\gamma_{j,h})_{j,h}$ as the interaction matrix. In the empirical interpretation of the model, Γ should be read as a reduced-form representation of technological interdependence across categories. We further assume the following conditions:

($\mathcal{A}1$) Γ is non-negative and such that $\mathbf{1}^\top\Gamma \leq \mathbf{1}^\top$ (where $\mathbf{1} = (1, \dots, 1)^\top$), i.e. $\sum_{j\in\mathcal{H}} \gamma_{j,h} \leq 1$ for each $h \in \mathcal{H}$ so that we have $P_{t,h} \in (0, 1)$;

($\mathcal{A}2$) Γ is irreducible, that is the graph with the Bernoulli processes as nodes and with Γ as the (weighted) adjacency matrix is strongly connected. (In Appendix C we will discuss the case of a reducible matrix Γ.[5])

Assumption ($\mathcal{A}1$) ensures that reinforcement does not make success probabilities explode beyond the unit interval, while Assumption ($\mathcal{A}2$) guarantees that no category is isolated from the rest of the system. Together, these conditions describe a setting in which successes can propagate across categories – consistently with the innovation literature emphasizing that technological outcomes are shaped by interdependence and spillovers across related domains rather than evolving in isolation (Jaffe, 1989; Taalbi, 2020)– but in a statistically novel and well-defined way[6]

More formally, at each time-step t, the probability $P_{t,h}$ that we will have a success for process h at time-step $t + 1$ can have an increasing dependence, not only on the number of successes observed in process h itself, according to a *self-reinforcement principle* (Pemantle, 2007), but also on the number $S_{t,j}$ of successes observed in any other process j until time-step t (a property that we call *cross-reinforcement*). The parameter $\gamma_{j,h}$ regulates this dependence (with $\gamma_{j,h} = 0$ meaning the absence of direct dependence). In other words, for each pair (j, h) of processes, the parameter $\gamma_{j,h}$ quantifies how much the appearance of a success in process j induces a potential future success in process h. It is important to note that the assumption of irreducibility of Γ entails an *interaction* in this sense among all the processes: process j can reinforce process h directly (the case when $\gamma_{j,h} > 0$) or indirectly, through the presence of at least one path joining them.

We first characterize the long-run behavior of success probabilities. This result is useful for interpretation, because it shows when influential innovations become rarer relative to the number of patent opportunities generated by the system.

Theorem 3.1. *Under assumptions ($\mathcal{A}1$) and ($\mathcal{A}2$), denote by $\gamma^* \in (0, 1]$ the Perron–Frobenius eigenvalue of Γ. Then, for each $h \in \mathcal{H}$, we have as the number t of patents grows*

$$t^{1-\gamma^*} P_{t,h} \xrightarrow{a.s.} P_{\infty,h},$$

where $P_{\infty,h}$ is a finite strictly positive random variable. Moreover, for each $h, j \in \mathcal{H}$, the ratio $P_{\infty,h}/P_{\infty,j}$ of the above limit random variables is almost surely equal to the deterministic quantity $u_h/u_j > 0$, where $\mathbf{u} = (u_h)_h$ is the (unique up to a multiplicative non-zero constant) left eigenvector of Γ associated to γ^*. Finally, the above convergence is also in quadratic mean.

We recall that – according to the Perron–Frobenius theory – the components of the vector $\mathbf{u}$ are all different from zero and with the same sign and, in graph theory, they correspond to the relative eigenvector centrality scores (with respect to $\Gamma^\top$).[7] Hence, the quantity u_h is a (relative) measure of how strongly category h is exposed to successes coming from categories that are themselves central in the system of technological interdependencies, consistently with the idea that technological position and network linkages shape both spillovers and subsequent innovative outcomes (Jaffe, 1989; Taalbi, 2020).

This result implies that

- in the case $\gamma^* < 1$ (that is when $\mathbf{1}^\top\Gamma \neq \mathbf{1}^\top$) the probability of observing a success converges almost surely toward zero, i.e. successful innovations continue to occur, but they become progressively rarer relative to the number of patent opportunities observed in the system and this convergence happens sublinearly at the same rate $1/t^{(1-\gamma^*)}$ for every process (category) h;
- in the case $\gamma^* = 1$ (that is when $\mathbf{1}^\top\Gamma = \mathbf{1}^\top$) the probability of observing a success in process h converges almost surely to a finite strictly positive random limit, which is the same for each

[5] This extension may be particularly relevant in contexts where the USPTO –or other patent offices—modify the structure of technological categories through reclassification (Lafond and Kim, 2019; Chae and Gim, 2019) thus changing the landscape of interaction among the latter.

[6] Namely, our framework relates to the literature on reinforced stochastic processes surveyed by Pemantle (2007), which includes generalized Pólya urns, interacting urn models, and other processes with reinforcement. Relative to that literature, our contribution is to formulate a multi-category system of patent-success processes with an economically interpretable interaction matrix and to derive joint predictions for success probabilities, cumulative successes, relative shares, cross-category dependence, and shock propagation.

The economic interpretation of the model is straightforward. A higher number of past successes in category h raises the probability of future success in the same category. At the same time, past successes in other categories may also raise the probability of success in h, capturing cross-category spillovers and technological interdependence. The parameters $\gamma_{j,h}$ govern the strength of these direct cross-category effects. In other words, given the sequence of patents observed up to time t, the probability that the next patent is a success in category h depends on the cumulative record of past successes both within category h and across other categories. In this sense, our framework provides a dynamic stochastic representation of how within-category cumulativeness and cross-category spillovers jointly shape the evolution of innovation success, capturing cumulativeness and path dependence within technological fields.

[7] The eigenvector centrality is a measure of the importance of a node in a graph with respect to its out-links (if it is computed for the adjacency matrix) or its in-links (if it is computed for the transpose of the adjacency matrix). A high eigenvector centrality score means that the node points to, or respectively is pointed to by, many nodes with high scores. In the above theorems, $\mathbf{u}$ is the vector of the relative eigenvector centrality scores with respect to $\Gamma^\top$ and, hence, by (1), a high value of u_h means that the probability of having a success for process h depends on the number of past successes observed in many processes that themselves have high scores.

process (note that in this case $\mathbf{u}$ is, up to a multiplicative non-zero constant, equal to the vector $\mathbf{1}$); in other words, at the steady state, the probability of having a success in a process is the same for each process of the system. Economically, this means that cross-category reinforcement is strong enough to prevent the long-run success probability from vanishing.

Remark 2. A clarification on the time scale is useful here. The quantity $P_{t,h}$ is defined in system time: it is the probability that the next patent in the sequence is a success in category h, conditional on the history of previously observed patents. Therefore, the statement that $P_{t,h}$ declines when $\gamma^* < 1$ should not be read as a purely calendar-time trend. Rather, it means that, conditional on the number of patent opportunities already generated by the system, additional successes become progressively rarer. This is precisely why the result is non-trivial. The model does not start from an assumed decline in $P_{t,h}$. Instead, the decline is derived from the reinforcement dynamics and the interaction structure summarized by Γ. The asymptotic relation for $S_{t,h}$ implies that success probabilities decrease because cumulative successes grow less than proportionally with the number of observed patent opportunities.

This distinction is useful for interpreting the model in innovation terms. When $\gamma^* < 1$, the system displays a form of subcritical reinforcement: past successes still matter, but not enough to sustain a constant rate of influential innovations as the system expands. When $\gamma^* = 1$, instead, interdependencies across categories are sufficiently strong to keep success probabilities persistently positive in the long run. This interpretation is consistent with the literature showing, on the one hand, that the absence of autocatalytic structures can weaken innovation dynamics (Napolitano et al., 2018) and, on the other hand, that cross-domain knowledge flows and technological linkages can sustain innovation opportunities across domains (Jaffe, 1989; Kim and Magee, 2017; Taalbi, 2020). As we will see in what follows, our data analysis points to the empirically relevant case $\gamma^* < 1$.

The most relevant implications for the empirical analysis concern the observable cumulative counts of successes. Unlike the conditional probabilities $P_{t,h}$, these quantities can be directly constructed from the patent data and therefore provide the main bridge between the model and the empirical patterns studied in Section 4.

In particular, from Theorem 3.1, we can also deduce the following result for the processes $(S_{t,h})_t$, with $h \in \mathcal{H}$, that represent the number of successes observed along time in each Bernoulli process h. We enforce here again that, differently from the previous quantities $P_{t,h}$, that are not observable (due to the presence of the unknown model parameters), the quantities $S_{t,h}$ *can be directly observed* and so used for data analyses.

Theorem 3.2. *Under assumptions ($\mathcal{A}1$) and ($\mathcal{A}2$), denote by $\gamma^* \in (0,1]$ the Perron–Frobenius eigenvalue of Γ and by $\mathbf{u} = (u_h)_h$ its corresponding (unique up to a multiplicative non-zero constant) left eigenvector. Then, for each $h \in \mathcal{H}$, we have*

$$\frac{S_{t,h}}{t^{\gamma^*}} \xrightarrow{a.s.} S_{\infty,h},$$

where $S_{\infty,h}$ is a finite strictly positive random variable. Moreover, for each $h, j \in \mathcal{H}$, we have

$$\frac{S_{t,h}}{S_{t,j}} \xrightarrow{a.s.} \frac{u_h}{u_j} \quad \text{and} \quad \frac{S_{t,h}}{\sum_{j=1}^{N} S_{t,j}} \xrightarrow{a.s.} \frac{u_h}{\sum_{j=1}^{N} u_j}.$$

This result states that the number $S_{t,h}$ of successes for all the processes h grows with the same Heaps' exponent $\gamma^* \in (0,1]$: that is $S_{t,h} \overset{a.s.}{\sim} S_{\infty,h}\, t^{\gamma^*}$. In addition, the ratio $S_{t,h}/S_{t,j}$ provides a strongly consistent estimator of the ratio u_h/u_j of the relative eigenvector centrality scores (with respect to $\Gamma^\top$) of the two nodes (processes) h and j and the share of successes observed for process h converges almost surely to the absolute eigenvector centrality score of h.

It is interesting to note that, in the innovation framework, Theorem 3.2 highlights how the long-run proportions of successful patents in each category reflect the connectedness and influence of that category within the broader innovation network. While this relationship is formally established in the present work, related insights have been suggested in the literature without formal proof. For example, Pichler et al. (2020) show that the innovation rate of a technological domain is influenced by the innovation rates of the domains it depends on. Similarly, Sampat and Ziedonis (2005), Zhang et al. (2025) emphasize the importance of the quality of a patent's prior connections as well as their technological domain.

Because the framework is stochastic, it also yields implications for fluctuations and cross-category comovement. We believe that the latter additional implication of the model may highlight interesting implications as the uncertainty (or volatility) of innovation processes within technological domains – with higher variance often associated with emerging or rapidly evolving fields – has long attracted the broader innovation literature (Jalonen and Lehtonen, 2011; Jalonen, 2012; Allen, 2013). This shows the potential of our setting, whose stochastic nature makes it possible to characterize not only average success dynamics but also how fluctuations and cross-category comovements evolve over time.

Notice that the framework abstracts from exogenous latent common factors. Therefore, the result below should be interpreted as a model-based implication of the interaction structure Γ, rather than as a claim that all observed comovement in the data must necessarily arise from direct interactions alone. The next theorem is a consequence of Theorem 3.1 and it characterizes how volatility and cross-category dependence evolve over time.

Theorem 3.3. *Under assumptions ($\mathcal{A}1$) and ($\mathcal{A}2$), denote by $\gamma^* \in (0,1]$ the Perron–Frobenius eigenvalue of Γ and by $\mathbf{u} = (u_h)_h$ its corresponding (unique up to a multiplicative non-zero constant) left eigenvector. Then, for each $h \in \mathcal{H}$, we have*

$$t^{1-\gamma^*} Var[X_{t+1,h}] \longrightarrow |u_h|\, \alpha(\mathbf{u}),$$

where $\alpha(\mathbf{u}) > 0$ (and such that $\alpha(C\mathbf{u}) = \alpha(\mathbf{u})/|C|$ for each constant $C \neq 0$). Moreover, for each pair $h \neq j$ of different processes, we get

$$t^{2(1-\gamma^*)} cov(X_{t+1,h}, X_{t+1,j}) \longrightarrow u_h u_j \sigma^2(\mathbf{u}),$$

where $\sigma^2(\mathbf{u}) > 0$ (and such that $\sigma^2(C\mathbf{u}) = \sigma^2(\mathbf{u})/C^2$ for each constant $C \neq 0$), so that

$$t^{1-\gamma^*} \rho(X_{t+1,h}, X_{t+1,j}) \longrightarrow \sqrt{u_h u_j}\, \frac{\sigma^2(\mathbf{u})}{\alpha(\mathbf{u})}.$$

Hence, the ratio of the variances $Var[X_{t,h}]/Var[X_{t,j}]$ converges to the ratio u_h/u_j of the relative eigenvector centrality scores (with respect to $\Gamma^\top$) of the two nodes (processes) h and j. Moreover, the correlation coefficient between the observations related to any pair of different processes converges to zero at the rate $1/t^{(1-\gamma^*)}$ when $\gamma^* < 1$ and converges to the same strictly positive value when $\gamma^* = 1$.

In short, a balanced system ($\gamma^* = 1$) supports persistent uncertainty and correlated success across technological categories. By contrast, an unbalanced system ($\gamma^* < 1$) leads over time to more asymmetric and progressively decoupled innovation dynamics. In other words, when interaction is strong enough to keep the system balanced, categories continue to display common fluctuations and persistent comovement; when reinforcement is subcritical, cross-category correlations vanish asymptotically and success dynamics become increasingly separated across categories.

The interpretation that follows, moreover, is consistent with the literature emphasizing that innovation outcomes depend on technological proximity and network linkages rather than evolving in isolation (Jaffe, 1989; Taalbi, 2020).

Finally, in order to extend the analysis from category-level success counts to source-target success counts, we develop Theorem 3.4. As

explained in the following, this theorem examines not only how many successful patents there are in each target category h, but also traces the source category k to which each successful patent originally belonged. Namely, in our particular context, we can enrich the model assuming that the number $S_{t,k,h}$ of successes until time-step t in category (process) h coming from belonging category k is of the form $S_{t,k,h} = \sum_{n=1}^{t} X_{n,h}Y_{n,k}$, where $Y_{n,k}$ takes value 1 if the category of patent n is k and 0 otherwise and

($\mathcal{A}3$) $\mathbf{Y}_n = (Y_{n,k})_{k\in\mathcal{H}}^{\top}$ is independent of $\mathbf{X}_n = (X_{n,h})_{h\in\mathcal{H}}^{\top}$ and of all the past until time-step $n-1$ with $P(Y_{n,k} = 1) = \pi_k \in (0,1)$ (where $\sum_{k\in\mathcal{H}} \pi_k = 1$).

Assumption ($\mathcal{A}3$) should be interpreted as a benchmark introduced for tractability. Its purpose is to isolate the role of the composition of patent activity across source categories, rather than to provide a full empirical description of source-category assignment. In particular, the assumption abstracts from the possibility that patents sort endogenously across source categories and from the possibility that source-category composition co-evolves with success dynamics.

Under these assumptions, we obtain the following result.

Theorem 3.4. *Under assumptions ($\mathcal{A}1$), ($\mathcal{A}2$) and ($\mathcal{A}3$), denote by $\gamma^* \in (0,1]$ the Perron–Frobenius eigenvalue of Γ. Then, for each pair $h,k \in \mathcal{H}$, we have*

$$\frac{S_{t,k,h}}{t^{\gamma^*}} \xrightarrow{a.s.} \pi_k S_{\infty,h}$$

and so $S_{t,k,h}/S_{t,j,h} \xrightarrow{a.s.} \pi_k/\pi_j$.

Theorem 3.4 provides a formal representation of how the composition of innovation inputs shapes the distribution of successful outcomes across technological domains. Specifically – under the benchmark assumption ($\mathcal{A}3$)– it shows that, asymptotically, the share of successful innovations in a target category reflects the origin distribution of patent activity. This result helps clarify how early category composition can propagate into long-run patterns of successful innovation across interacting technological domains.

The remaining material in this section provides visual illustrations of two implications of the framework that are especially relevant for the innovation context studied in the paper. First, because categories interact through Γ, localized increases in success can propagate across related categories. Second, because the framework is stochastic, it yields predictions not only for cumulative success dynamics but also for the evolution of volatility and cross-category comovement. These simulations help make the dynamic logic of the model more transparent before turning to the patent data.

3.1. Model implication: simulations with a shock

It is worthwhile to illustrate what the model implies for propagation conditional on the cross-category interaction structure Γ. If innovation success is cumulative and interdependent, then a shock affecting one category should not remain confined to that category, but should transmit to other connected categories through the inter-dependent reinforcement mechanism. The simulation below visualizes this implication of the model showing how the effect of a *shock* to one process propagates to the others through the interaction structure. A model that accounts for the propagation of innovation shocks, moreover, is useful because such spillovers are frequently observed in practice.[8]

We consider here a system with $N = 2$ processes starting with the same initial composition, $\theta_h = 1/2$ and $c_h = 1$. The purpose is to analyze two simple benchmark interaction structures. The first is a symmetric benchmark, in which the two categories are equally connected and spillovers are homogeneous across them. The second is an asymmetric benchmark, in which cross-category effects are directional and one category is more central in the interaction structure than the other. In this way, the simulation illustrates how shock propagation depends not only on the presence of interdependence, but also on its structure. We take into account the following two scenarios:

- mean-field interaction, that is a symmetric interaction matrix Γ of the form

$$\gamma_{j,h} = \begin{cases} \gamma^*(\iota/N + (1-\iota)) & \text{if } j = h \\ \gamma^*\iota/N & \text{if } j \neq h, \end{cases} \tag{2}$$

  with $\gamma^*, \iota \in (0,1]$,[9] where we have chosen $\gamma^* = 0.7$ and $\iota = 0.9$, so that

$$\Gamma = \begin{pmatrix} 0.385 & 0.315 \\ 0.315 & 0.385 \end{pmatrix};$$

- non-symmetric interaction matrix Γ with $\gamma^* = 0.685$ and $\mathbf{u} = (1.394, 0.575)^{\top}$, i.e.

$$\Gamma = \begin{pmatrix} 0.50 & 0.20 \\ 0.45 & 0.20 \end{pmatrix}.$$

The *shock* occurs at time $t_{shock} = 10^4$ and it acts as follows: at time $t_{shock} = 10^4$, the probability of having a success for process 1 is increased (that is, we give a positive shock to process 1) by replacing the parameters $\theta_1 = 1/2$ and $c_1 = 1$ by the new ones $\theta_{shock,1} = 1/2 + 10^4$ and $c_{shock,1} = 1 + 10^4$. These new values of the parameters remain in the dynamics of process 1 for all subsequent time-steps until $t = 10^7$. The parameters for process 2 are unchanged: $\theta_2 = 1/2$ and $c_2 = 1$ remain the same throughout. The number of successes for the two processes, namely $S_{t,1}$ and $S_{t,2}$, are then observed. Fig. 1 shows that the positive shock causes the number of successes for process 1 to rise and that this effect also propagates to process 2 through the interaction terms in the dynamics. It also shows that, although the shock shifts the levels of the trajectories, it does not alter their long-run scaling properties, which remain governed by the interaction structure summarized by Γ.

3.2. Model implication: simulations on variance and covariances

In this section, we illustrate through simulations the implications of Theorem 3.3 in the empirically relevant unbalanced case $\gamma^* < 1$. The point is to show that the stochastic predictions of the model are not limited to asymptotic statements: even in finite simulations, the framework implies that fluctuations in category-specific success and cross-category comovement weaken over time when reinforcement is subcritical. In other words, as successful innovations become rarer per patent opportunity, uncertainty at the category level also declines and success dynamics across categories become progressively less correlated. Fig. 2 shows these implications in the same two benchmark interaction structures considered in Section 3.1.

Specifically, we run 10^4 independent simulations of $N = 2$ processes up to a final time of $t = 10^4$, considering the scenarios of mean-field and non-symmetric interaction outlined in Section 3.1. In Fig. 2, we observe that the variances of both the processes $(X_{t,1})_t$ and $(X_{t,2})_t$ decay according to a power law with exponent $1-\gamma^*$ (left panels). Similarly, the covariance between the processes $(X_{t,1})_t$ and $(X_{t,2})_t$ exhibits a

[8] A vivid example of this phenomenon is the surge in pharmaceutical innovation during the COVID-19 pandemic, which triggered subsequent innovations in other fields such as logistics (Dovbischuk, 2022) and IT (Li et al., 2022), and also prompted a wave of open innovation practices through which inventors widely disseminated their contributions (Brem et al., 2021; Lee and Trimi, 2021; Ho, 2023).

[9] We refer to this kind of interaction as the mean-field interaction, because in this case the term $\sum_{j\in\mathcal{H}} \gamma_{j,h}S_{t,j}$ in (1) becomes $\gamma^*\left(\iota\sum_{j\in\mathcal{H}} S_{t,j}/N + (1-\iota)S_{t,h}\right)$ and so the conditional probability $P_{t,h}$ depends on a convex combination between $S_{t,h}$ and the averaged value of all the $S_{t,j}$ in the system.

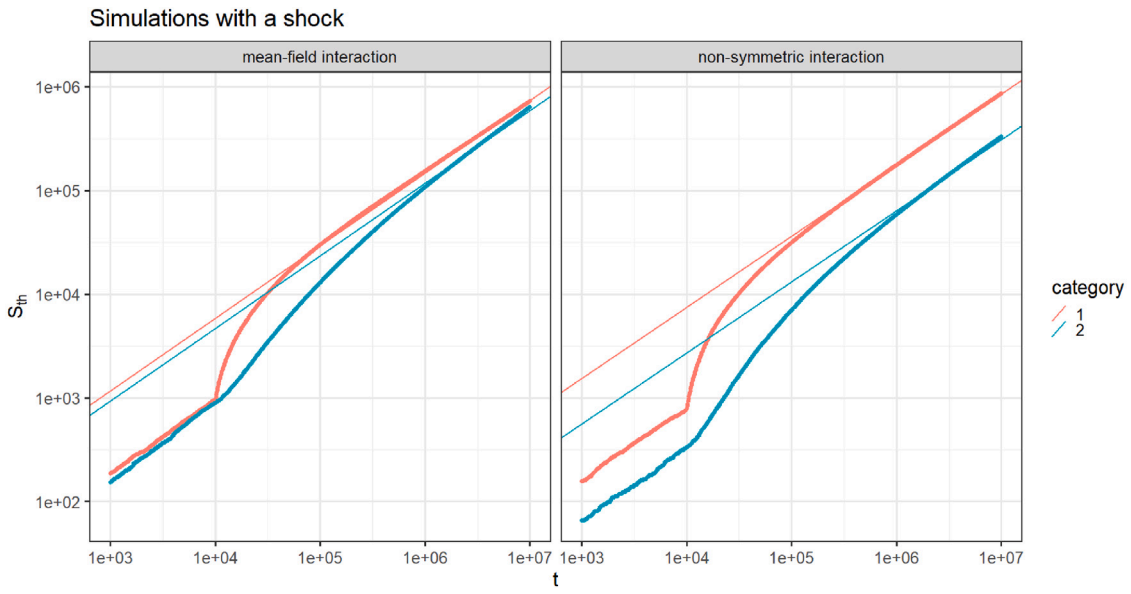


Fig. 1. Linear behavior of the successes counter $S_{t,1}$ (red line) and $S_{t,2}$ (blue line) along t, in $\log_{10} - \log_{10}$ scale, in two different scenarios with a positive shock on process 1. The two scenarios are distinguished by the type of interaction, i.e. by the form of the interaction matrix Γ used in the simulations. The shock on process 1 occurs at time $t_{shock} = 10^4$ and the system is observed until time-step $t = 10^7$. We can observe the sudden rise of $S_{t,1}$ (red line) caused by the shock and how the positive effect of the shock on process 1 propagates to process 2 (consequent progressive rise of the blue line). We can also note that the slope of the two lines and the distance between their intercepts are the same before and after the shock, because the interaction matrix Γ does not change: indeed, the slope is equal to γ^* and the distance between the intercepts corresponds to the quantity $\log_{10}(u_1/u_2) = \log_{10}(u_1) - \log_{10}(u_2)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

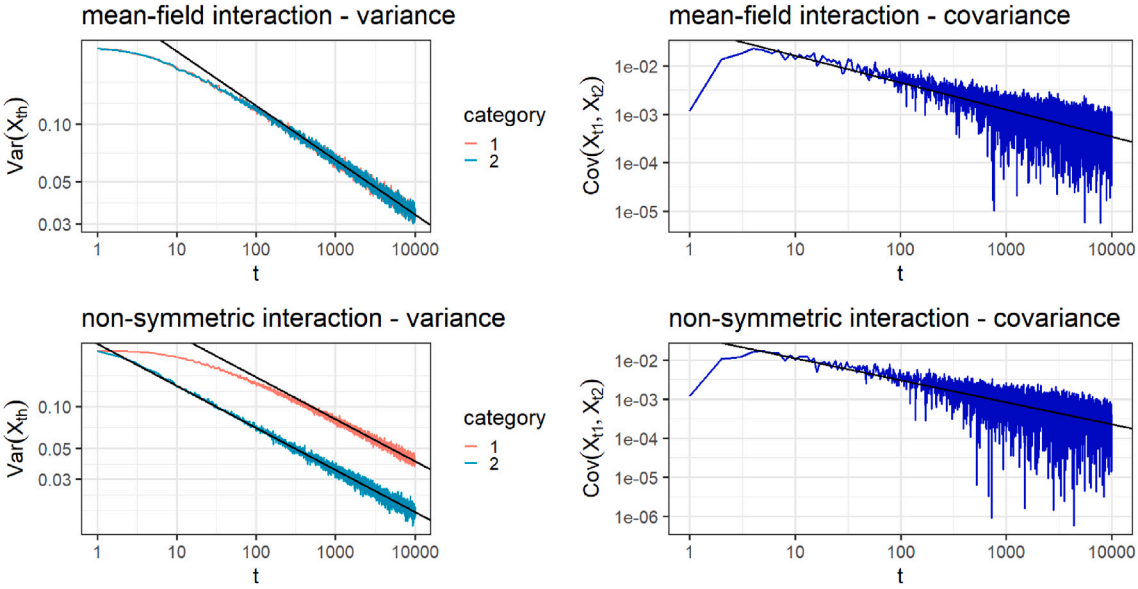


Fig. 2. Linear behavior of $Var[X_{t,h}]$ (left panels) and $cov[X_{t,1}, X_{t,2}]$ (right panels) along t, in $\log_{10} - \log_{10}$ scale, in two different scenarios. These quantities have been estimated by performing 10^4 independent simulations and the system is observed until time-step $t = 10^4$. The two scenarios are distinguished by the type of interaction, i.e. by the form of the interaction matrix Γ used in the simulation.

comparable asymptotic behavior, with an exponent of $2(1-\gamma^*)$ (right panels).

The plots are presented in a log–log scale, where the power-law decay appears as linear behavior, with the exponent corresponding to the slope of the fitted line obtained from linear regression. In the left panels of Fig. 2, the difference between the intercepts can be interpreted as the logarithmic ratio between the variances of the two categories, i.e., $\frac{\mathrm{Var}[X_{t+1,1}]}{\mathrm{Var}[X_{t+1,2}]}$, which converges to $\frac{u_1}{u_2}$.

In the first scenario of mean-field interaction, the estimated common slope of −0.286 in Fig. 2 (top-left panel) is in excellent agreement with the theoretical value of $-(1-\gamma^*) = -0.3$. In the second scenario of nonsymmetric interaction, the estimated common slope (bottom-left panel) is −0.302, which is also very close to the theoretical value of $-(1-\gamma^*) = -0.315$. The difference in intercepts of the two processes in the two scenarios is 0.00023 and 0.367, respectively, which aligns well with the theoretical predictions: $\log_{10}(u_1/u_2) = 0$ for the mean-field case, and $\log_{10}(u_1/u_2) = \log_{10}(1.394/0.575) = 0.385$ in the nonsymmetric interaction scenario.

For the covariances, Fig. 2 shows that the estimated slope in the mean-field scenario (top-right panel) is −0.559, which is in very close

agreement with the theoretical value of $-2(1-\gamma^*) = -0.6$. Similarly, in the nonsymmetric interaction scenario, the estimated slope (bottom-right panel) is -0.563, which is in good agreement with the theoretical value of $-2(1-\gamma^*) = -0.629$.

Overall, the simulations confirm that, as successful innovations become rarer per patent opportunity, innovation dynamics across categories also become progressively less correlated. This complements the earlier results on sublinear growth and stable long-run shares, and helps clarify the broader implication of the model: cumulative and interdependent innovation need not remain equally volatile or tightly coupled as the system expands.

4. Data analysis

Starting from the category-specific success indicators defined in Section 2, we now turn to their empirical evolution in the patent data following the results derived in Section 3. The most direct quantities to examine are the cumulative success counts and their relative behavior across categories, because these are the observable objects through which the model yields its main aggregate implications. We therefore study their scaling properties, their long-run proportions, and their source-target decomposition, and then report a complementary inference exercise on interaction intensity under a mean-field restriction.

Our dataset consists of the GLOBAL PATSTAT database.[10] More specifically, we collected all the granted US patent families (DOCDB) published in the period [1980 – 2018], with their exact (full) date of publication and their CPC-1 category. Moreover, for each patent n, we know if it has been cited by subsequent patents (published in the considered period) and which are the citing patents. As observed in Section 2, patents may be associated with multiple CPC-1 classes,[11] and this multiplicity is preserved when constructing target-category citation counts. By contrast, measures based on the source category, which require a unique origin, assign each patent to the modal first-level CPC letter among its CPC codes.[12] In the considered dataset, ties were very rare and were resolved by choosing the first CPC code in PATSTAT.

We limit ourselves to the CPC categories $h \in \mathcal{H} = \{A, B, C, D, E, F, G, H\}$ (i.e. we exclude category Y[13]). We take the size T of the time-window for the computation of the index $I_{n,h}$ equal to 5 years (coherently with other previous studies, e.g. Squicciarini et al. (2013)). Therefore, since the available dataset includes patents with publication year less than or equal to $T_f = 2018$, we perform the analysis on the patents with publication year in the period between 1980 and $(T_f - T) = 2013$, so that, for all of them, we have the entire time-window to compute the index. As explained in Section 2, this is done in order to neutralize the mechanical right-truncation problem associated with unequal citation windows. Indeed, in our analysis, all the considered patents have the *same* time-window of $T = 5$ years to receive citations. Finally, we fix the threshold $\tau = 0.8$ (see Appendix D for some details on the choice of the threshold). We thus obtain a matrix with $N = \text{card}(\mathcal{H}) = 8$ columns, where each row corresponds to a patent n. The patents (and so the rows) are ordered with respect to their publication (full) date. The total number of rows is $n_{tot} = 5\,004\,253$. The matrix entry $x_{n,h}$ is equal to 1 if the patent n is a success for category h, i.e. if its value of the index is above the threshold τ, or equal to 0 if n is not a success. Hence, the constructed matrix can be seen as the realization of a finite system $\{X^{(h)} = (X_{n,h})_n : h \in \mathcal{H}\}$ of Bernoulli processes. The patent ordering induces the relevant notion of *system time* used throughout the paper, and the remainder of the section asks whether the aggregate behavior of the real data $\{x_{n,h}\}$ is in agreement with the theoretical results of the previous section.

4.1. Aggregate empirical patterns and their connection with the model

We begin from the cumulative success counts, since these are the observable quantities that most directly correspond to the main implications of the model. Theorem 3.2 predicts that, under the maintained assumptions, the number of cumulative successes in each category should grow with the same Heaps' exponent γ^*. Fig. 3 examines this first implication in the patent data.

Fig. 3 provides the asymptotic behavior, in $\log_{10} - \log_{10}$ scale, of every process $S_{t,h}$, that represents the number of successes observed in category h until time-step (patent) t. We can appreciate how the lines exhibit the same slope, which indicates that the processes have the same Heaps' exponent. This is exactly in accordance with Theorem 3.2. The emergence of a common Heaps' exponent across all categories is therefore consistent with the model prediction that category-specific cumulative successes share a common long-run scaling parameter. At the same time, and coherently with the discussion in Section 3, this pattern should be interpreted with some care: it is compatible with the presence of systemic interdependence, although it does not by itself distinguish direct cross-category interaction from co-movement induced by broader common factors. We provide here, therefore, evidence consistent with the view that innovation success across categories evolves in an interdependent way rather than as a collection of completely separate category-specific processes.

The value of the common Heaps' exponent, estimated as the common slope of the lines in the $\log_{10} - \log_{10}$ plot, is $\hat{\gamma}^* = 0.689$.[14] The estimate $\hat{\gamma}^* = 0.689$ therefore implies sublinear growth of cumulative successes in system time. Interpreted through the model, this means that the probability that the next observed patent is a success declines as the number of patents already observed increases. Put differently, successful patents continue to occur, but they become less frequent relative to the expanding set of patent opportunities. This interpretation refers to system time rather than calendar time: Fig. 3 does not imply that yearly success rates must decline for purely mechanical reasons, but that cumulative success grows less than proportionally with the scale of inventive activity. This is the economically relevant sense in which success becomes harder to obtain as the system expands. More broadly, the common sublinear pattern across categories points to a form of bounded scalability in innovation outputs: innovation continues, but the marginal probability of producing an additional success in any category diminishes, consistently with the literature on declining research productivity and increasing knowledge burdens (Bloom et al., 2020; Jones, 2009).

The second empirical implication emphasized in the Introduction and derived in Theorem 3.2 concerns not the common growth exponent itself, but the long-run relative proportions of successes across categories. If cumulative success counts scale with the same exponent, then their ratios should stabilize over time. Fig. 4 examines this implication.

[10] The dataset is maintained at IMT School for Advanced Studies Lucca.

[11] Approximately 35% of the patents in the dataset belong to multiple CPC-1 classes.

[12] Given a patent n belonging to more than one category, we take as source category c_n the modal first-level CPC letter among its CPC codes; while, if a patent n' belonging to more than one category cites patent n, then we count patent n' in the computation of the index $I_{n,h}$ for each category h to which patent n' belongs.

[13] This category has been excluded due to its broad scope and ambiguity, which are well-documented in the literature (Leydesdorff et al., 2017; Rainville et al., 2025). Although it is often broadly referred to as the class encompassing green patents, it is widely recognized in the green innovation literature that only the subclasses $Y02$ and $Y04$ accurately represent genuine green technologies (Corrocher and Mancusi, 2021; Barbieri et al., 2023).

[14] In order to measure the uncertainty of the slopes estimated for each category h at a given time-step, we have performed a linear regression with random effects, on both intercept and slope and using a subsample of size 10^4 of the data set. The estimated values of the slopes, one for each category h, present a mean equal to $\hat{\gamma}^* = 0.689$ and standard deviation equal to 0.044.

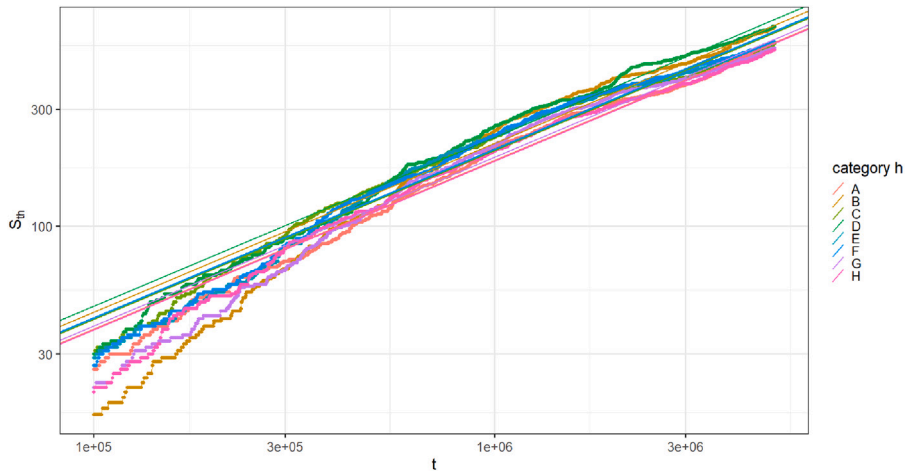


Fig. 3. Linear behavior of the successes counter $S_{t,h}$ along t, in $\log_{10} - \log_{10}$ scale, for each category h. The lines are obtained by a least square interpolation using a subsample of size 10^4 of the data set. We can appreciate how the lines exhibit the same slope, which has been estimated to be equal to $\hat{\gamma}^* = 0.689$. Indeed, the goodness of fit R^2 index obtained imposing a common slope is 0.969 and so basically equal to the one obtained allowing the slopes to be different across categories, i.e. 0.972. The emergence of a common Heaps' exponent across all categories is suggestive of shared interdependence, i.e. of the fact that these innovation processes are not evolving independently but are shaped by shared systemic dynamics. Moreover, the sub-linear growth reflects diminishing returns to cumulative knowledge, increasing innovation complexity, and a bounded scalability in innovation outputs over time. The figure is plotted in system time, so the estimated sublinear scaling should be interpreted as a decline in success per patent opportunity rather than as a mechanical statement about calendar-time trends.

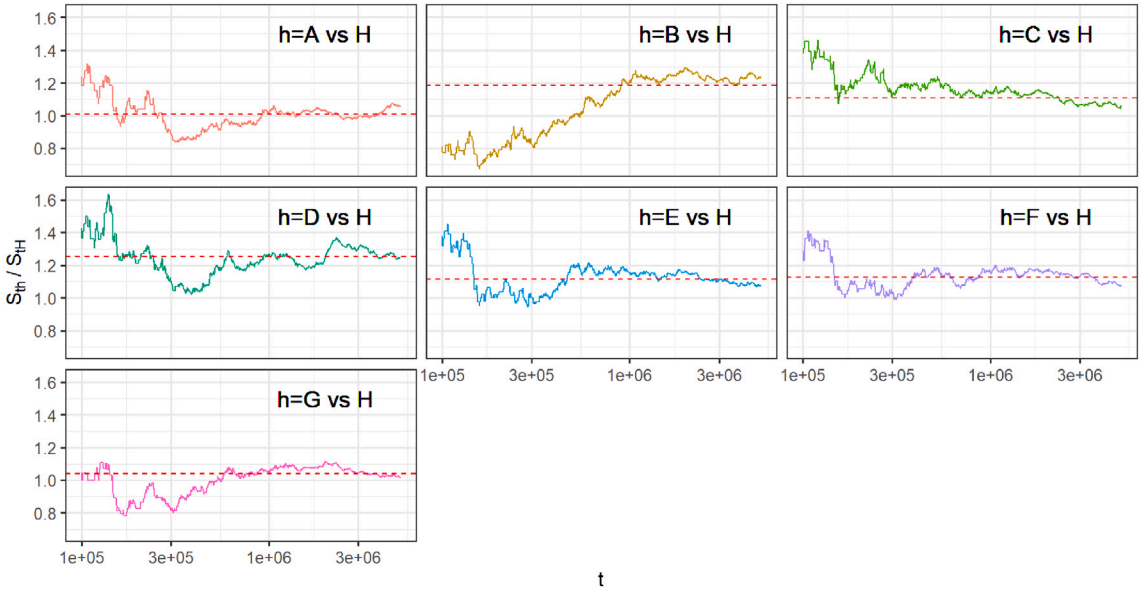


Fig. 4. Plot of the process $S_{t,h}/S_{t,H}$ along time t, for all the categories $h \neq H$. Category H is arbitrarily chosen to be the baseline category. The horizontal dashed red lines represent an estimation of the ratio u_h/u_H, obtained as $10^{d_{h,H}}$, where $d_{h,H}$ is the difference between the intercepts of the regression lines in Fig. 3 for the categories h and H. This long-run convergence across categories highlights a structural stabilization in innovation outcomes, despite increasing complexity and declining marginal returns to knowledge production. Moreover, note that all the estimated ratios u_h/u_H are around 1. This means that in the long run the numbers of successes in the different categories tend to coincide.

Fig. 4 shows, for each category h, the convergence of the process $(S_{t,h}/S_{t,H})_t$ toward the quantity u_h/u_H, estimated as $10^{d_{h,H}}$ where $d_{h,H}$ is the difference between the intercepts of the regression lines in Fig. 3 for the two processes related to the pair (h, H) of categories.[15] Also this fact is in accordance with Theorem 3.2 and it means that the relative number of successes across categories stabilizes over time. In other words, although categories may differ in levels, their long-run relative success shares tend to settle around stable proportions. This interpretation is in line with work emphasizing that technological interdependencies shape persistent patterns of innovation across domains (Taalbi, 2020), even when producing additional influential innovations becomes progressively harder over time (Bloom et al., 2020). Moreover, the limit corresponds to the ratio of the respective components of the eigenvector-centrality scores $\mathbf{u}$ (with respect to $\Gamma^\top$).

Taken together, Figs. 3 and 4 provide the empirical counterpart of the two main aggregate implications highlighted in the Introduction: cumulative successes grow sublinearly in system time and their relative shares across categories stabilize in the long run. For this reason, they constitute the core empirical evidence of the paper.

[15] Category H is arbitrarily chosen to be the baseline category. Any other category can be used as a baseline.

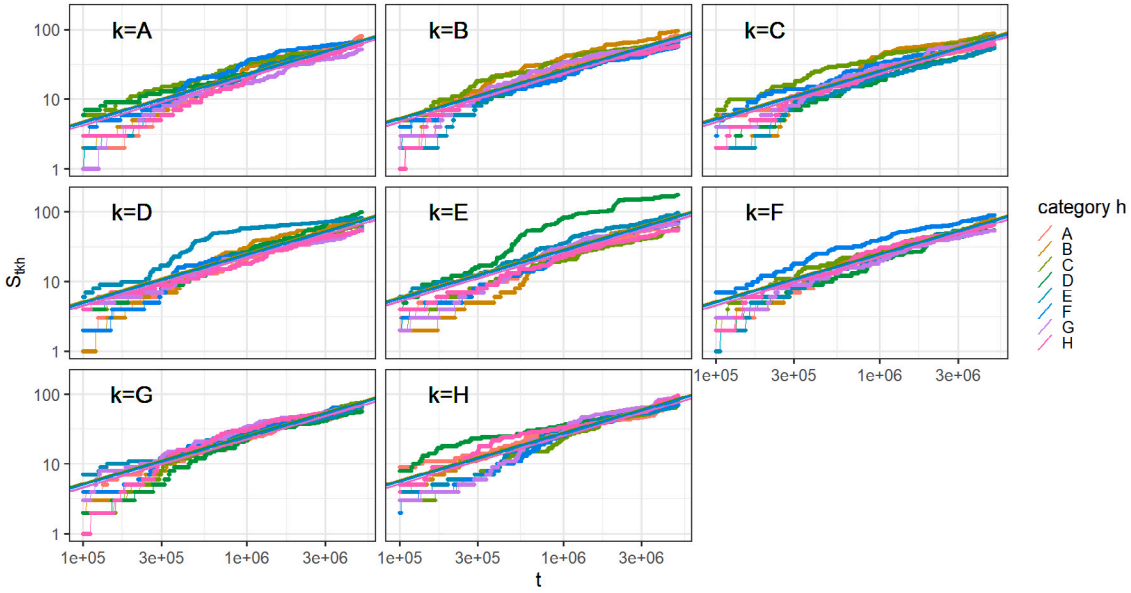


Fig. 5. Linear behavior of the successes counter $S_{t,k,h}$ along t in $\log_{10} - \log_{10}$ scale, for each pair (k,h) of categories (k =sub-figure, h =color). The lines are obtained by a least square interpolation based on a subsample of the dataset with a slope equal to the previously estimated value $\widehat{\gamma}^* = 0.689$. The goodness of fit R^2 index is 0.851, which is basically the same as the one obtained allowing the slopes to be different across the pairs (k,h) of categories, i.e. 0.855. The common slope suggests once again the presence of interaction across the categories. Moreover, the sub-linear growth reflects a consistent sub-linear scaling of cross-domain innovation flows, suggesting that while knowledge recombination persists, its marginal productivity diminishes over time across all category pairs.

We now consider an extension that tracks successful patents jointly by target category and source category. This part of the analysis is included because Theorem 3.4 refines the aggregate cumulative-success dynamics by showing how they decompose according to the source category from which successful patents originate. In this sense, the following evidence is not a separate empirical claim, but a finer decomposition of the aggregate patterns already documented above.

Fig. 5 provides the asymptotic behavior, in $\log_{10} - \log_{10}$ scale, of every process $S_{t,k,h}$, that represents the number of successes for category h coming from category k observed until time-step (patent) t. We can appreciate how the lines exhibit the same slope, equal to the previously estimated value $\widehat{\gamma}^* = 0.689$ in accordance with Theorem 3.4. In this sense, the source-target decomposition inherits the same aggregate scaling structure documented at the category level.

Finally, each panel of Fig. 6 refers to the successes observed in a given category h and shows, for each coming category k, the convergence of the process $S_{t,k,h}/S_{t,H,h}$ toward the quantity π_k/π_H, estimated as 10 to the power of the difference between the intercepts of the corresponding regression lines in Fig. 5.[16] Also this fact is coherent with Theorem 3.4. In the long run, therefore, the number of successful innovations in each category ends up matching how much patent activity started in each category. Hence, if a certain type of technology had a greater patent activity early on, it will also tend to have more successful innovations over time.[17]

Although the main empirical analysis of this section focuses on cumulative counts and long-run proportions, it is useful to recall that Section 3 also derived an additional implication concerning fluctuations and cross-category correlations in the empirically relevant unbalanced case $\gamma^* < 1$. In the theoretical development, this result serves to show that the stochastic framework yields implications not only for average success dynamics but also for the way uncertainty and comovement evolve over time. In the empirical logic of the paper, this implication complements the two core aggregate patterns discussed above by clarifying why sublinear cumulative growth and stable long-run shares do not imply persistent synchronization in category-specific success dynamics.

4.2. Statistical inference on the interaction intensity under the mean-field assumption

The evidence reported so far compares the patent data with the aggregate implications of the model under a general interaction structure. We now add a complementary exercise under a more restrictive specification. Namely, assuming the system is subjected to a mean-field interaction, we can summarize cross-category dependence by a single average interaction parameter and perform a tractable benchmark inference exercise on its magnitude. This part of the section is included for a different purpose than the figures above. The previous evidence asks whether the aggregate patterns predicted by the framework are present in the data; the present exercise asks whether the observed category-specific success counts move similarly enough to be compatible with a given level of average interaction intensity under a symmetry restriction.

A standard approach in innovation studies is to construct measures of technological proximity or network relatedness and then use them in regression or forecasting exercises to assess their effect on innovative performance (Jaffe, 1989; Pichler et al., 2020; Taalbi, 2020). By contrast, under the mean-field specification, our test is model-based: it assesses whether the observed category-specific success counts move similarly enough to be compatible with a given level of average interaction intensity. In this sense, the contribution is not to show that technological interdependence matters — which is well established in the literature — but to provide an inferential procedure that is structurally tied to the interacting reinforced Bernoulli framework.

Namely, assuming the system is subjected to a mean-field interaction (note that in this case we have $u_h/u_j = 1$ for each $h,j \in \mathcal{H}$), i.e. assuming the interaction matrix Γ to be of the form (2), we can perform a suitable test on the parameter ι, that rules the intensity of

[16] Category H in Fig. 6 is arbitrarily chosen to be the baseline category. Any other category can be used as a baseline.

[17] These patterns are consistent with the benchmark implication in Theorem 3.4, although that result relies on the stylized assumption ($\mathcal{A}3$)and should therefore be read as an illustrative benchmark rather than as a full structural account of source-category assignment.

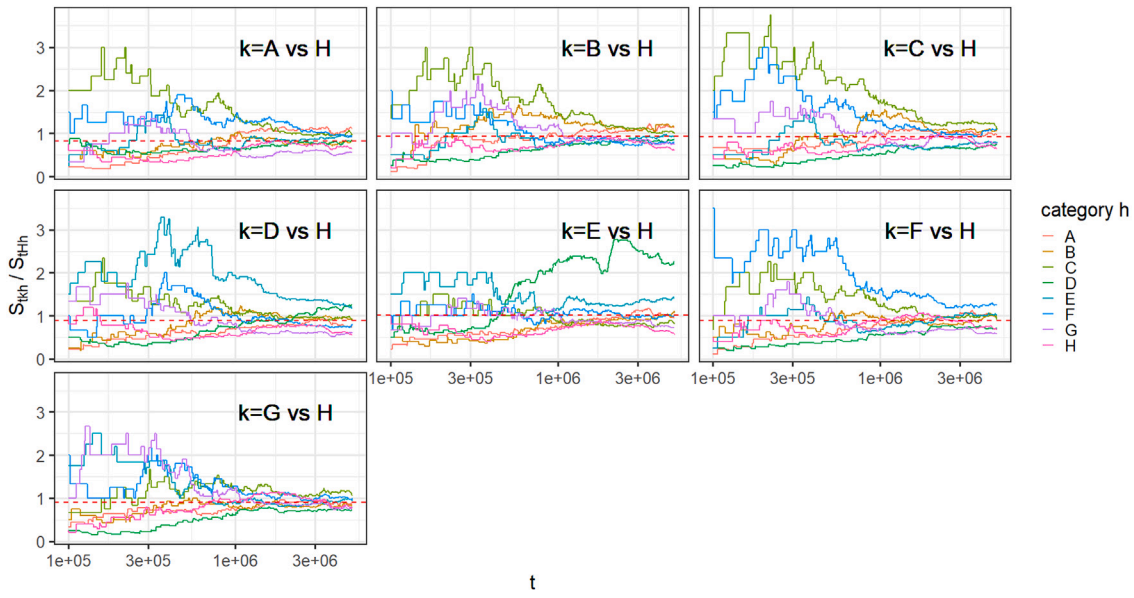


Fig. 6. Plot of the process $S_{t,k,h}/S_{t,H,h}$ along time t, for all the pairs (k,h) of categories (k =sub-figure, h =color), with $k \neq H$. Category H is arbitrarily chosen to be the baseline category. The horizontal dashed red lines represent the value 10 to the power of the differences between the intercepts of the regression lines in Fig. 5 for the pairs of categories (k,h) and (H,h). This long-run proportionality links early patenting intensity to long-term innovation success, highlighting the path-dependent and self-reinforcing nature of technological trajectories.

the interaction. This inferential exercise should be interpreted as a parsimonious benchmark under a symmetry restriction on the interaction structure. Under the mean-field specification, heterogeneous pairwise technological linkages are summarized by a single average interaction parameter, rather than being allowed to vary across category pairs. The restriction is useful for tractable model-based inference, but it is not intended as a full description of the heterogeneous and directional interdependencies emphasized in the innovation literature (Jaffe, 1986; Breschi et al., 2003; Colladon et al., 2025)

Notice that, under the mean-field specification, the parameter ι measures how strongly categories are tied together through a common interaction component. The test checks whether the category-specific success counts move similarly enough to be compatible with a given level of average interaction across categories.

For a two-sided test with $H_0 : \iota = \iota_0$, where $\iota_0 \leq 1$ and $\iota_0 > 1/2$ (the required condition in the theoretical result for having the second eigenvalue of Γ strictly smaller than $\gamma^*/2$), we can use the test statistic (see Appendix B for the technical details)

$$2\Delta_0 \frac{\left\|\mathbf{S}_t - \widetilde{S}_t \mathbf{1}\right\|^2}{\widetilde{S}_t} \overset{d}{\sim} \chi^2(N-1) \quad \text{under } H_0, \tag{3}$$

where $\Delta_0 = \iota_0 - \frac{1}{2}$ and $\widetilde{S}_t = \frac{\mathbf{1}^\top \mathbf{S}_t}{N} = \sum_{h\in\mathcal{H}} S_{t,h}/N$. Since this statistic is increasing in ι_0, it works well also for the one-sided test with $H_0 : \iota \geq \iota_0$, with $\iota_0 \in (1/2, 1]$. Applying this one-sided test to our real dataset of patents, we obtain different p-values, one for each tested value ι_0, that are collected in Table 1. We can see that, if we choose the significance level $\alpha = 0.05$, then the minimum value of ι_0 at which we can reject the null hypothesis is $\iota_0 = 0.75$.

Taking the value of the parameter γ^* equal to the estimated one, i.e. $\widehat{\gamma}^* = 0.689$, we have also computed an estimate of the parameter of interest ι maximizing the (joint) likelihood of the processes $X^{(h)} = (X_{n,h})_n$, $h \in \mathcal{H}$, i.e.

$$L(\iota \mid x_{n,h},\, h \in \mathcal{H},\, n = 1, \dots, n_{tot}) = \prod_{n=1}^{n_{tot}} \prod_{h\in\mathcal{H}} P_{h,n}^{x_{h,n}} (1 - P_{h,n})^{1-x_{h,n}}.$$

The obtained estimate is in accordance with the results of Table 1. Indeed, the maximum-likelihood estimate of the interaction intensity ι is $\widehat{\iota} = 0.643$ and from Table 1 we get that, at the level $\alpha = 0.05$, we can reject the null hypothesis $H_0 : \iota \geq 0.75$, while we do not have enough statistical evidence to reject for smaller values of ι_0.

Table 1
p-values associated to the hypothesis test with $H_0 : \iota \geq \iota_0$ under the mean-field interaction .

ι_0	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
p-value	0.893	0.562	0.273	0.113	0.042	0.015	0.005	0.002	< 0.001

The estimated value $\hat{\iota} = 0.643$ points to a positive but non-maximal degree of average interaction across technological categories. Read together with the hypothesis-test results in Table 1, this suggests that the patent data are not compatible with a very high degree of homogeneous interaction across categories, while they are clearly inconsistent with the view that categories evolve independently. In substantive terms, the evidence points to an intermediate configuration: cross-category spillovers are strong enough to generate co-movement in innovative success, but not so strong as to eliminate category-specific dynamics. This interpretation is consistent with the innovation literature showing that technological interdependence and relatedness matter for innovative performance, while substantial heterogeneity across domains remains an important empirical feature (Jaffe, 1989; Taalbi, 2020; Pichler et al., 2020). Importantly, our contribution is not to compare $\hat{\iota}$ to a standard coefficient already used in that literature, but to provide a model-based estimate of interaction intensity that is structurally tied to the interacting reinforced Bernoulli framework.

5. Conclusions

This paper has asked whether innovation success across technological categories can be represented within a single stochastic framework that jointly captures cumulativeness within categories and spillovers across categories. To address this question, we have proposed a model of interacting reinforced Bernoulli processes in which the probability that the next observed patent is a success in category h depends on past successes both within that category and across the others.

The contribution of the paper is to bring together, within one (novel) parsimonious dynamic framework, cumulative reinforcement,

cross-category interdependence, and aggregate regularities in innovation outcomes. The value added of the framework, in particular, is that it derives, within a single stochastic structure, joint implications for success probabilities, cumulative successes, relative success shares, and cross-category dependence.

On the empirical side, using granted US patent families from GLOBAL PATSTAT classified at the CPC-1 level, we have constructed category-specific success indicators based on a cohort-normalized forward-citation measure. The empirical analysis shows that the patent data are consistent with the two main aggregate implications emphasized in the paper: cumulative successes grow sublinearly in system time, and their relative shares across categories stabilize in the long run. We have also shown, under the benchmark source-target extension, that the same aggregate scaling logic carries over to the decomposition of successful patents by source category. In addition, under a mean-field restriction, we provide a model-based inference exercise indicating a positive but non-maximal degree of average interaction across categories.

These findings support the paper's core view of innovation success as cumulative and interdependent, while still allowing for category-specific dynamics. They also reinforce an important interpretive point of the analysis: the sublinear result is formulated in *system time*, so it should be read as success becoming rarer per patent opportunity as the system expands, not as a mechanical claim about calendar-time trends. More broadly, the stochastic structure of the framework makes it possible to study not only average success trajectories but also fluctuations, comovement, and shock propagation.

The analysis remains subject to limitations. The interaction matrix is treated as reduced-form and is not itself endogenized; the source-target result is derived under a stylized benchmark assumption introduced for tractability; and the inferential exercise under the mean-field specification is intentionally restrictive, since it summarizes heterogeneous interactions through a single average parameter. Moreover, while the framework abstracts from exogenous latent common factors, some observed co-movement in the data could in principle also reflect broader common influences rather than direct interaction alone.

These limitations also point to directions for further research, including applications based on alternative taxonomies or alternative indicators of innovative influence and inference for richer interaction structures and empirical testing of the behavior along time of the fluctuations and of cross-category comovement of the observations. In other words, it could be interesting to investigate how the variability of the observations $X_{t,h}$ around their local mean, and their correlations across categories, evolve over time by devising measures of dispersion within rolling windows.

Overall, the paper provides a parsimonious and empirically interpretable framework for studying how successful innovations accumulate across technological categories in the presence of reinforcement and interdependence. By linking local success dynamics to observable aggregate regularities in patent data, it contributes to the broader effort to formalize innovation as a cumulative, interdependent, and stochastic process.

CRediT authorship contribution statement

Giacomo Aletti: Writing, Methodology, Formal analysis, Data analysis, Conceptualization. **Irene Crimaldi:** Writing, Methodology, Formal analysis, Data Analysis, Conceptualization. **Andrea Ghiglietti:** Writing, Methodology, Formal analysis, Data Analysis, Conceptualization. **Federico Nutarelli:** Writing, Methodology, Formal analysis, Data Analysis, Conceptualization.

Ethics statements

This research did not involve human participants, intervention, or personally identifiable information, and did not require ethics approval. No animals were used in this research.

Funding

The work is partially supported by the European Union – NextGenerationEU through the Italian Ministry of University and Research under the National Recovery and Resilience Plan (PNRR) – M4C2 – Investment 1.3, title Public Sector Indicators for Sustainability and Well- being (PUWELL) – Program Growing Resilient, INclusive and Sustainable (GRINS) – PE18 – CUP J33C22002910001. This research was supported by the "Resilienza Economica e Digitale" (RED) project (CUP D67G23000060001) funded by the Italian Ministry of University and Research (MUR) as Department of Excellence (Dipartimenti di Eccellenza 2023–2027, Ministerial Decree no. 230/2022).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are thankful to the editor and to the anonymous referees for their useful comments and suggestions.

Appendix A. Definitions of novelty in the literature

In patent law, novelty is a legal requirement: for a patent to be granted, the invention must be new relative to the prior art at the time of filing (Mueller, 2024; Bekkers et al., 2020). By contrast, the innovation literature often uses the term "novelty" more broadly to refer to concepts such as radicalness, unusual recombination, technological distance, distinctiveness, or ex post technological impact (Abbas et al., 2014; Fleming, 2001; Dahlin and Behrens, 2005; Verhoeven et al., 2016). In this appendix, we use "novelty" in that broader academic sense unless otherwise stated. The question is therefore not whether a patent is legally novel, but which observable features the literature uses to characterize patents as more or less novel in an economic and technological sense. Table A.2 provides a selective (non-exhaustive) overview of representative approaches, including citation-based, recombination-based, text-based, and composite measures. Given the richness of the literature about the subject, as aforementioned, the papers that address "patent novelty" adopt their own definitions of the term. Consequently, if we attempt to classify the main streams of literature on patent novelty, we may identify the following:

(i) The older and more established literature is the one that defines novelty based on the forward and backward citations (Trajtenberg et al., 1997; Squicciarini et al., 2013). Moreover, in the document about PATSTAT database we can read: "The number of citing patent documents can be an indicator of the importance of the patent. A frequently cited patent can be an indication of a core technology". This literature has been recently criticized. In Abbas et al. (2014), for instance, the authors primarily discuss patent novelty in the context of citation analysis, classification systems, and keywords. They highlight how backward citations (patents cited by a new patent) are commonly used to infer novelty—specifically, fewer backward citations can signal higher novelty because the patent draws less from existing technologies. However, they criticize citation-based methods, noting that they do not always capture the full semantic or functional novelty of a patent.

Table A.2
This Table lists the various novelty indices brought by the literature. As noted in Lanjouw and Schankerman (2004), however, no single indicator can fully capture the quality or value of a patent and multiple indicators (citations, claims, oppositions, family size) provide a more comprehensive and nuanced understanding of both the technological importance and the economic potential of innovations.

Source	Novelty Index	Quick explanation	Based on
Squicciarini et al. (2013)	$\text{Originality}_p = 1 - \sum_j^{n_p} s_{pj}^2$	s_{pj} is the percentage of citations made by patent p to patent class j out of the n_p IPC 4-digit (or 7-digit) patent codes contained in the patents cited by patent p. Citation measures are built on EPO patents and account for patent equivalents.	Citations
Squicciarini et al. (2013)	$\text{Radicalness}_p = \sum_j^{n_p} \frac{\text{CT}_j}{n_p}$; $\text{IPC}_{pj} \neq \text{IPC}_p$	CT_j denotes the count of IPC-4 digit codes IPC_{pj} of patent j cited in patent p that is not allocated to patent p, out of n IPC classes in the backward citations counted at the most disaggregated level available (up to the 5th hierarchical level). The higher the ratio, the more diversified the array of technologies on which the patent relies upon.	Citations Tech.Knowledge
Own elaboration based on intuitions in Hall et al. (2000) and Lanjouw and Schankerman (2004)	$\text{ANPCI}_{i,t} = \frac{C_{i,t}}{\lambda_t \cdot \left(\frac{1}{n_t}\sum_{j=1}^{n_t} C_{j,t}\right)}$ where: - $C_{i,t}$ is the number of citations received by patent i in year t, - n_t is the number of patents filed in year t, - $\frac{1}{n_t}\sum_{j=1}^{n_t} C_{j,t}$ is the average number of citations per patent in year t, -λ_t is a correction factor to account for citation inflation, which adjusts the average citation counts for inter-year comparability. Is the average number of citations per patent across all fields in year t, divided by a baseline year (e.g., t_0)	Intra-year comparison: index is normalized by the average citations within the same cohort (year and field). Inter-year comparison: adjusting for citation inflation λ_t should mitigate the bias of comparing older patents with more citations to newer ones that have not had as much time to accumulate citations.	Citations
Trajtenberg et al. (1997) (Modified better version)	$G_X = 1 - \sum_{j=1}^{M_i} \left(\frac{1}{N}\sum_{i=1}^{N} \frac{T_{ji}^n}{T_i^n}\right)^2$	X is the focal patent with Y_i patents citing the focal patent X, with $i = 1, \ldots, N$. T_i^n is the total number of IPC n-digit classes in y_i T_{ji}^n is the total number of IPC n-digit classes in the jth IPC4 digit class in y_i and $j = 1 \ldots M_i$ where M_i is the cardinality of all IPC4-digit classes in y_i	Citations Tech.Knowledge
Various	Patent quality: Composite indices	Typically based on patent citations, claims, patent renewals and patent family size. Usually different compositions as follows: [i]Patent quality index 4 – 4 components: number of forward citations (up to 5 years after publication); patent family size; number of claims; and the patent generality index. Only granted patents are covered by the index. [ii] Patent quality index 4b – 4 components, bis: number of forward citations (up to 5 years after publication); patent family size; corrected claims; and the patent generality index. Only granted patents are covered by the index. [iii] Patent quality index 6 – 6 components: covers the same components as above, plus the number of backward citations and the grant lag index. Entropy based approaches: Defines a weighting scheme (W) for all indicators of a patent (citations, claims...) based on entropy. With M patents, it then constructs $M \cdot W$. A set of negative (patents with the lowest weighted scores across the majority of indicators) and positive patents is then identified and divided; a similarity across good and negative patents is then computed. The following steps are then performed: Maximum Similarity: For each patent in M^R, the maximum similarity to any patent in M^N is determined. If the maximum similarity $S(M_i^R, M^N)$ exceeds a predefined threshold τ, that remaining patent is marked as negative. The process repeats iteratively. Newly marked negative patents are moved to M^N, and the model re-evaluates the remaining patents in M^R. The iteration stops when the number of remaining patents M^R is less than a threshold θ, or no new patents are marked as negative. The final output is a small set of patents in M^R that are not marked as negative, indicating their higher potential for technological innovation based on distinctiveness and low similarity to the negative patents.	Various

(ii) In Fleming (2001) novelty is primarily defined through the recombination of existing knowledge. The core idea is that innovations arise by combining different components (Fleming uses U.S. patent classification codes as a proxy for these components by examining how frequently different combinations of these classifications appear together in patents, Fleming can assess the novelty of the recombination[18]) in new ways. A patent is considered novel when it brings together elements (or knowledge) that have not been combined before, thus creating something new. Fleming emphasizes that recombinant novelty can vary in terms of uncertainty and outcomes: some combinations lead to breakthroughs, while others fail to generate impactful results. Fleming also makes a distinction between simple recombination, which involves putting together elements that are closely related or have been combined before, and more radical recombination, where the components are from different, often unrelated, technological fields. The latter leads to higher novelty and a greater chance of disruption but also comes with more uncertainty in terms of success. Thus, novelty in this context is seen as the creation of new combinations of existing components, with the degree of novelty depending on how distant or different these components are from one another.

(iii) In the early 2000s also another stream of literature developed following the influential publication of Aghion et al. (2005) according to which novelties were characterized by a certain degree of creative destruction. This idea has been incorporated in some studies on patents the most insightful of which is maybe Autor et al. (2020). The idea, though not explicitly stated, is that a novel patent is one that "destroys" similar patents (e.g. reducing the sales of products sold with patents in the same sector). Though this idea is very appealing it cannot be easily implemented.

(iv) The most recent stream is the one of Natural Language Programming (NLP) and text analyses of patents (Abbas et al., 2014; Gerken and Moehrle, 2012). In essence in Gerken and Moehrle (2012) the process of calculating patent novelty compares the text of a new patent with all previously filed patents. The similarity between the new patent and each prior one is measured, and the highest similarity score (the patent's "oldness") is subtracted from 1 to determine the novelty score. Essentially, the more distinct a new patent is from earlier patents, the higher its novelty. This approach emphasizes the unique content of a patent by minimizing similarities with existing technologies. In Abbas et al. (2014), semantic analysis is mentioned as a developing method, with an emphasis on how linguistic analysis (e.g., analyzing keywords, abstracts) is becoming crucial in identifying novel patents. But, their definition of novelty primarily revolves around citation gaps[19] and technological distance[20] (e.g., patents in new technological classes). More recently Kelly et al. (2021) defined, again, patent novelty by focusing on textual similarity. They measure novelty by comparing the text content of patents to both previous and future patents. A patent is considered novel if it is distinct from earlier patents but also relevant to future innovations. Their approach uses textual analysis to identify patents that contribute significantly to technological progress, and these are often predictive of future citations and market value.

Table A.2 summarizes the above indices. Note that different indices may capture different aspects of a patent. For example, the forward citations of a patent n measure its impact on subsequent "production" or innovation, whereas the composition of patent n across categories reflects its degree of "interdisciplinarity", which can be interpreted as a form of "originality". Therefore, it is not necessarily the case that all indices are positively correlated with one another.

In the main text we adopt a category-version of the Squicciarini et al. (2013) index as our primary measure of patent novelty due to its widespread recognition and methodological robustness within the economics of innovation literature (Higham et al., 2021; Dehghani et al., 2023). This index, endorsed and implemented by the OECD in its patent quality indicators framework, captures the diversity of technological fields cited by a patent-interpreted as a proxy for knowledge recombination and inventive originality. It has also been applied across a range of recent empirical studies (Dehghani et al., 2023; Gao and Lazarova, 2024; Angori et al., 2024) reflecting its growing standardization and comparability across datasets and time. Compared to newer text-based or machine-learning approaches, the Squicciarini et al. (2013)'s index offers a better replicability, and ease of modification in a cross-category setting as ours. Notice, however that our setting is general enough to allow for other measures of patent novelty as well.

Appendix B. Proofs of the main results and technical details for the statistical test

We here provide the rigorous proofs of the main results and the technical details that explain the performed statistical test. These proofs are based on the general theory we develop in Appendix B.1 and which has the merit to provide a general mathematical framework that can be applied also in other settings.

Proof of Theorem 3.1. For each time step t, set $\mathbf{X}_t = (X_{t,1}, \dots, X_{t,N})^\top$, $\mathbf{P}_t = (P_{t,1}, \dots, P_{t,N})^\top$ and $\boldsymbol{\theta} = (\theta_1, \dots, \theta_N)^\top$. Hence, the vectorial dynamics for the random vector $\mathbf{P}_t$ is, for $t \geq 0$,

$$\begin{aligned} &\mathbf{P}_0 = \boldsymbol{\theta}/c \neq \mathbf{0} \quad \text{a.s.} \\ &\mathbf{P}_{t+1} = \Big(1 - \frac{1}{t+1}\Big)\mathbf{P}_t + \frac{1}{t+1}\Gamma^\top \mathbf{X}_{t+1} + O(1/t^2)\mathbf{1} \\ &\qquad = \mathbf{P}_t - \frac{1}{t+1}(I - \Gamma^\top)\mathbf{P}_t + \frac{1}{t+1}\Gamma^\top \Delta\mathbf{M}'_{t+1} + O(1/t^2)\mathbf{1}, \end{aligned} \tag{B.1}$$

where $\Delta\mathbf{M}'_{t+1} = \mathbf{X}_{t+1} - \mathbf{P}_t$.

Now, we fix $x > 0$ and set

$$\zeta_0(x) = 1, \qquad \zeta_t(x) = \frac{\Gamma(t+x)}{\Gamma(t)} \sim t^x \uparrow +\infty\,. \tag{B.2}$$

More precisely, from Gouet (1993, Lemma 4.1) we have

$$\zeta_t(x) = t^x + O(t^{x-1}) \quad \text{and} \quad \frac{1}{\zeta_t(x)} = \frac{1}{t^x} + O(1/t^{x+1}),$$

and so

$$\begin{aligned} &\frac{1}{\zeta_{t+1}(x)\zeta_{t+1}(1-x)} = \\ &\Big(\frac{1}{(t+1)^x} + O(1/t^{x+1})\Big)\Big(\frac{1}{(t+1)^{1-x}} + O(1/t^{2-x})\Big) = \\ &\frac{1}{t+1} + O(1/t^2). \end{aligned} \tag{B.3}$$

Hence, multiplying (B.1) by $\zeta_{t+1}(1-\gamma^*)$ and using the relation

$$\frac{\zeta_{t+1}(x)}{\zeta_t(x)} = \frac{\Gamma(t+x+1)}{\Gamma(t+1)}\frac{\Gamma(t)}{\Gamma(t+x)} = 1 + \frac{x}{t} = 1 + \frac{x}{t+1} + O(1/t^2),$$

[18] Specifically, he looks at how often certain technological combinations have been used before. If a patent combines classifications that rarely appear together or have not been combined before, it is considered more novel. This method helps to quantify novelty based on how "new" or "unusual" the combination of technological elements is, rather than just looking at the number of citations or backward links. Fleming also explores the uncertainty associated with novelty, showing that the more novel the combination (i.e., the less frequently those components have been combined in the past), the greater the uncertainty in terms of the patent's success.

[19] If a patent does not reference many previous patents (or none at all), it suggests that the new patent is introducing ideas not heavily reliant on earlier work. This is seen as a sign of novelty because it indicates that the patent is not just building on existing knowledge, but possibly presenting something new.

[20] This refers to how different a patent is from those in similar fields. If a patent belongs to a new or less crowded technological class (based on how patents are categorized), it suggests that the patent is more novel because it explores a technological area that has not been fully developed yet.

with $x = 1-\gamma^*$, we get the following dynamics for $\mathbf{A}_t = \zeta_t(1-\gamma^*)\mathbf{P}_t$, where we set $\Delta\mathbf{M}_{t+1} = \zeta_{t+1}(1-\gamma^*)\Delta\mathbf{M}'_{t+1}$:

$$\begin{aligned}\mathbf{A}_{t+1} &= \Big(1-\frac{1}{t+1}(I-\Gamma^\top)\Big)\frac{\zeta_{t+1}(1-\gamma^*)}{\zeta_t(1-\gamma^*)}\mathbf{A}_t + \frac{1}{t+1}\Gamma^\top\Delta\mathbf{M}_{t+1} \\ &\quad + O\Big(\frac{\zeta_{t+1}(1-\gamma^*)}{t^2}\Big)\mathbf{1} \\ &= \Big(1-\frac{1}{t+1}(I-\Gamma^\top)\Big)\Big(1+\frac{1-\gamma^*}{t+1}+O(1/t^2)\Big)\mathbf{A}_t \\ &\quad + \frac{1}{t+1}\Gamma^\top\Delta\mathbf{M}_{t+1} + O\Big(\frac{\zeta_{t+1}(1-\gamma^*)}{t^2}\Big)\mathbf{1} \\ &= \mathbf{A}_t - \frac{1}{t+1}(\gamma^* I-\Gamma^\top)\mathbf{A}_t + \frac{1}{t+1}\Gamma^\top\Delta\mathbf{M}_{t+1} + O\Big(\frac{\zeta_{t+1}(1-\gamma^*)}{t^2}\Big)\mathbf{1},\end{aligned}$$

with $O(\zeta_t(1-\gamma^*)/t^2) = O(1/t^{1+\gamma^*})$.

Now, we are going to apply Theorem B.2 of Appendix B.1 with $\Phi = \Gamma$, $\phi^* = \gamma^*$ and $\mathcal{F} = (\mathcal{F}_t)_t$ the natural filtration associated to the model, i.e. $\mathcal{F}_t = \sigma(X_{n,h} : n \le t, h = 1,\dots,N)$. To this purpose, we choose $\mathbf{u}$ and $\mathbf{v}$ as the left and the right eigenvectors of Γ associated to γ^*, with strictly positive components and such that $\mathbf{v}^\top\mathbf{1} = 1$ and $\mathbf{v}^\top\mathbf{u} = 1$. (Recall that γ^* is simple and it is possible to choose the components of these vectors all strictly positive because of the Frobenius–Perron theory.) We set $\widetilde{A}_t = \mathbf{v}^\top\mathbf{A}_t = \zeta_t(1-\gamma^*)\mathbf{v}^\top\mathbf{P}_t$. First of all, we observe that assumptions (i) and (ii) of Theorem B.2 are satisfied because of $(\mathcal{A}1)$ and $(\mathcal{A}2)$, that imply $\gamma^* \in (0,1]$. Moreover, also assumption (iii) is verified: indeed, we have $\sup_t E[\widetilde{A}_t] < +\infty$ as $\widetilde{A}_t$ is non-negative (see Remark 3). Finally, we have

$$W_t = \sum_{j=1}^N E[(\Delta M_{t+1,j})^2 \mid \mathcal{F}_t] = \zeta_{t+1}(1-\gamma^*)^2\sum_{j=1}^N P_{j,t}(1-P_{j,t}) \le \zeta_{t+1}(1-\gamma^*)^2\sum_{j=1}^N P_{j,t}.$$

Then, denoting by $v_{\min} > 0$ the minimum element of $\mathbf{v}$, we obtain

$$\zeta_{t+1}(1-\gamma^*)^2\sum_{j=1}^N P_{j,t} \le \zeta_{t+1}(1-\gamma^*)^2\mathbf{v}^\top\mathbf{P}_t/v_{\min} \le 2\zeta_{t+1}(1-\gamma^*)\widetilde{A}_t/v_{\min}.$$

and so

$$W_t \le \frac{1}{\gamma^* v_{\min}}\zeta_{t+1}(1-\gamma^*)\widetilde{A}_t.$$

Therefore, recalling that $\sup_t E[\widetilde{A}_t] < +\infty$ as told before, we get

$$\sum_t \frac{1}{(t+1)^2}E[W_t] \le \frac{1}{\gamma^* v_{\min}}\sup_t E[\widetilde{A}_t]\sum_t \frac{\zeta_{t+1}(1-\gamma^*)}{(t+1)^2} < +\infty.$$

Hence, also assumption (iv) in Theorem B.2 is satisfied and we get

$$t^{1-\gamma^*}\mathbf{P}_t \overset{a.s./L^2}{\sim} \zeta_t(1-\gamma^*)\mathbf{P}_t = \mathbf{A}_t \overset{a.s.}{\longrightarrow} \mathbf{P}_\infty = \widetilde{P}_\infty\mathbf{u},$$

where $\widetilde{P}_\infty = \widetilde{A}_\infty$ is a square-integrable non-negative random variable, which is the almost sure limit of the above defined $\widetilde{A}_t$, or equivalently (by (B.2)), of $t^{1-\gamma^*}\mathbf{v}^\top\mathbf{P}_t$.

It remains to prove that $P(\widetilde{P}_\infty > 0) = 1$. To this purpose, we use (Aletti et al., 2023, Theorem S1.3). Indeed, setting $\widetilde{P}_t = \mathbf{v}^\top\mathbf{P}_t$, we have

$$\begin{aligned}\widetilde{P}_0 &= \mathbf{v}^\top\mathbf{P}_0 = \frac{1}{c}\mathbf{v}^\top\theta > 0 \\ \widetilde{P}_{t+1} &= \Big(1-\frac{1}{t+1}\Big)\widetilde{P}_t + \frac{\gamma^*}{t+1}\widetilde{X}_{t+1} + O(1/t^2), \quad t \ge 0,\end{aligned} \tag{B.4}$$

with $\widetilde{X}_{t+1} = \mathbf{v}^\top\mathbf{X}_{t+1}$. Hence, if we define the stochastic process $\mathcal{V} = (\mathcal{V}_t)_{t\ge0}$, taking values in the interval $[0,1]$, as

$$\mathcal{V}_0 = \widetilde{P}_0 > 0, \qquad \mathcal{V}_{t+1} = \Big(1-\frac{1}{t+2}\Big)\mathcal{V}_t + \frac{1}{t+2}Y_{t+1}, \quad t \ge 0,$$

where $Y_{t+1} = \gamma^*\widetilde{X}_{t+1}$ (that takes values in $[0,1]$, since $0 < \gamma^* \le 1$, $X_{t+1,j} \in \{0,1\}$ and $\mathbf{v}^\top\mathbf{1} = 1$), then by the technical result in Appendix B.2 (applied to $\mathcal{W}_t = \widetilde{P}_t$ with $\beta = 1$) we have

$$|\widetilde{P}_t - \mathcal{V}_t| = O(\ln(t)/t) \to 0$$

and also

$$t^{1-\gamma^*}|\widetilde{P}_t - \mathcal{V}_t| = O(t^{1-\gamma^*}\ln(t)/t) = O(\ln(t)/t^{\gamma^*}) \to 0.$$

Hence, from Aletti et al. (2023, Theorem S1.3) applied to $\mathcal{V} = (\mathcal{V}_t)$ with $\delta = \gamma^*$ (note that $E[Y_{t+1}|\mathcal{F}_t] = \gamma^*\widetilde{P}_t = \gamma^*\mathcal{V}_t + O(\ln(t)/t)$, we get that $t^{1-\gamma^*}\mathcal{V}_t$ converges almost surely to a strictly positive finite random variable. This random variable is obviously also the almost sure limit of $t^{1-\gamma^*}\widetilde{P}_t$ and so we can conclude that $P(\widetilde{P}_\infty > 0) = 1$. □

Proof of Theorem 3.2. From Theorem 3.1, using the same notation adopted in its proof, we get

$$S_{t,h} = \sum_{n=1}^t X_{n,h} \qquad \text{with } E[X_{t+1,h} \mid \mathcal{F}_t] = P_{t,h} \overset{a.s.}{\sim} \frac{P_{\infty,h}}{t^{1-\gamma^*}}$$

and so, by Williams (1991, sec. 12.15), we get

$$S_{t,h} \overset{a.s.}{\sim} S_{\infty,h}\,t^{\gamma^*} \quad \text{with } S_{\infty,h} = \frac{P_{\infty,h}}{\gamma^*}. \tag{B.5}$$

(We can also note that, similarly, since the convergence in Theorem 3.1 is also in mean, we also have $E[S_{t,h}] = \sum_{n=0}^t E[X_{n,h}] = \sum_{n=0}^t E[P_{n,h}] \sim t^{\gamma^*}E[S_{\infty,h}]$.) As a consequence of (B.5) and the fact that $P_{\infty,h}/P_{\infty,j} = u_h/u_j$, we obtain

$$\frac{S_{t,h}}{S_{t,j}} \overset{a.s.}{\longrightarrow} \frac{S_{\infty,h}}{S_{\infty,j}} = \frac{u_h}{u_j}. \quad \square$$

Proof of Theorem 3.3. In the proof of Theorem 3.1, we have proven that the limit random vector of $t^{1-\gamma^*}\mathbf{P}_t$ is $\mathbf{P}_\infty = \widetilde{P}_\infty\mathbf{u}$, where here $\widetilde{P}_\infty$ refers to a precise choice for the vector $\mathbf{u}$, that is $\widetilde{P}_\infty = \widetilde{P}_\infty(\mathbf{u})$. If we choose a different (left) eigenvector $\mathbf{u}'$ of Γ associated to γ^*, then we necessarily have $\mathbf{u}' = C\mathbf{u}$ with $C \ne 0$ and so we can write the limit random vector as $\mathbf{P}_\infty = \widetilde{P}_\infty(\mathbf{u}')\mathbf{u}'$ with $\widetilde{P}_\infty(\mathbf{u}') = \widetilde{P}_\infty(\mathbf{u})/C$. Summing up, for any choice of the (left) eigenvector $\mathbf{u}$, the limit random vector can be written as $\mathbf{P}_\infty = \widetilde{P}_\infty(\mathbf{u})\mathbf{u}$ with a suitable square-integrable random variable $\widetilde{P}_\infty(\mathbf{u})$ such that $P(\widetilde{P}_\infty(\mathbf{u}) \ne 0) = 1$ and $\widetilde{P}_\infty(C\mathbf{u}) = \widetilde{P}_\infty(\mathbf{u})/C$. In other words, each random variable $P_{\infty,h}$ can be factorized in the product of a deterministic term specific for each h, i.e. u_h, and a common random term, i.e. $\widetilde{P}_\infty(\mathbf{u})$.

Since $X_{t+1,h} \in \{0,1\}$ with $E[X_{t+1,h} \mid \mathcal{F}_t] = P_{t,h}$ and the convergence of $t^{1-\gamma^*}\mathbf{P}_t$ toward $\mathbf{P}_\infty$ is also in quadratic mean, we obtain

$$\begin{aligned}&t^{1-\gamma^*}Var[X_{t+1,h}] = t^{1-\gamma^*}E[P_{t,h}(1-P_{t,h})] \\ &\longrightarrow \begin{cases} E[P_{\infty,h}] = u_h E[\widetilde{P}_\infty(\mathbf{u})] & \text{for } \gamma^* < 1 \\ E[P_{\infty,h}(1-P_{\infty,h})] = u_h E\left[\widetilde{P}_\infty(\mathbf{u})(1-u_h\widetilde{P}_\infty(\mathbf{u}))\right] & \text{for } \gamma^* = 1. \end{cases}\end{aligned}$$

Therefore, since the sign of the components of $\mathbf{u}$ (which is the same for all of them) necessarily coincides with the one of $\widetilde{P}_\infty(\mathbf{u})$, in order to obtain the first limit relation in Theorem 3.3, we can set $\alpha(\mathbf{u})$ equal to $|E[\widetilde{P}_\infty(u)]|$ when $\gamma^* < 1$ and equal to $|E\left[\widetilde{P}_\infty(\mathbf{u})(1-u_h\widetilde{P}_\infty(\mathbf{u}))\right]|$ when $\gamma^* = 1$, so that the two possible cases in the above formula can be summarized as $|u_h|\alpha(\mathbf{u})$. Moreover, for each pair $h \ne j$, since $X_{t+1,h}$ and $X_{t+1,j}$ are conditionally independent given $\mathcal{F}_t$, we get

$$\begin{aligned}&t^{2(1-\gamma^*)}cov(X_{t+1,h}, X_{t+1,j}) = t^{2(1-\gamma^*)}cov(P_{t,h}, P_{t,j}) \longrightarrow cov(P_{\infty,h}, P_{\infty,j}) \\ &= u_h u_j Var[\widetilde{P}_\infty(\mathbf{u})].\end{aligned}$$

Hence, in order to obtain the second desired relation, we can set $\sigma^2(\mathbf{u}) = Var[\widetilde{P}_\infty(\mathbf{u})]$. Finally, as a consequence, for the correlation coefficients, we have

$$t^{1-\gamma^*}\rho(X_{t+1,h}, X_{t+1,j}) = \frac{t^{2(1-\gamma^*)}cov(X_{t+1,h}, X_{t+1,j})}{\sqrt{t^{(1-\gamma^*)}Var[X_{t+1,h}]}\sqrt{t^{(1-\gamma^*)}Var[X_{t+1,j}]}} \longrightarrow \sqrt{u_h u_j}\frac{\sigma^2(\mathbf{u})}{\alpha(\mathbf{u})}. \quad \square$$

Proof of Theorem 3.4. From Theorems 3.1 and 3.2, using the same notation adopted in their proofs, we have by $(\mathcal{A}3)$

$$\begin{aligned}E[X_{t+1,h}Y_{t+1,k} \mid \mathcal{F}_t] &= E[Y_{t+1,k}]E[X_{t+1,h} \mid \mathcal{F}_t] = \pi_k P_{t,h} \\ &\overset{a.s.}{\sim} t^{-(1-\gamma^*)}\pi_k P_{\infty,h} = t^{-(1-\gamma^*)}\pi_k\gamma^* S_{\infty,h}\end{aligned}$$

and so it is enough to apply (Williams, 1991, sec. 12.15) in order to obtain $S_{t,k,h} = \sum_{n=1}^t X_{n,h}Y_{n,k} \overset{a.s.}{\sim} \pi_k S_{\infty,h}\,t^{\gamma^*}$. □

Technical details for the statistical inference (Section 4.2)

We here use the same choice of the eigenvectors $\mathbf{v}$ and $\mathbf{u}$ and the same notation adopted in the proof of Theorem 3.1. We recall that the dynamics of the above defined process $\mathbf{A}_t = \zeta_t(1-\gamma^*)\mathbf{P}_t$ with $\zeta_t(\cdot)$ defined in (B.2), is

$$\mathbf{A}_0 = \theta/c$$
$$\mathbf{A}_{t+1} = \mathbf{A}_t - \frac{1}{t+1}(\gamma^* I - \Gamma^\top)\mathbf{A}_t + \frac{1}{t+1}\Gamma^\top \Delta\mathbf{M}_{t+1} + O\left(\frac{\zeta_{t+1}(1-\gamma^*)}{t^2}\right)\mathbf{1}\,,$$

where $O(\zeta_t(1-\gamma^*)/t^2) = O(1/t^{1+\gamma^*})$ and $\Delta\mathbf{M}_{t+1} = \zeta_{t+1}(1-\gamma^*)\Delta\mathbf{M}'_{t+1}$ with $\Delta\mathbf{M}'_{t+1} = \mathbf{X}_{t+1} - \mathbf{P}_t$. Moreover, setting $\mathbf{S}_t = (S_{t,1}, \dots, S_{t,N})$ and $\mathbf{B}_t = \frac{1}{\zeta_t(\gamma^*)}\mathbf{S}_t$, we find the following vectorial dynamics:

$$\mathbf{B}_0 = \mathbf{0}$$
$$\begin{aligned}\mathbf{B}_{t+1} &= \frac{\zeta_t(\gamma^*)}{\zeta_{t+1}(\gamma^*)}\mathbf{B}_t + \frac{1}{\zeta_{t+1}(\gamma^*)}\mathbf{X}_{t+1}\\ &= \left(1 - \frac{\gamma^*}{t+1}\right)\mathbf{B}_t + \frac{1}{\zeta_{t+1}(\gamma^*)}\Delta\mathbf{M}'_{t+1} + \frac{1}{\zeta_{t+1}(\gamma^*)}\mathbf{P}_t\\ &= \left(1 - \frac{\gamma^*}{t+1}\right)\mathbf{B}_t + \frac{1}{\zeta_{t+1}(\gamma^*)\zeta_{t+1}(1-\gamma^*)}\Delta\mathbf{M}_{t+1}\\ &\quad + \frac{1}{\zeta_{t+1}(\gamma^*)\zeta_{t+1}(1-\gamma^*)}\frac{\zeta_{t+1}(1-\gamma^*)}{\zeta_t(1-\gamma^*)}\mathbf{A}_t\,.\end{aligned}$$

Using (B.3) and the relation $\zeta_{t+1}(x)/\zeta_t(x) = 1 + O(1/t)$, we obtain

$$\begin{aligned}\mathbf{B}_{t+1} = \mathbf{B}_t &- \frac{1}{t+1}\left(\gamma^*\mathbf{B}_t - \mathbf{A}_t\right)\\ &+ \frac{1}{t+1}\Delta\mathbf{M}_{t+1} + O(\zeta_t(1-\gamma^*)/t^2)\mathbf{1}\,,\end{aligned} \tag{B.6}$$

where again $O(\zeta_t(1-\gamma^*)/t^2) = O(1/t^{1+\gamma^*})$. Finally, we observe that, by (B.2), we have

$$\begin{aligned}t^{-(1-\gamma^*)}E[\Delta\mathbf{M}_{t+1}\Delta\mathbf{M}_{t+1}^\top|\mathcal{F}_t] &\overset{a.s.}{\sim} t^{1-\gamma^*}diag\left(P_{t,1}(1-P_{1,t}), \dots, P_{t,N}(1-P_{t,N})\right)\\ &\xrightarrow{a.s.} \begin{cases} diag(\mathbf{P}_\infty) & \text{for } \gamma^* < 1\\ diag(P_{\infty,1}(1-P_{\infty,1}), \dots, P_{\infty,N}(1-P_{\infty,N})) & \text{for } \gamma^* = 1 \text{ (and so } \mathbf{u} = \mathbf{1})\end{cases}\\ &= \begin{cases} \widetilde{P}_\infty diag(\mathbf{u}) & \text{for } \gamma^* < 1\\ \widetilde{P}_\infty(1-\widetilde{P}_\infty)I & \text{for } \gamma^* = 1 \text{ (and so } \mathbf{u} = \mathbf{1})\end{cases}\end{aligned}$$

and we recall that, as proven before, $\widetilde{A}_t = \mathbf{v}^\top\mathbf{A}_t = \zeta_t(1-\gamma^*)\mathbf{v}^\top\mathbf{P}_t \overset{a.s.}{\to} \widetilde{P}_\infty$. Therefore, the pair $(\mathbf{A}_t, \mathbf{B}_t)_t$ satisfies the dynamics and the conditions required in the general central limit theorem proven in Aletti et al. (2026, Appendix A), provided we assume $\mathcal{R}e(\gamma_2^*)/\gamma^* < 1/2$, where γ_2^* is an eigenvalue of Γ different from γ^* with highest real part, that is $\gamma_2^* \in Sp(\Gamma)\setminus\{\gamma^*\}$ with $\mathcal{R}e(\gamma_2^*) = \max\{\mathcal{R}e(\gamma) : \gamma \in Sp(\Gamma)\setminus\{\gamma^*\}\}$. Hence, we can apply the statistical tools based on that result and described in Aletti et al. (2026, Appendix B and C). In particular, in the mean-field case we have $\mathbf{v} = N^{-1}\mathbf{1}$, $\mathbf{u} = \mathbf{1}$ and $\gamma_2^* = \gamma^*(1-\iota)$ so that, when $\iota > 1/2$, we have

$$\begin{aligned}&(2\iota-1)\frac{\|\mathbf{S}_t - (\mathbf{v}^\top\mathbf{S}_t)\mathbf{1}\|^2}{\mathbf{v}^\top\mathbf{S}_t}\\ &= (2\iota-1)\zeta_t(\gamma^*)\frac{\|\mathbf{B}_t - (\mathbf{v}^\top\mathbf{B}_t)\mathbf{1}\|^2}{\mathbf{v}^\top\mathbf{B}_t} \overset{a.s.}{\sim} (2\iota-1)t^{\gamma^*}\frac{\|\mathbf{B}_t - (\mathbf{v}^\top\mathbf{B}_t)\mathbf{1}\|^2}{\mathbf{v}^\top\mathbf{B}_t} \xrightarrow{d} \chi^2(N-1).\end{aligned}$$

B.1. General results

Let $\mathbf{A}_t = (A_{t,1}, \dots, A_{t,N})^\top$, with $t \geq 0$, be a multi-dimensional real stochastic process, adapted to a filtration $(\mathcal{F}_t)_t$, with the following dynamics:

$$\mathbf{A}_{t+1} = \mathbf{A}_t - \frac{1}{t+1}(\phi^* I - \Phi^\top)\mathbf{A}_t + \frac{1}{t+1}\Phi^\top\Delta\mathbf{M}_{t+1} + \mathbf{R}_{A,t+1} \tag{B.7}$$

where $\mathbf{A}_0$ is *integrable* and

(i) $\Phi^\top$ is a non-negative irreducible matrix with leading eigenvalue $0 < \phi^* \leq 1$;

(ii) $\mathbf{R}_{A,t+1} = O(t^{-(1+\beta)})\mathbf{1}$ for some $\beta > 0$.[21]

Let $\mathbf{u}$ and $\mathbf{v}$ be the left and the right eigenvectors of Φ associated to ϕ^*, with strictly positive components and such that $\mathbf{v}^\top\mathbf{1} = 1$ and $\mathbf{v}^\top\mathbf{u} = 1$. (Recall that ϕ^* is real and simple and it is possible to choose the components of these vectors all strictly positive because of the Frobenius–Perron theory). Set $\widetilde{A}_t = \mathbf{v}^\top\mathbf{A}_t$.

Theorem B.1. *Under (i) and (ii) and assuming*

(iii) $\sup_t E[\,|\widetilde{A}_t|\,] < +\infty$,

we have $\widetilde{A}_t \overset{a.s.}{\longrightarrow} \widetilde{A}_\infty$, *where* $\widetilde{A}_\infty$ *is an integrable random variable. Moreover, if we also assume*

(iv) $\mathbf{A}_0$ *square-integrable and* $\sum_t w_t/(t+1)^2 < +\infty$ *where* $w_t = \sum_{j=1}^N E[(\Delta M_{t+1,j})^2]$,

then we have $\sup_t E[(\widetilde{A}_t)^2] < +\infty$ *and so* $\widetilde{A}_t$ *converges to* $\widetilde{A}_\infty$ *also in quadratic mean (i.e. in* L^2*).*

Note that (iii) is verified when $(\widetilde{A}_t)_t$ is uniformly integrable and in this case we also have that the convergence is in mean.

Proof. By multiplying Eq. (B.7) by $\mathbf{v}^\top$ we obtain

$$\widetilde{A}_{t+1} = \widetilde{A}_t + \frac{1}{t+1}\phi^*\Delta\widetilde{M}_{t+1} + \mathbf{v}^\top\mathbf{R}_{A,t+1}\,, \tag{B.8}$$

where $\Delta\widetilde{M}_{t+1} = \mathbf{v}^\top\Delta\mathbf{M}_{t+1}$. Setting $\check{M}_t = \sum_{n=1}^t \frac{1}{n}\Delta\widetilde{M}_n$ and $\widetilde{R}_{A,t+1} = \mathbf{v}^\top\mathbf{R}_{A,t+1} = O(1/t^{1+\beta})$ (by (ii)), we have

$$\widetilde{A}_{t+1} - \widetilde{A}_0 = \sum_{n=0}^t(\widetilde{A}_{n+1} - \widetilde{A}_n) = \phi^*\check{M}_{t+1} + \sum_{n=0}^t \widetilde{R}_{A,n+1}\,. \tag{B.9}$$

Hence, since $\sup_t E[\,|\widetilde{A}_t|\,] < +\infty$ (by (iii)) and $\sum_t 1/t^{1+\beta} < +\infty$, we also have $\sup_t E[\,|\check{M}_t|\,] < +\infty$. Therefore, $(\check{M}_t)_t$ is a martingale bounded in L^1 and so it converges almost surely to an integrable random variable $\check{M}_\infty$. It follows from (B.9) that $(\widetilde{A}_t)$ converges almost surely to an integrable random variable $\widetilde{A}_\infty$. Moreover, we obtain

$$\begin{aligned}(\widetilde{A}_t - \widetilde{A}_\infty)^2 &= \left(\phi^*(\check{M}_t - \check{M}_\infty) + \sum_{n\geq t}\widetilde{R}_{A,n+1}\right)^2\\ &= (\phi^*)^2(\check{M}_t - \check{M}_\infty)^2 + O(1/t^\beta)\,.\end{aligned} \tag{B.10}$$

Now, we are going to prove that, under assumption (iv), we have $\sup_t(\widetilde{A}_t)^2 < +\infty$ so that (by (B.9)) $(\check{M}_t)_t$ is a martingale bounded in L^2 and so $\check{M}_\infty$ is square-integrable and $\check{M}_t$ converges in quadratic mean to it. By (B.10) this fact obviously implies that $\widetilde{A}_t$ converges in quadratic mean to $\widetilde{A}_\infty$.

We observe that, from (iv) and the fact that $(\Delta\widetilde{M}_{t+1})^2 \leq C\sum_{j=1}^N(\Delta M_{t+1,j})^2$, we obtain

$$E\left[\left|\frac{\Delta\widetilde{M}_{t+1}}{t+1}\widetilde{R}_{A,t+1}\right|\right]^2 \leq E\left[\frac{(\Delta\widetilde{M}_{t+1})^2}{(t+1)^2}\right]E[(\widetilde{R}_{A,t+1})^2] \leq C\frac{w_t}{(t+1)^2}O(1/t^{2(1+\beta)})$$

and so

$$E\left[\left|\frac{\Delta\widetilde{M}_{t+1}}{t+1}\widetilde{R}_{A,t+1}\right|\right] = o(1/t^{1+\beta}).$$

Therefore, from (B.8), since $\sup_t E[\widetilde{A}_t] < +\infty$, we get

$$\begin{aligned}E[(\widetilde{A}_{t+1})^2] &\leq E[(\widetilde{A}_t)^2] + (\phi^*)^2C\frac{w_t}{(t+1)^2} + O\left(\frac{1}{t^{2(1+\beta)}}\right) + O\left(\frac{\sup_t E[\widetilde{A}_t]}{t^{1+\beta}}\right) + o\left(\frac{1}{t^{1+\beta}}\right),\\ &= E[(\widetilde{A}_t)^2] + (\phi^*)^2C\frac{w_t}{(t+1)^2} + O(1/t^{1+\beta})\,.\end{aligned}$$

[21] The notation $R_t = O(s_t)$ means that R_t is a (possibly random) remainder term such that $|R_t| \leq Cs_t$ for a suitable constant C and for t large enough.

Then, we find

$$|E[(\widetilde{A}_t)^2] - E[(\widetilde{A}_0)^2]| \le \sum_{n=0}^{t-1} |E[(\widetilde{A}_{n+1})^2] - E[(\widetilde{A}_n)^2]| \le (\phi^*)^2 C \sum_n \frac{w_n}{(n+1)^2} + \sum_n O(1/n^{1+\beta}) < +\infty,$$

where we have used (iv) in order to say that the first series is finite. Therefore, assuming $\mathbf{A}_0$ (and so $\widetilde{A}_0$) square-integrable, we have $\sup_t E[(\widetilde{A}_t)^2] < +\infty$. □

Remark 3 (*Non-Negative Case*)**.** Condition (iii) is verified when $\widetilde{A}_t$ is non-negative for each t. Indeed, for each t, we have

$$|E[\widetilde{A}_t] - E[\widetilde{A}_0]| \le \sum_{n=0}^{t-1} |E[\widetilde{A}_{n+1}] - E[\widetilde{A}_n]| \le \sum_n O(1/n^{1+\beta})$$

and thus, since the last series is finite and $\widetilde{A}_0$ is integrable, it follows $\sup_t E[\widetilde{A}_t] < +\infty$. Hence, if $\widetilde{A}_t$ is non-negative, we can conclude that (iii) is verified.

Theorem B.2. *Assuming (i), (ii), (iii) and (iv), we have*

$$\mathbf{A}_t \overset{a.s./L^2}{\longrightarrow} \widetilde{A}_\infty \mathbf{u},$$

where $\widetilde{A}_\infty$ is a square-integrable random variable.

Proof. We firstly want to prove that we can neglect the term $R_{A,t+1}$ in the dynamics (B.7) of $\mathbf{A}_t$.

We recall that the matrix $\boldsymbol{\Phi}^\top$ can be decomposed as

$$\boldsymbol{\Phi}^\top = \phi^* \mathbf{u}\mathbf{v}^\top + U D V^\top,$$

where D is the diagonal matrix whose elements are the eigenvalues of $\boldsymbol{\Phi}$ (i.e. of $\boldsymbol{\Phi}^\top$) different from ϕ^* and U and V denote the matrices whose columns are the left (right) and the right (left) eigenvectors of $\boldsymbol{\Phi}$ (of $\boldsymbol{\Phi}^\top$, respectively) associated to these eigenvalues, so that we have

$$V^\top \mathbf{u} = U^\top \mathbf{v} = 0, \quad V^\top U = U^\top V = I, \quad I = \mathbf{u}\mathbf{v}^\top + U V^\top. \tag{B.11}$$

Therefore the dynamics of $\mathbf{A}_t$ can be rewritten as follows:

$$\mathbf{A}_{t+1} = \Big(I - \frac{1}{t+1} U(I\phi^* - D)V^\top\Big)\mathbf{A}_t + \frac{1}{t+1}\boldsymbol{\Phi}^\top \Delta\mathbf{M}_{t+1} + \mathbf{R}_{A,t+1}. \tag{B.12}$$

Let $\alpha_j = \phi^* - \phi_j$ with ϕ_j eigenvalue of $\boldsymbol{\Phi}$ different from ϕ^*. Then $\mathcal{R}e(\alpha_j) > 0$. Moreover, we have

$$\begin{aligned}\mathbf{A}_{t+1} &= C_{m_0,t}\mathbf{A}_{m_0} + \sum_{k=m_0}^{t} \frac{1}{k+1} C_{k+1,t}\boldsymbol{\Phi}^\top \Delta\mathbf{M}_{k+1} + \sum_{k=m_0}^{t} C_{k+1,t}\mathbf{R}_{\mathbf{A},k+1} \\ &= C_{m_0,t}\mathbf{A}_{m_0} + \sum_{k=m_0}^{t} \frac{1}{k+1} C_{k+1,t}\boldsymbol{\Phi}^\top \Delta\mathbf{M}_{k+1} + \boldsymbol{\rho}_{t+1},\end{aligned} \tag{B.13}$$

where m_0 is such that $\mathcal{R}e(\alpha_j)/(m_0+1) < 1$ for each j and

$$C_{k+1,t} = \prod_{m=k+1}^{t}\Big(I - \frac{1}{m+1}(I\phi^* - \boldsymbol{\Phi}^\top)\Big) = \prod_{m=k+1}^{t}\Big(I - \frac{1}{m+1}U(I\phi^* - D)V^\top\Big) = U\Big(\prod_{m=k+1}^{t}\Big(I - \frac{1}{m+1}(I\phi^* - D)\Big)\Big)V^\top,$$

and so we can write $C_{k+1,t} = U A_{k+1,t} V^\top$ and

$$A_{k+1,t} = \prod_{m=k+1}^{t}\Big(I - \frac{1}{m+1}(I\phi^* - D)\Big).$$

Moreover, setting for any $x \in \mathbb{C}$ with $\mathcal{R}e(x)/(m_0+1) < 1$, $p_{m_0}(x) = 1$ and $p_k(x) = \prod_{m=m_0}^{k}(1 - \frac{x}{m+1})$ for $k \ge m_0$ and $F_{k+1,t} = \frac{p_t(x)}{p_k(x)}$ for $m_0 - 1 \le k \le t-1$, from Aletti et al. (2019, Lemma A.5) we get

$$[A_{k+1,t}]_{jj} = F_{k+1,t}(\alpha_j).$$

We now prove that $|\boldsymbol{\rho}_{t+1}| \overset{a.s.}{\to} 0$. To this end, first notice that $O(|C_{k+1,t}|) = O(|A_{k+1,t}|)$ and, setting $a_2^* = \mathcal{R}e(\alpha_2^*) = \phi^* - \mathcal{R}e(\phi_2^*)$ with ϕ_2^* eigenvalue of $\boldsymbol{\Phi}$, different from ϕ^*, such that $\mathcal{R}e(\phi_2^*) = \max_j\{\mathcal{R}e(\phi_j)\}$, we have (see Aletti et al. (2017, Lemma A.4))

$$|A_{k+1,t}| = O\Big(\frac{|p_t(\alpha_2^*)|}{|p_k(\alpha_2^*)|}\Big) = O\Big(\Big(\frac{k}{t}\Big)^{a_2^*}\Big) \quad \text{for } k = m_0, \dots, t-1,$$

and simply $|A_{t+1,t}| = O(1)$ for $k = t$. Moreover, recalling that $\mathbf{R}_{A,t+1} = O(t^{-(1+\beta)})\mathbf{1}$ for some $\beta > 0$, we have

$$\begin{aligned}|\boldsymbol{\rho}_{t+1}| &= \Big|\sum_{k=m_0}^{t} C_{k+1,t}\mathbf{R}_{A,k+1}\Big| = O\Big(\sum_{k=m_0}^{t-1}\Big(\frac{k}{t}\Big)^{a_2^*}\frac{1}{k^{1+\beta}}\Big) + O(1/t^{1+\beta}) \\ &= O\Big(\frac{1}{t^{a_2^*}}\sum_{k=m_0}^{t-1} k^{a_2^*-1-\beta}\Big) + O(1/t^{1+\beta}) \to 0,\end{aligned}$$

because $a_2^* > 0$ and $\beta > 0$.

Therefore, in all the sequel, without loss of generality, we can assume that $\mathbf{A}_t$ follows the dynamics (B.7) with $\mathbf{R}_{A,t+1} = \mathbf{0}$.

We now decompose the vectorial process $\mathbf{A}_t$ by means of the Jordan representation of the matrix $\boldsymbol{\Phi}$. Specifically, for any $\phi \in Sp(\boldsymbol{\Phi}) \setminus \{\phi^*\}$, we can denote as J_ϕ the Jordan block and with U_ϕ and V_ϕ the matrices whose columns are, respectively, the left and right (possibly generalized) eigenvectors of $\boldsymbol{\Phi}$ associated to the eigenvalue ϕ, i.e.

$$\boldsymbol{\Phi} V_\phi = V_\phi J_\phi \qquad \text{and} \qquad U_\phi^\top \boldsymbol{\Phi} = J_\phi U_\phi^\top.$$

Then, we can consider the decomposition

$$\mathbf{A}_t = \widetilde{A}_t \mathbf{u} + \sum_{\phi \in Sp(\boldsymbol{\Phi})\setminus\{\phi^*\}} \mathbf{A}_{\phi,t},$$

where $\widetilde{A}_t = \mathbf{v}^\top \mathbf{A}_t$ (as defined above) and $\mathbf{A}_{\phi,t} = U_\phi V_\phi^\top \mathbf{A}_t$.

We have already proven the almost sure convergence and the convergence in L^2 for $\widetilde{A}_t$ under (i), (ii), (iii) and (iv). In the following steps we are going to show that, under (i), (ii) and (iv), each $\mathbf{A}_{\phi,t}$ converges almost surely and in L^2 to zero. In particular, this last task will be done separately for the eigenvalues with $|\phi| < \phi^*$ and with $|\phi| = \phi^*$. Remember that the assumption that $\boldsymbol{\Phi}$ (or, equivalently, $\boldsymbol{\Phi}^\top$) is irreducible ensures that ϕ^* is real, simple and $|\phi| \le \phi^*$ for any $\phi \in Sp(\boldsymbol{\Phi})$. Moreover, let us set $W_t = \sum_{j=1}^{N} E[(\Delta M_{t+1,j})^2 \mid \mathcal{F}_t]$ and observe that assumption (iv) means $E[\sum_t W_t/(t+1)^2] = \sum_t E[W_t]/(t+1)^2 = \sum_t w_t/(1+t)^2 < +\infty$, which also implies $\sum_t W_t/(t+1)^2 < +\infty$ almost surely.

Study of $\mathbf{A}_{\phi,t}$ with $|\phi| < \phi^*$.

Let $\check{\mathbf{A}}_t = V_\phi^\top \mathbf{A}_t$ and since $\mathbf{A}_{\phi,t} = U_\phi V_\phi^\top \mathbf{A}_t = U_\phi \check{\mathbf{A}}_t$, it is enough to prove that $\|\check{\mathbf{A}}_t\|^2$ converges a.s. and in L^2 to zero. To this end, by multiplying Eq. (B.7) by $V_\phi^\top$, we have

$$\check{\mathbf{A}}_{t+1} = \Big(I - \frac{1}{t+1}(\phi^* I - J_\phi^\top)\Big)\check{\mathbf{A}}_t + \frac{1}{t+1} J_\phi^\top V_\phi^\top \Delta\mathbf{M}_{t+1}.$$

Then, since for any real matrix Q we can write

$$E[\Delta\mathbf{M}_{t+1}^\top Q \Delta\mathbf{M}_{t+1} | \mathcal{F}_t] = \sum_{j=1}^{N} q_{jj} E[(\Delta M_{j,t+1})^2 | \mathcal{F}_t] \le \max_j q_{jj} W_t, \tag{B.14}$$

we have that

$$
\begin{aligned}
E[\|\breve{\mathbf{A}}_{t+1}\|^2|\mathcal{F}_t] &\le \left\|\left(\left(1-\frac{\phi^*}{t+1}\right)I+\frac{1}{t+1}J_\phi\right)\breve{\mathbf{A}}_t\right\|^2 \\
&\quad+\frac{1}{(t+1)^2}\sum_{j=1}^{N}[\bar{V}_\phi\bar{J}_\phi J_\phi^\top V_\phi^\top]^2_{jj}E[(\Delta M_{j,t+1})^2|\mathcal{F}_t] \\
&\le \left(1-\frac{\phi^*}{t+1}+\frac{\|J_\phi\|_{2,2}}{t+1}\right)^2\|\breve{\mathbf{A}}_t\|^2 \\
&\quad+\frac{1}{(t+1)^2}\max_j\{[\bar{V}_\phi\bar{J}_\phi J_\phi^\top V_\phi^\top]^2_{jj}\}W_t.
\end{aligned}
$$

Then, regarding the first term, we note that

$$
\left(1-\frac{\phi^*}{t+1}+\frac{\|J_\phi\|_{2,2}}{t+1}\right)^2 \le \left(1-\frac{\phi^*}{t+1}+\frac{|\phi|+\phi^*}{2(t+1)}\right)^2 = \left(1-\frac{\phi^*-|\phi|}{2(t+1)}\right)^2,
$$

and so

$$
E[\|\breve{\mathbf{A}}_{t+1}\|^2|\mathcal{F}_t] \le \left(1-\frac{\phi^*-|\phi|}{2(t+1)}\right)^2\|\breve{\mathbf{A}}_t\|^2+\frac{C}{(t+1)^2}W_t.
$$

Therefore, since $\phi^* > |\phi|$ and by (iv), the process $\|\breve{\mathbf{A}}_t\|^2$ is a non-negative almost supermartingale so that it converges almost surely (see Robbins and Siegmund (1971)). Moreover, by applying the expectation, we obtain

$$
E[\|\breve{\mathbf{A}}_{t+1}\|^2] \le \left(1-\frac{\phi^*-|\phi|}{2(t+1)}\right)^2E[\|\breve{\mathbf{A}}_t\|^2]+\frac{C}{(t+1)^2}E[W_t],
$$

which, since $\sum_t(\phi^*-|\phi|)/(t+1)=+\infty$, by (iv) and Aletti et al. (2023, Lemma S1.6), we can conclude that $\|\breve{\mathbf{A}}_t\| \overset{a.s./L^2}{\longrightarrow} 0$, and hence $\breve{\mathbf{A}}_t \overset{a.s./L^2}{\longrightarrow} \mathbf{0}$.

Study of $\mathbf{A}_{\phi,t}$ with $|\phi|=\phi^*$.

From the Frobenius–Perron theory, we know that each eigenvalue with maximum modulus is simple. Let us denote the corresponding right and left eigenvectors by $\mathbf{v}_\phi$ and $\mathbf{u}_\phi$ of $\boldsymbol{\Phi}$. Then, set $a_{\phi,t}=\mathbf{v}_\phi^\top\mathbf{A}_t$ so that, since we have $\mathbf{A}_{\phi,t}=\mathbf{u}_\phi\mathbf{v}_\phi^\top\mathbf{A}_t=\mathbf{u}_\phi a_{\phi,t}$, it is enough to prove that $|a_{\phi,t}|$ converges to zero almost surely and in L^2. To this end, by multiplying Eq. (B.7) by $\mathbf{v}_\phi^\top$, we have

$$
a_{\phi,t+1}=\left(1-\frac{1}{t+1}(\phi^*-\phi)\right)a_{\phi,t}+\frac{\phi}{t+1}\mathbf{v}_\phi^\top\Delta\mathbf{M}_{t+1}.
$$

Then, using (B.14), we have that

$$
\begin{aligned}
E[|a_{\phi,t+1}|^2|\mathcal{F}_t] &\le \left|1-\frac{\phi^*}{t+1}+\frac{\phi}{t+1}\right|^2|a_{\phi,t}|^2+\frac{|\phi|^2}{(t+1)^2}\sum_{j=1}^{N}|v_j|^2E[(\Delta M_{j,t+1})^2|\mathcal{F}_t] \\
&\le \left|1-\frac{\phi^*}{t+1}+\frac{\phi}{t+1}\right|^2|a_{\phi,t}|^2+\frac{|\phi|^2}{(t+1)^2}\max_j\{|v_j|^2\}W_t.
\end{aligned}
$$

Then, regarding the first term we have that

$$
\begin{aligned}
\left|1-\frac{\phi^*}{t+1}+\frac{\phi}{t+1}\right|^2 &= \left(1-\frac{\phi^*}{t+1}+\frac{\mathcal{R}e(\phi)}{t+1}\right)^2+\left(\frac{\mathcal{I}m(\phi)}{t+1}\right)^2 \\
&= 1+\left(\frac{\phi^*-\mathcal{R}e(\phi)}{t+1}\right)^2-2\left(\frac{\phi^*-\mathcal{R}e(\phi)}{t+1}\right)+\left(\frac{\mathcal{I}m(\phi)}{t+1}\right)^2 \\
&= 1-\left(\frac{2(\phi^*-\mathcal{R}e(\phi))}{t+1}\right)+\left(\frac{\phi^{*2}-2\phi^*\mathcal{R}e(\phi)+\mathcal{R}e(\phi)^2+\mathcal{I}m(\phi)^2}{(t+1)^2}\right) \\
&= 1-\left(\frac{2(\phi^*-\mathcal{R}e(\phi))}{t+1}\right)+\left(\frac{2\phi^*(\phi^*-\mathcal{R}e(\phi))}{(t+1)^2}\right) \\
&= 1-2\left(\frac{1}{t+1}-\frac{\phi^*}{(t+1)^2}\right)(\phi^*-\mathcal{R}e(\phi))
\end{aligned}
$$

and so

$$
E[|a_{\phi,t+1}|^2|\mathcal{F}_t] \le \left(1-2\left(\frac{1}{t+1}-\frac{\phi^*}{(t+1)^2}\right)(\phi^*-\mathcal{R}e(\phi))\right)|a_{\phi,t}|^2+\frac{C}{(t+1)^2}W_t.
$$

Therefore, since $\phi^* > \mathcal{R}e(\phi)$ and by (iv), the process $|a_{\phi,t}|^2$ is a non-negative almost supermartingale so that it converges almost surely (see Robbins and Siegmund (1971)). Moreover, by applying the expectation, we obtain

$$
E\left[|a_{\phi,t+1}|^2\right] \le \left(1-2\left(\frac{1}{t+1}-\frac{\phi^*}{(t+1)^2}\right)(\phi^*-\mathcal{R}e(\phi))\right)E\left[|a_{\phi,t}|^2\right]+\frac{C}{(t+1)^2}E[W_t].
$$

Since $\sum_t(1/(t+1)-\phi^*/(t+1)^2)=+\infty$ and by (iv) and Aletti et al. (2023, Lemma S1.6), we can conclude that $|a_{\phi,t}| \overset{a.s./L^2}{\longrightarrow} 0$, and hence $a_{\phi,t} \overset{a.s./L^2}{\longrightarrow} 0$. □

B.2. Technical result

Let $\mathcal{W}=(\mathcal{W}_t)_{t\ge0}$ be a bounded stochastic process with the following dynamics

$$
\mathcal{W}_{t+1}=\left(1-\frac{1}{t+1}\right)\mathcal{W}_t+\frac{1}{t+1}Y_{t+1}+R_{t+1},\quad t\ge0,
$$

where $(Y_t)_t$ is a bounded stochastic process and $R_t=O(1/t^{1+\beta})$ with $\beta>0$. Define a bounded stochastic process $\mathcal{V}=(\mathcal{V}_t)_{t\ge0}$ with dynamics

$$
\mathcal{V}_{t+1}=\left(1-\frac{1}{t+2}\right)\mathcal{V}_t+\frac{1}{t+2}Y_{t+1},\quad t\ge0.
$$

Then

$$
|\mathcal{W}_t-\mathcal{V}_t|=O(1/t)|\mathcal{W}_0-\mathcal{V}_0|+\frac{1}{t}\sum_{n=0}^{t}O(n^{-\beta'}),
$$

where $\beta'=\min(\beta,1)$.

Indeed, we can observe that we can write

$$
\mathcal{W}_{t+1}=\left(1-\frac{1}{t+2}\right)\mathcal{W}_t+\frac{1}{t+2}Y_{t+1}+R'_{t+1},\quad t\ge0,
$$

with $R'_{t+1}=R_{t+1}+O(1/t^2)=O(1/t^{1+\beta'})$ and so

$$
\mathcal{W}_{t+1}=C_{0,t}\mathcal{W}_0+\sum_{n=0}^{t}C_{n+1,t}\frac{Y_{n+1}}{n+2}+\sum_{n=0}^{t}C_{n+1,t}R'_{n+1}
$$

where

$$
C_{0,t}=\prod_{m=0}^{t}\left(1-\frac{1}{m+2}\right)=O(1/t)
$$

$$
C_{n+1,t}=\prod_{m=n+1}^{t}\left(1-\frac{1}{m+2}\right)=\frac{\prod_{m=0}^{t}\left(1-\frac{1}{m+2}\right)}{\prod_{m=0}^{n}\left(1-\frac{1}{m+2}\right)}=O(n/t)
$$

Similarly, we have

$$
\mathcal{V}_{t+1}=C_{0,t}\mathcal{V}_0+\sum_{n=0}^{t}C_{n+1,t}\frac{Y_{n+1}}{n+2}.
$$

Hence, we obtain

$$
\begin{aligned}
|\mathcal{W}_{t+1}-\mathcal{V}_{t+1}| &\le C_{0,t}|\mathcal{W}_0-\mathcal{V}_0|+\left|\sum_{n=0}^{t}C_{n+1,t}R'_{n+1}\right| \\
&= O(1/t)|\mathcal{W}_0-\mathcal{V}_0|+\sum_{n=0}^{t}O\left(\frac{n}{t}\frac{1}{n^{1+\beta'}}\right) \\
&= O(1/t)|\mathcal{W}_0-\mathcal{V}_0|+\frac{1}{t}\sum_{n=0}^{t}O(n^{-\beta'}).
\end{aligned}
$$

Appendix C. The case of a reducible interaction matrix

When the interaction matrix Γ is not irreducible, it is possible to decompose it in irreducible sub-matrices such that the union of the spectra of the sub-matrices coincides with the spectrum of the original matrix. In the following, we will describe a heuristic argument (also employed in Iacopini et al. (2020) and Aletti et al. (2023)), useful in order to detect the rate at which each $S_{t,h}$ grows along time in the case of a general matrix Γ.

The dynamics that rules the vectorial process $\mathbf{S}_t=(S_{t,1},\dots,S_{t,N})^\top$ can be approximated (as $t\to+\infty$) by the linear system of (deterministic) differential equations

$$
\dot{\mathbf{s}}(t)=\Gamma\frac{\mathbf{s}(t)}{t}
$$

Table D.3
For each category h, percentage of patents with a value of the index $I_{n,h}$ greater than $\tau = 0.1, 0.3, 0.5, 0.7, 0.9$.

Category	$(I_{n,h} > 0.1)\%$	$(I_{n,h} > 0.3)\%$	$(I_{n,h} > 0.5)\%$	$(I_{n,h} > 0.7)\%$	$(I_{n,h} > 0.9)\%$
A	1.640%	0.185%	0.045%	0.017%	0.008%
B	2.440%	0.250%	0.059%	0.021%	0.008%
C	1.180%	0.154%	0.042%	0.017%	0.008%
D	0.554%	0.1410%	0.031%	0.017%	0.011%
E	1.070%	0.149%	0.038%	0.017%	0.009%
F	1.420%	0.170%	0.043%	0.018%	0.008%
G	1.450%	0.153%	0.038%	0.016%	0.008%
H	1.350%	0.150%	0.037%	0.016%	0.008%

and hence we can say that $\mathbf{S}_t \approx \mathbf{s}(t)$ for $t \to +\infty$. By the change of variable $t = e^z$, we get

$$\dot{\mathbf{s}}(z) = \Gamma \mathbf{s}(z),$$

whose general solution is given by $\mathbf{s}(z) = e^{\Gamma z}\mathbf{c}$. Now, the term $e^{\Gamma z}$ can be expressed using the canonical Jordan form of the matrix Γ, so that we obtain

$$\mathbf{s}(z) = \sum_{k=1}^{r} e^{\gamma_k z} \sum_{i=0}^{p_k-1} z^i \mathbf{c}_i,$$

where $\gamma_1, \dots, \gamma_r$ are the distinct eigenvalues of Γ, $p_1, \dots, p_r$ are the sizes of the corresponding Jordan blocks and $\mathbf{c}_i$ are suitable vectors related to $\mathbf{c}$ and to the generalized eigenvectors of Γ. Indeed, we can write Γ as PJP^{-1}, where J is its canonical Jordan form and P is a suitable invertible matrix of generalized eigenvectors. Therefore, we have $e^{\Gamma z} = Pe^{Jz}P^{-1}$, where e^{Jz} is a block matrix with blocks of the form $e^{J_k z}$ with J_k block in J. On the other hand, if $J_k = \gamma_k I + N_k$ is a generic Jordan block of Γ with size p_k and associated to the eigenvalue γ_k, we have

$$e^{J_k z} = e^{\gamma_k z} e^{N_k z} = e^{\gamma_k z} \sum_{i=0}^{p_k-1} \frac{z^i}{i!} N_k^i.$$

Changing the variable from z to t, we find

$$\mathbf{S}_t \approx \mathbf{s}(t) = \sum_{k=1}^{r} t^{\gamma_k} \sum_{i=0}^{p_k-1} \ln^i(t) \mathbf{c}_i \tag{C.1}$$

and so the rate at which $S_{t,h}$ increases is given by the leading term in the expression of $s_h(t)$.

Appendix D. Choice of the threshold (robustness check)

In this section we discuss the choice of the threshold τ that has been used in the data analysis of Section 4 to define when a given patent n can be considered a success for (or in) category h, that is, if and only if $I_{n,h} > \tau$. Since we want to identify as a "success" only patents that have an extraordinary impact on at least one category, the value of the threshold τ should be greater than the index value of the vast majority of patents in the data set. However, this requirement is not so stringent in the sense that the percentage of patents with $(I_{n,h} > \tau)$ is already around 1% for $\tau = 0.1$ and is below 0.1% for any $\tau > 0.5$ (see Fig. D.7 and Table D.3 for further details). This means that the majority of patents in the data set have an index value very low (precisely, below 0.1). This is in accordance with Squicciarini et al. (2013), who observed that only a very small subset of patents typically receives a large number of forward citations and the mean value of the forward index decreases along time.

Then, to verify the robustness of the results of the paper with respect to the choice of the threshold τ, we performed the data analysis presented in Section 4 multiple times, each one using a different threshold value. In particular, we have performed a linear regression for every process $S_{t,h}$, in the $\log_{10} - \log_{10}$ scale, allowing the slopes to be different between categories and also imposing a common slope. The results are collected in Fig. D.8, where we can see that the variability presented

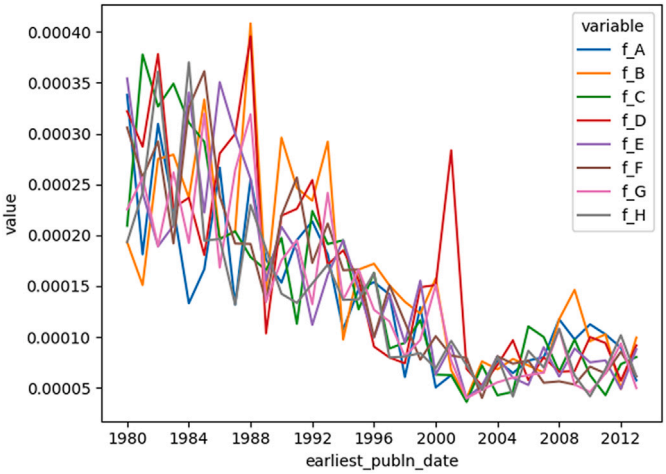


Fig. D.7. Behavior of the mean index per category over the years.

by the slopes is essentially very similar for any value of $\tau > 0.2$. In addition, when we perform the linear regression that imposes the same slope for all categories, the estimated common slope $\hat{\gamma}^*$ is always very close to the value 0.689, that is the one estimated with $\tau = 0.8$ in Section 4. Finally, we have calculated the goodness-of-fit index R^2 obtained imposing a common slope and the one obtained allowing the slopes to be different across categories, and they are always very close to each other and always higher than 0.95 for any value of τ (see Fig. D.9).

Appendix E. Econometric robustness check

As a complementary robustness check, we estimate a reduced-form econometric specification that asks whether the empirical patterns emphasized by the model are also visible in an aggregated panel of technological categories over time. The purpose of the exercise is to verify, in a more standard econometric setting, whether the successes in a target category are positively associated with past successes in the same category (path dependence) and with past successes in the other categories (cross-fertilization) – both results suggested in the main analysis. In other words, the present Appendix investigates whether the two qualitative channels underlying the theoretical framework also emerge in a conventional finite-sample econometric setting: persistence associated with previous success in the same category and predictive association with previous success in other technological categories.

To this end, we aggregate the patent-level success indicators $X_{n,h}$ to the target-category–year level. Let $Y_{h,\tau}$ denote the number of patents published in year τ that qualify as successes in target category h, and let N_τ denote the total number of patents published in year τ.[22] We then define the target-category success rate as

$$r_{h,\tau} = \frac{Y_{h,\tau}}{N_\tau}.$$

Using this rate-based measure, we estimate the following linear specification:

$$r_{h,\tau} = \alpha_h + \lambda_\tau + \beta_1 r_{h,\tau-1} + \beta_2 \sum_{j \neq h} r_{j,\tau-1} + \varepsilon_{h,\tau},$$

[22] Because $Y_{h,\tau}$ and N_τ are contemporaneous flows, the success rate $r_{h,\tau}$ can be constructed without observing the stock of patents or successful patents accumulated before 1980. Consequently, the beginning of the sample does not mechanically truncate the dependent variable, as it would in an analysis based on cumulative patent counts. This does not imply, however, that the pre-1980 technological history is irrelevant; it may still affect post-1980 outcomes through the unobserved initial state of the innovation system.

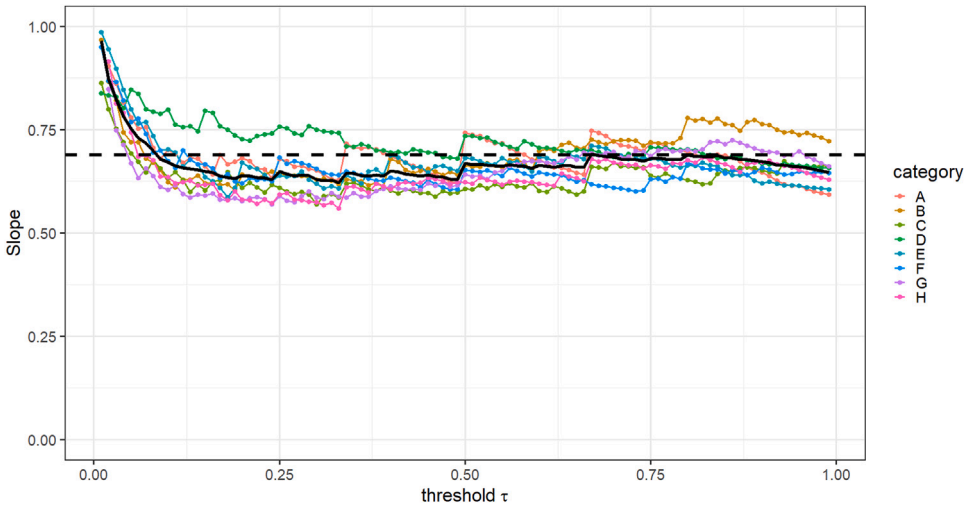


Fig. D.8. Different slopes estimated through linear regression on $S_{t,h}$, in $\log_{10} - \log_{10}$ scale, for each category h (color line) and each value of the threshold τ (x-axis). The solid black line indicates the common slope $\hat{\gamma}^*$ estimated imposing the same slope for all categories. The dashed horizontal black line indicates the common slope $\hat{\gamma}^* = 0.689$ estimated for $\tau = 0.8$ in Section 4.

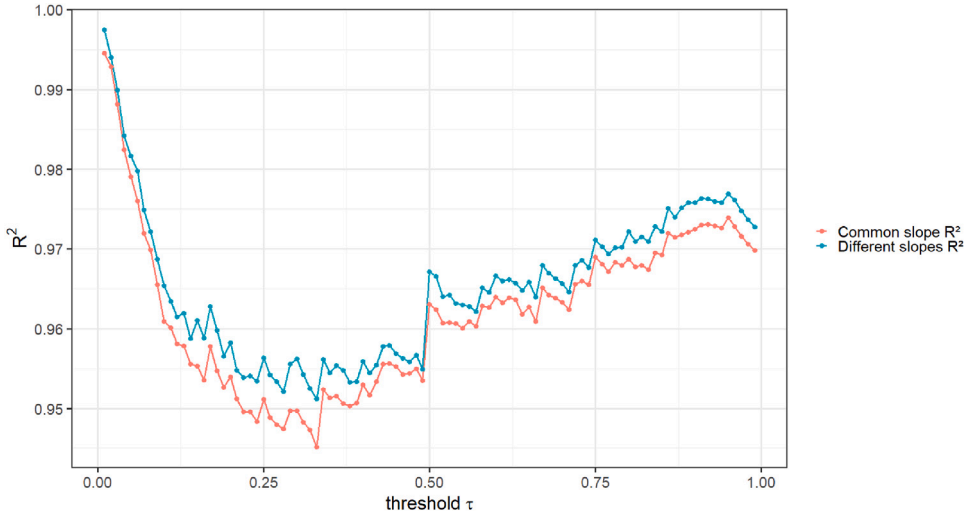


Fig. D.9. Goodness-of-fit index R^2 of the linear regression on $S_{t,h}$, in $\log_{10} - \log_{10}$ scale, for each value of the threshold τ (x-axis). The red line indicates the R^2 obtained imposing a common slope and the blue line the R^2 obtained allowing the slopes to be different across categories. Both the lines are always very close to each other and always higher than 0.95 for any value of τ. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where α_h are category fixed effects and λ_τ are year fixed effects. The coefficient β_1 captures path dependence within category h, while β_2 captures cross-fertilization from the other categories.

This formulation is motivated by the empirical structure of the paper. Since category-specific success is defined for each patent and each target category, the relevant aggregated outcome is the number of successes in target category h per patent opportunity in year τ. The rate specification therefore provides a natural reduced-form counterpart to the paper's interpretation of success becoming rarer or more persistent relative to the scale of inventive opportunities. At the same time, it avoids the extreme sparsity of the patent-level binary specification.

Table E.4 reports the results.

The coefficient on the lagged own-category success rate is positive and marginally significant in all specifications, while the coefficient on the lagged success rate in the other categories is positive and strongly significant. Hence, the reduced-form evidence is consistent with both path dependence and cross-fertilization, with stronger support for the latter. We notice here that these findings should be interpreted as complementary to the main analysis of the paper: they do not replace the model-based results, but they show that the two central mechanisms highlighted by the theoretical framework are also detectable in a conventional econometric specification.

Column (1) regresses the target-category rate on the lag own and other categories' rates. Column (2) adds category and year fixed effects. Finally, Column (3) repeats the same linear rate specification as Column (2), but augments it with category-year characteristics. These controls summarize the observable composition of patenting activity within each technological category and year, including the average size of patent families, the average number of applicants, the average number of inventors, and the relative importance of different filing routes. The results leave the main picture broadly unchanged. The coefficient on the lagged own-category success rate remains positive, although it

Table E.4
Econometric robustness check: target-category success rates.

	(1) Baseline	(2) Baseline + FE	(3) Baseline + Characteristics
Lagged own-category success rate, $r_{h,\tau-1}$	0.162*	0.116*	0.127*
	(0.088)	(0.060)	(0.067)
Lagged other-categories success rate, $\sum_{j\neq h} r_{j,\tau-1}$	0.103***	0.113***	0.137***
	(0.013)	(0.021)	(0.049)
Category FE	No	✓	✓
Year FE	No	✓	✓
Characteristics	No	No	✓
Observations	272	272	252
R^2	0.620	0.771	0.759

Notes: The dependent variable is the success rate in target category h and year τ, defined as $r_{h,\tau} = Y_{h,\tau}/N_\tau$, where $Y_{h,\tau}$ is the number of patents published in year τ that qualify as successes in target category h, and N_τ is the total number of patents published in year τ. Robust standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

becomes less precisely estimated, while the coefficient on the lagged success rate in the other categories remains positive and statistically significant.

Data availability

We have access to GLOBAL PATSTAT which is not public. We can share a subsample.

References

Abbas, A., Zhang, L., Khan, S.U., 2014. A literature review on the state-of-the-art in patent analysis. World Pat. Inf. 37, 3–13.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P., 2005. Competition and innovation: An inverted-U relationship. Q. J. Econ. 120 (2), 701–728.

Aletti, G., Crimaldi, I., Ghiglietti, A., 2017. Synchronization of reinforced stochastic processes with a network-based interaction. Ann. Appl. Probab. 27 (6), 3787–3844. http://dx.doi.org/10.1214/17-AAP1296.

Aletti, G., Crimaldi, I., Ghiglietti, A., 2019. Networks of reinforced stochastic processes: asymptotics for the empirical means. Bernoulli 25 (4 B), 3339–3378. http://dx.doi.org/10.3150/18-BEJ1092.

Aletti, G., Crimaldi, I., Ghiglietti, A., 2023. Interacting innovation processes. Sci. Rep. 13 (1), 17187. http://dx.doi.org/10.1038/s41598-023-43967-1.

Aletti, G., Crimaldi, I., Ghiglietti, A., 2026. Central limit theorems for interacting innovation processes, related statistical tools and general results. Stochastic Process. Appl. 201, 105039. http://dx.doi.org/10.1016/j.spa.2026.105039.

Allen, P., 2013. Complexity, uncertainty and innovation. Econ. Innov. New Technol. 22 (7), 702–725.

Angori, G., Marzocchi, C., Ramaciotti, L., Rizzo, U., 2024. A patent-based analysis of the evolution of basic, mission-oriented, and applied research in European universities. J. Technol. Transf. 49 (2), 609–641.

Autor, D., Dorn, D., Hanson, G.H., Pisano, G., Shu, P., 2020. Foreign competition and domestic innovation: Evidence from US patents. Am. Econ. Rev.: Insights 2 (3), 357–374.

Barbieri, N., Consoli, D., Napolitano, L., et al., 2023. Regional technological capabilities and green opportunities in Europe. J. Technol. Transf. 48, 749–778.

Bekkers, R., Martinelli, A., Tamagni, F., 2020. The impact of including standards-related documentation in patent prior art: Evidence from an EPO policy change. Res. Policy 49 (7), 104007.

Bloom, N., Jones, C.I., Van Reenen, J., Webb, M., 2020. Are ideas getting harder to find? Am. Econ. Rev. 110 (4), 1104–1144.

Brem, A., Viardot, E., Nylund, P.A., 2021. Implications of the coronavirus (COVID-19) outbreak for innovation: Which technologies will improve our lives? Technol. Forecast. Soc. Change 163, 120451. http://dx.doi.org/10.1016/j.techfore.2020.120451, URL: https://www.sciencedirect.com/science/article/pii/S0040162520312774.

Breschi, S., Lissoni, F., Malerba, F., 2003. Knowledge-relatedness in firm technological diversification. Res. Policy 32 (1), 69–87.

Chae, S., Gim, J., 2019. A study on trend analysis of applicants based on patent classification systems. Information 10 (12), 364.

Colladon, A.F., Guardabascio, B., Venturini, F., 2025. A new mapping of technological interdependence. Res. Policy 54 (1), 105126.

Corrocher, N., Mancusi, M.L., 2021. International collaborations in green energy technologies: What is the role of distance in environmental policy stringency? Energy Policy 156, 112470.

Dahlin, K.B., Behrens, D.M., 2005. When is an invention really radical?: Defining and measuring technological radicalness. Res. Policy 34 (5), 717–737. http://dx.doi.org/10.1016/j.respol.2005.03.009, URL: https://www.sciencedirect.com/science/article/pii/S0048733305000764.

Dehghani, M.A., Karavidas, D., Panourgias, N., Hutchinson, M., O'Reilly, P., 2023. Assessing the quality of financial technology patents through the development of a patent quality index for comparing jurisdictions, technical domains, and leading organizations. IEEE Trans. Eng. Manage. 71, 3934–3950.

Dovbischuk, I., 2022. Innovation-oriented dynamic capabilities of logistics service providers, dynamic resilience and firm performance during the COVID-19 pandemic. Int. J. Logist. Manag. 33 (2), 499–519.

Evenson, R.E., 1968. The Contribution of Agricultural Research and Extension to Agricultural Production (Ph.D. thesis). The University of Chicago.

Fleming, L., 2001. Recombinant uncertainty in technological search. Manag. Sci. 47 (1), 117–132. http://dx.doi.org/10.1287/mnsc.47.1.117.10671.

Gao, Y., Lazarova, E., 2024. Ex-ante Novelty and Invention Quality: A Cross-country Sectoral Empirical Study. University of East Anglia School of Economics Working Paper Series 2024–02, School of Economics, University of East Anglia, Norwich, UK. URL: https://ideas.repec.org/p/uea/ueaeco/2024-02.html.

Gerken, J.M., Moehrle, M.G., 2012. A new instrument for technology monitoring: novelty in patents measured by semantic patent analysis. Scientometrics 91 (3), 645–670.

Gouet, R., 1993. Martingale functional central limit theorems for a generalized Pólya urn. Ann. Probab. 1624–1639.

Griliches, Z., 1958. Research costs and social returns: Hybrid corn and related innovations. J. Political Econ. 66 (5), 419–431.

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

Hall, B.H., Jaffe, A.B., Trajtenberg, M., 2000. Market Value and Patent Citations: A First Look. http://dx.doi.org/10.3386/w7741, Working Paper 7741, National Bureau of Economic Research, URL: http://www.nber.org/papers/w7741.

Higham, K., De Rassenfosse, G., Jaffe, A.B., 2021. Patent quality: Towards a systematic framework for analysis and measurement. Res. Policy 50 (4), 104215.

Ho, C.M., 2023. Research on interaction of innovation spillovers in the AI, Fin-Tech, and IoT industries: considering structural changes accelerated by COVID-19. Financ. Innov. 9 (1), 7.

Iacopini, I., Di Bona, G., Ubaldi, E., Loreto, V., Latora, V., 2020. Interacting discovery processes on complex networks. Phys. Rev. Lett. 125, 248301. http://dx.doi.org/10.1103/PhysRevLett.125.248301.

Jaffe, A.B., 1986. Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value. National Bureau of Economic Research Cambridge, Mass., USA.

Jaffe, A.B., 1989. Characterizing the "technological position" of firms, with application to quantifying technological opportunity and research spillovers. Res. Policy 18 (2), 87–97.

Jalonen, H., 2012. The uncertainty of innovation: a systematic review of the literature. J. Manag. Res. 4 (1), 1–47.

Jalonen, H., Lehtonen, A., 2011. Uncertainty in the innovation process. In: European Conference on Innovation and Entrepreneurship. Academic Conferences International Limited, 51–51.

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

Jones, C.I., 1995b. Time series tests of endogenous growth models. Q. J. Econ. 110 (2), 495–525.

Jones, B.F., 2009. The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? Rev. Econ. Stud. 76 (1), 283–317.

Kauffman, S.A., 2000. Investigations. Oxford University Press.

Kelly, B., Papanikolaou, D., Seru, A., Taddy, M., 2021. Measuring technological innovation over the long run. Am. Econ. Rev.: Insights 3 (3), 303–320.

Kim, J., Magee, C.L., 2017. Dynamic patterns of knowledge flows across technological domains: empirical results and link prediction. arXiv:1706.07140, URL: https://arxiv.org/abs/1706.07140.

Lafond, F., Kim, D., 2019. Long-run dynamics of the US patent classification system. J. Evol. Econ. 29 (2), 631–664.

Lanjouw, J.O., Schankerman, M., 2004. Patent quality and research productivity: Measuring innovation with multiple indicators. Econ. J. 114 (495), 441–465.
Lee, S.M., Trimi, S., 2021. Convergence innovation in the digital age and in the COVID-19 pandemic crisis. J. Bus. Res. 123, 14–22.
Levin, R.C., Cohen, W.M., Mowery, D.C., 1985. R & D appropriability, opportunity, and market structure: new evidence on some schumpeterian hypotheses. Am. Econ. Rev. 75 (2), 20–24.
Leydesdorff, L., Kogler, D.F., Yan, B., 2017. Mapping patent classifications: portfolio and statistical analysis, and the comparison of strengths and weaknesses. Scientometrics 112, 1573–1591.
Li, L., Tong, Y., Wei, L., Yang, S., 2022. Digital technology-enabled dynamic capabilities and their impacts on firm performance: Evidence from the COVID-19 pandemic. Inf. Manag. 59 (8), 103689.
Madsen, J., Minniti, A., Venturini, F., 2024. Declining research productivity and income inequality: A centenary perspective. J. Econom. Dynam. Control 167, 104924.
Malerba, F., 2002. Sectoral systems of innovation and production. Res. Policy 31 (2), 247–264.
Malerba, F., Orsenigo, L., 1996. Schumpeterian patterns of innovation are technology-specific. Res. Policy 25 (3), 451–478.
Monechi, B., Ruiz-Serrano, A., Tria, F., Loreto, V., 2017. Waves of novelties in the expansion into the adjacent possible. PloS One 12 (6), e0179303.
Mueller, J., 2024. Aspen Treatise for Patent Law. Aspen Publishing, Aspen Treatise Series.
Napolitano, L., Evangelou, E., Pugliese, E., Zeppini, P., Room, G., 2018. Technology networks: The autocatalytic origins of innovation. R. Soc. Open Sci. 5 (6), 172445.
Park, M., Leahey, E., Funk, R.J., 2023. Papers and patents are becoming less disruptive over time. Nature 613 (7942), 138–144.
Pemantle, R., 2007. A survey of random processes with reinforcement. Probab. Surv. 4 (none), 1–79. http://dx.doi.org/10.1214/07-PS094.
Pichler, A., Lafond, F., Farmer, J.D., 2020. Technological interdependencies predict innovation dynamics. arXiv preprint arXiv:2003.00580.
Rainville, A., Dikker, I., Buggenhagen, M., 2025. Tracking innovation via green patent classification systems: Are we truly capturing circular economy progress? J. Clean. Prod. 486, 144385.
Robbins, H., Siegmund, D., 1971. A convergence theorem for non negative almost supermartingales and some applications. In: Rustagi, J.S. (Ed.), Optimizing Methods in Statistics. Academic Press, New York, pp. 233–257. http://dx.doi.org/10.1016/B978-0-12-604550-5.50015-8.
Rosenberg, N., 1979. Technological interdependence in the American economy. Technol. Cult. 20 (1), 25–50.
Sampat, B.N., Ziedonis, A.A., 2005. Patent citations and the economic value of patents. In: Moed, H.F., Glänzel, W., Schmoch, U. (Eds.), Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems. Springer Netherlands, Dordrecht, pp. 277–298. http://dx.doi.org/10.1007/1-4020-2755-9_13.
Scherer, F.M., 1982. Inter-industry technology flows in the United States. Res. Policy 11 (4), 227–245.
Schmookler, J., 1954. The level of inventive activity. Rev. Econ. Stat. 183–190.
Squicciarini, M., Dernis, H., Criscuolo, C., 2013. Measuring patent quality: Indicators of technological and economic value. http://dx.doi.org/10.1787/5k4522wkw1r8-en, Technology and Industry Working Papers 2013/3, OECD Publishing.
Stafford, A.B., 1952. Is the rate of invention declining? Am. J. Sociol. 57 (6), 539–545.
Taalbi, J., 2020. Evolution and structure of technological systems-an innovation output network. Res. Policy 49 (8), 104010.
Trajtenberg, M., Henderson, R., Jaffe, A., 1997. University versus corporate patents: A window on the basicness of invention. Econ. Innov. New Technol. 5 (1), 19–50.
Verhoeven, D., Bakker, J., Veugelers, R., 2016. Measuring technological novelty with patent-based indicators. Res. Policy 45 (3), 707–723.
Williams, D., 1991. Probability with Martingales. Cambridge University Press, Cambridge.
Zhang, C., Zhang, D., Pan, Y., Wang, Y., 2025. Whom you connect with matters: Innovation collaboration network centrality and innovative productivity in Chinese cities. Growth Chang. 56 (1), e70015.