

Dynastic dynamics: Modelling powerful naming choices with stochastic prestige

Pablo Rosillo-Rodes,^{1,2,3,*} Santiago Lamata-Otín,^{4,5,3,*} David Soriano-Paños,^{6,5}
Laurent Hébert-Dufresne,^{3,7,8} and Peter Sheridan Dodds^{2,3,7,8}

¹*Institute for Cross-Disciplinary Physics and Complex Systems IFISC (UIB-CSIC),
Campus Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain*

²*Computational Story Lab, University of Vermont, Burlington, VT 05405, US*

³*Vermont Complex Systems Institute, University of Vermont, USA*

⁴*Department of Condensed Matter Physics, University of Zaragoza, 50009 Zaragoza, Spain*

⁵*GOTHAM lab, Institute of Biocomputation and Physics of Complex
Systems (BIFI), University of Zaragoza, 50018 Zaragoza, Spain*

⁶*Departament d'Enginyeria Informàtica i Matemàtiques,
Universitat Rovira i Virgili, 43007 Tarragona, Spain*

⁷*Santa Fe Institute, 1399 Hyde Park Rd, Santa Fe, NM 87501, US*

⁸*Complexity Science Hub, Metternichgasse 8, 1030 Vienna, Austria*

(Dated: July 10, 2026)

The naming choices of powerful rulers encode millennia of cultural, political, and institutional influences. Consequently, the long-term onomastic dynamics within some of the most powerful dynasties cannot be explained by simple mechanisms such as frequency-dependent reinforcement or random name reuse. Here, we propose a minimal model that encapsulates these influences in a single highly stochastic variable, prestige, and assume that name selection is driven by the prestige accumulated by previous rulers bearing the same name within each dynasty. Using an extensive dataset spanning ten dynasties, we show that the model reproduces the pronounced inequalities in name frequencies, the long-term persistence of dominant names, and the abrupt rises in popularity observed in historical records. Our results suggest that the naming traditions of powerful rulers preserve institutional memory through reinforcement while remaining sensitive to rare historical events that can reshape the hierarchy of names across generations.

INTRODUCTION

How do the powerful choose their names? The naming choices of political and spiritual rulers constitute a paradigmatic case study in cultural evolution [1, 2]. The selection of a regnal or ruling name forms a discrete sequence of institutional choices, shaped by the interplay of power, tradition, and reform [3–5]. In this sense, the names of the powerful are symbolic assets rather than neutral labels. They encode social memory, institutional continuity, and collective values, transforming each act of naming into a public claim about inheritance, legitimacy, and change [6–9].

Throughout history, rulers have routinely chosen names linked to the land, the divine, and conflict, cultivating associations with strength, perfection, protection, power, and military victory [8]. In ancient Egypt, pharaonic names were used to announce royal policies and to imitate the nomenclature of illustrious predecessors [8]. In Roman imperial history, names likewise evolved for social and political reasons, often serving to consolidate the civil and military legacy of earlier rulers [10, 11]. A similar logic appears in the centuries-long sequence of pontificates during which newly elected popes repeatedly adopted names already chosen by their predecessors [3, 4, 12]. Taken together, these examples showcase that the accumulated prestige of a name is a central mechanism behind dynastic naming traditions.

To quantify how past prestige weighs on present choices, we frame ruling names as onomastic time series whose evolution can be studied through the lens of quantitative complex-systems approaches. Such approaches have been used to model the evolution and governance of societies, language, and culture [9, 13–23]. In particular, quantitative studies of onomastics have examined cultural drift or papal-name prediction, among other phenomena [24–28]. Yet, the long-term dynamics of ruling-name choices remain largely unexplored.

A natural formal language for modeling such repeated choices is provided by urn models, in which options are repeatedly selected from a repertoire that may change over time [29–31]. Urn schemes encompass mechanisms ranging from simple random sampling with [32] or without [33, 34] replacement to reinforcement processes in which past selections bias future ones [35–39], including heterogeneous reinforcement distributions [40, 41] and extensions that include innovation through the introduction of new elements in the repertoire [42–44]. In complex systems, reinforcement is commonly associated with rich-get-richer (RGR) dynamics and the Matthew effect [45], providing a minimal mechanism for heavy-tailed distributions and scaling laws across many empirical domains [46–52]; related principles also appear in cumulative-advantage models [53–55] and growing-network models [56, 57]. However, in regnal naming, reinforcement is not only frequency-based: rare

but highly prestigious reigns may abruptly reshape name popularity and disrupt existing name hierarchies.

In this work, we introduce a minimal stochastic model formulated within the framework of randomly reinforced urns [39] with innovation, in which reinforcement is interpreted as accumulated regnal prestige. Crucially, prestige is modeled through a heterogeneous heavy-tailed distribution unlike most reinforced urn model assuming uniform reinforcements. We validate our model with a broad set of powerful ruling dynasties, including Egyptian pharaohs, Roman emperors, popes, Russian tsars, patriarchs of Constantinople, and monarchs of several regions. Using the calibrated parameters of the model, we show that the role of prestige heterogeneity is more pronounced in non-hereditary dynasties, whereby the establishment of successful lineages is more shaped by the individual performance of the predecessor than by the historical name frequency within the dynasty.

RESULTS

Onomastics of powerful dynasties

We analyze the name timelines across 10 distinct ruling systems: Egyptian pharaohs [58], Roman Emperors [59], popes [12], Russian tsars [60], the patriarchs of Constantinople [61], Ottoman sultans, Holy Roman emperors, and the monarchs of Spain [62], Denmark [63], and England [64]. Each discrete event in these timelines corresponds to the succession of a new ruler following the death, abdication, or deposition of their predecessor. We show in Fig. 1 the dynastic dynamics for popes (top row), Egyptian pharaohs (middle row), and Roman emperors (bottom row), and we present the remaining datasets in Supplementary Fig. 1 in Supplementary Note 1.

Fig. 1a-c displays the temporal dynamics of name abundance, showing that name usage is bursty rather than smooth. Names can remain dormant for long periods before being rapidly reused, generating cascades in their cumulative abundance. At the same time, new names continue to enter the repertoire, highlighting the interplay between innovation and reuse. This balance is summarized by the evolution of the number of distinct names $M(t)$, as we show in the insets of Fig. 1d-f. At early times, $M(t)$ grows rapidly, indicating a regime dominated by onomastic innovation. As the sequences progress, this growth slows down, and the curves bend away from linearity, showing that the introduction of new names becomes less frequent and reuse increasingly dominates the dynamics. In hereditary monarchies, extended periods of low innovation often reflect dynastic constraints and lineage-based naming traditions (see Supplementary Fig. 1). In contrast, the papal sequence reflects the consolidation of an institutional tradition of

choosing names already present in the repertoire, leading to a long period of almost uninterrupted reuse until the election of Francis I in 2013.

Finally, in Fig. 1g-i we show that bursty reuse produces increasing inequality in name frequencies. The variance of the name-frequency distribution $\mathcal{V}(t)$ grows over time (see top rows), indicating that repeated use progressively concentrates on a small subset of names. Consistently, the rank-frequency ($r-f$) distributions measured at the end of the simulations display a broad, power-law-like hierarchy, compatible with the data. Here, names are ordered by descending frequency ($r=1$ for the most common), revealing a structure dominated by a few highly frequent names alongside a long tail of rare ones.

Modeling rare but prestigious reigns

To understand these patterns, we introduce the Reinforcement and Innovation through Prestige (RIP) model, a minimal stochastic framework assuming that naming choices are determined by the interplay between innovation dynamics and prestige-driven name reuse. Note that, depending on the historical dataset, the meaning of prestige can vary (see Supplementary Note 1). In our model, at each step, with probability $\rho(t)$ a succession event introduces a previously unused name rather than reusing one from the existing repertoire. Otherwise, an already used name is chosen according to the accumulated prestige by its predecessors. After a name is chosen, the prestige of that name is incremented considering a prestige-increment distribution $P(K)$, with mean given by $\mu = \mathbb{E}[K]$ and variance $\sigma^2 = \text{Var}(K) = \mathbb{E}[K^2] - \mu^2$. Note that $\sigma \ll \mu$ ($\sigma \gg \mu$) produces homogeneous (heterogeneous) individual contributions to the prestige accumulated by each name in the dynasty. Note that, when $\sigma^2 = 0$, prestige becomes proportional to abundance, and this recovers rich-get-richer dynamics.

In Fig. 1d-f we show a realization of the model for each dataset using the innovation function $\rho(t)$ averaged over a centered rolling window of length 3 as input data and the distribution $P(K)$ fitted as described in Methods. Comparing with the real data in Fig. 1a-c, we appreciate that the RIP model captures the abrupt rises in the long-term popularity of individual names. These cascades are then a direct consequence of rare events of prestige gain, since heavy-tailed reinforcement naturally leads to extended periods dominated by a single name.

Furthermore, the top rows in Fig. 1g-i show that the RIP model also matches the macroscopic statistics of onomastics. The variance of the simulated name distribution $\bar{\mathcal{V}}(t)$ averaged over 1000 runs of the model aligns closely with the data. Moreover, in Eqs. (15)-(S.52) in the Methods section, we have derived an analytical pre-

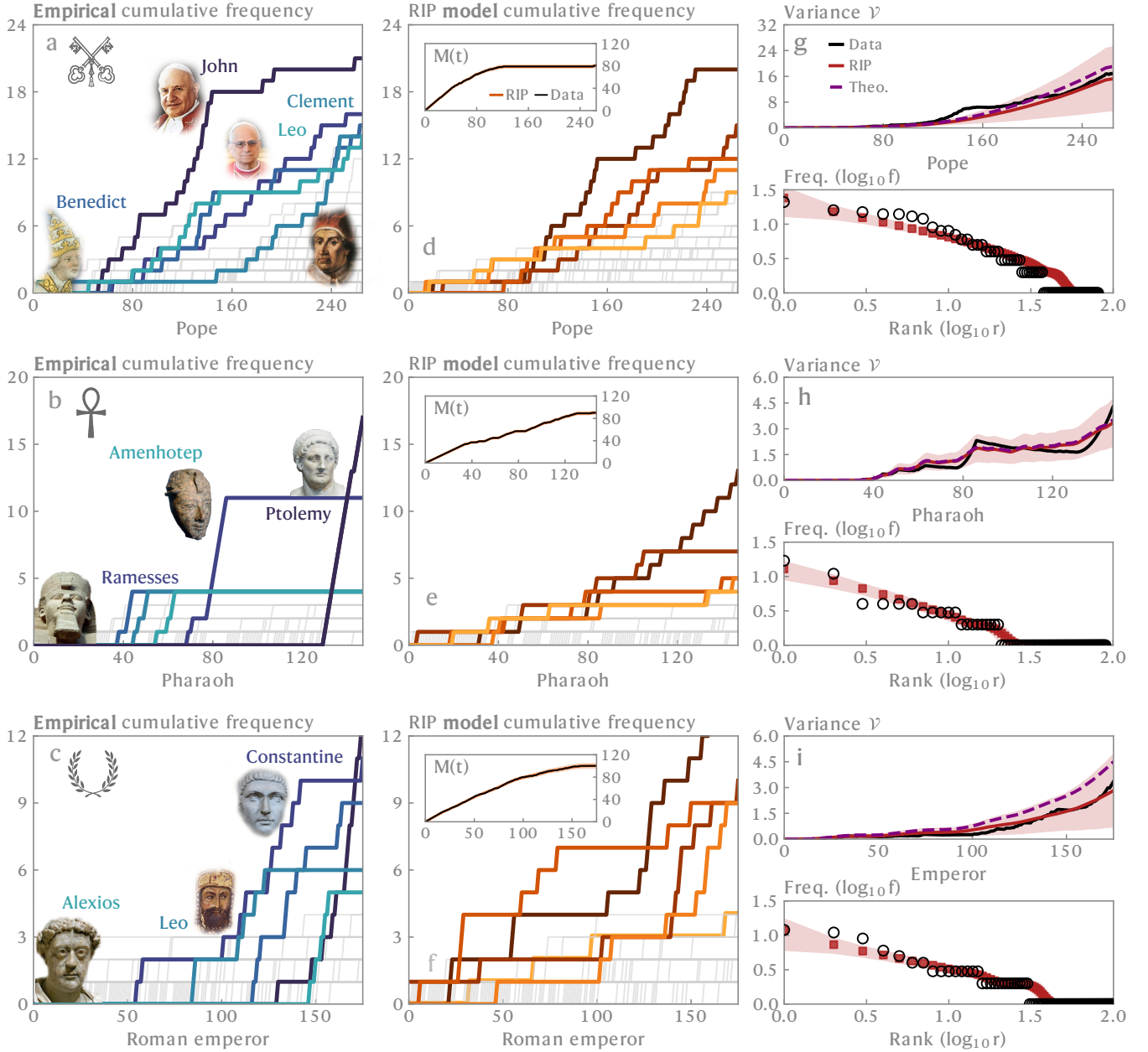


FIG. 1. **Innovation and rare but prestigious reigns explain the onomastics of powerful dynasties.** **a-c** Number of popes, Egyptian pharaohs, and Roman emperors respectively, bearing each name across chronologically ordered reigns. **d-f** Output of a single run of the RIP model. The insets show the evolution of the number of distinct names ($M(t)$ in bottom rows) over time. **g-i** Top: Variance of name frequencies in the data compared with 1000 simulations of the RIP model and with the analytical expression from Eqs. (15)-(S.52). Bottom: rank-frequency distribution of names. The plots show the frequency f_i of each name i as a function of its rank r_i , comparing the empirical data with the mean outcome of 1000 simulations of the RIP model $\pm 1\sigma$. Note that we standardize the name data by isolating the primary component of authority in each dynastic tradition (see Methods).

diction for the time evolution of the variance that fairly reproduces the data. From this analytical treatment we also find that, for constant innovation $\rho(t) = \rho$, the variance is asymptotically bounded only above a critical rate $\rho_c = 1/2$. Below this value, $\bar{\mathcal{V}}(t)$ diverges as $t \rightarrow \infty$. Since innovation decays below ρ_c in all datasets (see

Supplementary Fig. 3), the continued growth of the variance trajectories in Fig. 1g-i is consistent with the model framework.

Finally, the RIP model produces strong inequality in the frequency of names compatible with the one observed in the actual data (see the bottom rows in Fig. 1g-

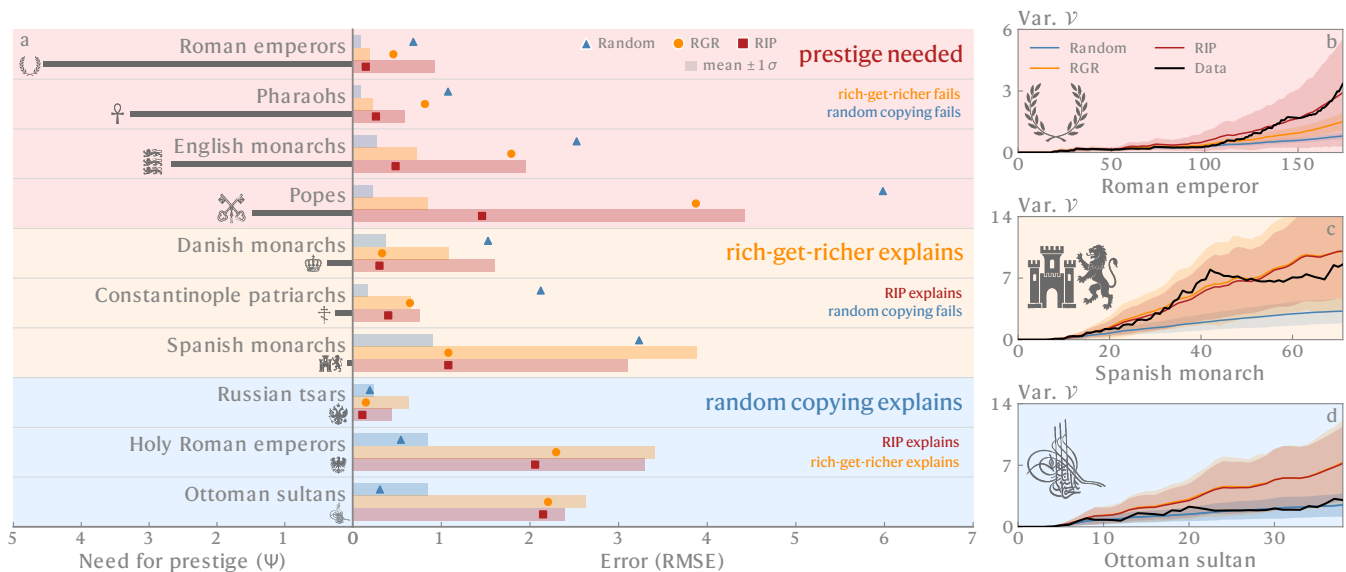


FIG. 2. **Prestige explains what random copying and rich-get-richer dynamics miss about onomastics.** **a** Root mean square error (RMSE) of the variance of the name distribution for the random, rich-get-richer, and RIP models. Shaded bars represent the RMSE that a trajectory at distance $\pm 1\sigma$ of the mean of the trajectory would produce. If a point is inside its corresponding bar, the empirical trajectory is compatible with a random realization of the model at 1σ confidence. Prestige bars below labels indicate the prestige indicator Ψ . **b-d** Examples of datasets where prestige is or is not required for explaining onomastics. Background colors group the traditions based on the simplest mechanism that accounts for their variance trajectory. Specifically, blue denotes cases explained by the random model (as well as rich-get-richer and RIP), orange refers to dynamics accounted for by rich-get-richer (and RIP), and red highlights instances where prestige (RIP) is explicitly required to explain the trajectory.

i). In the Supplementary Fig. 1 we show the agreement for the remaining datasets.

On the necessity of rare but prestigious reigns

The agreement between the RIP model and the empirical trajectories raises a natural question: is heterogeneous prestige necessary, or can the observed dynamics be explained by simpler reuse mechanisms? To address this, we compare the RIP model with two baselines: a random-choice model, in which reused names are selected uniformly from the existing repertoire, and a rich-get-richer model, in which reused names are selected proportionally to their previous abundance (note that this corresponds to the limit case $\sigma^2 = 0$ in the RIP model). In Fig. 2a we plot with a point the root mean square error (RMSE) of the average variance of 1000 simulations of the name distribution with respect to the actual data for the random, rich-get-richer, and RIP models, and with shaded bars we plot the RMSE that a trajectory at distance $\pm 1\sigma$ of the average of the simulations would produce. In this sense, if the point is inside the bar, it means that the empirical data is compatible at 1σ confidence with an arbitrary realization of the model. Supplementary Fig. 2 in Supplementary Note 2 shows an alternative benchmark based on Bayesian information criteria.

In all datasets except Ottoman sultans, Russian tsars, and Holy Roman emperors, the Random model's tendency is not consistent with the variance evolution. In the aforementioned three specific datasets, the dynamics do not require either prestige or a rich-get-richer mechanism to be explained. Indeed, Fig. 2d shows that the Ottoman sultans trend is better explained by the Random case. Out of the other datasets, the rich-get-richer mechanism with innovation turns out to be as likely as the RIP model to explain Danish monarchs, Constantinople patriarchs, and Spanish monarchs (see Fig. 2c to see the evolution of the variance for the Spanish monarchs). More importantly, however, for several of the most powerful dynasties in history (Pharaohs, English monarchs, Roman emperors, and popes), prestige is crucially required for explaining dynastic dynamics.

Notice that in those datasets where the RIP model outperforms the RGR and Random alternatives it is due to the fact that the variance of the distribution of names in the RIP model (\bar{V}_{RIP}) is substantially larger than what RGR (\bar{V}_{RGR}) and Random (\bar{V}_{R}) models can produce (see Fig. 2b for the Roman emperors case). We can compare the performance of the three models analytically when assuming that the prestige-driven model starts from an initial condition such that the initial prestiges are given by the abundances multiplied by the average of the prestige distribution, μ . As derived in Sup-

plementary Note 3, we obtain that

$$\bar{V}_{\text{RIP}}(t) \geq \bar{V}_{\text{RGR}}(t) \geq \bar{V}_{\text{R}}(t). \quad (1)$$

The first inequality follows from the additional heterogeneity induced by random prestige increments, whereas the second follows from the reinforcement of already abundant names in the rich-get-richer mechanism. Random reuse lacks both amplification mechanisms and therefore produces the smallest abundance variance within this comparison.

The condition in Eq. (1) holds in general. Nonetheless, to make the first inequality explicit, we consider an analytically tractable non-innovation regime, such as the one appearing in papal names for over a millennium (see the inset in Fig. 1d). In this closed-repertoire limit, the evolution of the variance is fully described by the number of initial name choices N_0 and the number of different names M_0 . Starting from the same initial conditions, we derive in the Supplementary Note 3 that the variance in both the prestige-based and the rich-get-richer models grows quadratically with time (see Eqs. (21)-(22) in Methods). From those expressions, the excess heterogeneity generated by the prestige model is given by

$$\bar{V}_{\text{RIP}}(t) - \bar{V}_{\text{RGR}}(t) \sim \frac{1}{M_0 N_0} \frac{\sigma^2}{\mu^2} (N_0 + t)^2. \quad (2)$$

Equation (2) shows that, although both mechanisms produce quadratic variance growth in the absence of innovation, heterogeneous prestige increments increase the prefactor of this growth. Thus, prestige-driven reinforcement yields a strictly larger variance than rich-get-richer dynamics whenever $\sigma^2 > 0$.

Based on Eq. (1), we can define an indicator for the need of prestige to describe the dynamics in terms of the first two moments of the prestige distribution as $\Psi = \sigma^2/\mu^2$. Below each label in Fig. 2a, we plot a bar with length Ψ . Clearly, those datasets that are better explained by the RIP model are those with a larger value of Ψ , meaning that in those cases where the prediction of the RIP model is similar to that of the rich-get-richer or random models are those where the distribution of prestiges that best describes the data does not feature rare events with large prestige.

DISCUSSION

For a ruler, choosing a name is both a symbolic act and a historical marker. Our analysis indicates that long-term patterns in dynastic names can be described using statistical features found in all kinds of natural systems, being stochastic extreme and rare increments in the accumulated prestige of some lineages necessary to describe many powerful dynasties.

The Reinforcement and Innovation through Prestige (RIP) model offers a compact quantitative framework

to understand the onomastic trends in a wide range of powerful dynasties. With minimal assumptions, it reproduces the marked inequality in name frequencies, the long persistence of preferred names, and the abrupt changes triggered by rare prestige increments. Rather than reducing these patterns to standard frequency-based reinforcement, the RIP model defines a prestige-driven reinforcement process in which highly stochastic prestige increments drive naming choices.

The same formulation also admits a growing network interpretation, where rulers are prestige-endowed nodes, and name lineages are connected components whose growth is driven by the accumulated prestige of their members (see Supplementary Figs. 4 and 5 in Supplementary Note 4). In this sense, the model also provides a framework for network growth based on accumulated stochastic fitness, distinct from pure structural rich-get-richer dynamics [56, 57] or hybrid approaches where fitness acts as a multiplicative factor of connectivity [65].

We analyze several different naming traditions. While both secular and religious naming systems transition from initial linear innovation to periods with zero-innovation plateaus, their underlying mechanisms diverge. Secular monarchies rely on dynastic conservation, creating shorter, intermittent plateaus of name reuse. Conversely, the papacy features a unique, millennial-long institutional plateau of complete stagnation that was only broken by Pope Francis I (see Supplementary Note 1 for more context). These structural variations reveal why some traditions can be modeled by a rich-get-richer framework, whereas other systems like the papacy or Egyptian pharaohs require extreme, event-driven spikes in name prestige to introduce new trends.

Overall, in this work we show how historical sources examined through methods from complex systems can reveal stable structures beneath centuries of political change, cultural evolution and theological shifts. This framework can be applied to other naming traditions, leadership sequences, or institutional lineages, opening a quantitative path for analysing the evolution of social and symbolic conventions. Our results also suggest that long-lived symbolic institutions may evolve through the interplay of steady background dynamics and rare, high-impact events.

Crucially, our work addresses a fundamental open question in cultural evolution: whether random sampling is the primary driver of stochasticity compared to factors like innovation, bursty trends, or decision-making [52]. We approach this problem through the lens of onomastics, offering unique empirical insights drawn from the naming data of powerful historical dynasties: random sampling with innovation is not enough to describe the naming dynamics of some of the most powerful dynasties. However, the phenomenon of name prestige, for-

mulated here as stochastic extreme and rare reinforcing events, is not confined to royal naming conventions. Instead, our framework may apply broadly to socio-technical and cultural analyses of innovation, trends, and social norms, aligning with recent investigations into these dynamic processes [66–69].

DATA AVAILABILITY STATEMENT

The datasets analyzed and codes used during the current study are available in the `dynastic-dynamics` Github repository, accesible at <https://github.com/pablorosillo/dynastic-dynamics/>.

ACKNOWLEDGMENTS

We thank Alberto Aparici for valuable insights on the onomastic traditions of the Roman Empire and Teresa Benet-Casaña for guidance on catholic and papal matters. We also thank Pedro Jiménez-González, David Sánchez, Natalia Briñas-Pascual, Leah Keating, Federico Malizia, and Sara Oliver-Bonafoux for their comments on a first version of the manuscript. P.R.-R. and S.L.-O. thank the Vermont Complex Systems Institute for bringing them together in 2025 Fall term, the St. Joseph Oratory in Montreal (CA) for bringing inspiration, and Gema and Izarbe for enjoying the culmination of this manuscript in the *pirineu aragonés*. P.R.-R. acknowledges support by the Spanish State Research Agency (MICIU/AEI/10.13039/501100011033) and FEDER (UE) under the María de Maeztu project CEX2021-001164-M and the project COSASTI (PID2024-157493NB-C21 & PID2024-157493NB-C22). S.L.-O. acknowledges financial support from the Departamento de Industria e Innovación del Gobierno de Aragón y Fondo Social Europeo (FENOL group grant E36-23R), from Ministerio de Ciencia e Innovación (grant PID2020-113582GB-I00), and from Gobierno de Aragón through a doctoral fellowship. D.S.-P acknowledge Spanish Ministerio de Ciencia e Innovación (PID2024-158120NB-C21). We are grateful for National Science Foundation Award #2242829 (Science of Online Corpora, Knowledge, and Stories) and award #2419733, foundational support from MassMutual, and an anonymous philanthropic gift.

* These two authors contributed equally. Correspondence: prosillo@ifisc.uib-csic.es, santiagolaot@gmail.com.

- [1] S. Lieberson, *A matter of taste: How names, fashions, and culture change* (Yale University Press, 2000).
 [2] R. Alford, *Naming and identity: A cross-cultural study of personal naming practices*, Hraf Press (1987).

- [3] T. Nau, The names of the popes, *Onomastica Canadiana* **75** (1993).
 [4] C. De Vinne, Papal self-naming: Genesis of a tradition, *Onomastica Canadiana* **88** (2006).
 [5] R. L. Poole, The names and numbers of medieval popes, *The English Historical Review* **32**, 465 (1917).
 [6] P. Bourdieu, *Language and symbolic power* (Harvard University Press, 1991).
 [7] M. W. Hahn and R. A. Bentley, Drift as a mechanism for cultural change: an example from baby names, *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, S120 (2003).
 [8] Heba Nayef and Mohamed El-Nashar, “Towers of Strength and More”: A Thematic Analysis of Royal Titularies in Ancient Egypt, *International Journal of Society, Culture & Language* 10.22034/ijsc.2023.1990883.2952 (2023).
 [9] Peter Sheridan Dodds and Joshua R. Minot and Michael V. Arnold and Thayer Alshaabi and Jane Lydia Adams and David Rushing Dewhurst, Allotaxonomy and rank-turbulence divergence: a universal instrument for comparing complex systems, *EPJ Data Science* **12**, 37 (2023).
 [10] Mason Hammond, Imperial Elements in the Formula of the Roman Emperors during the First Two and a Half Centuries of the Empire, *Memoirs of the American Academy in Rome* **25**, 17 (1957).
 [11] Benet Salway, What’s in a Name? A Survey of Roman Onomastic Practice from c. 700 B.C. to A.D. 700, *The Journal of Roman Studies* **84**, 124 (1994).
 [12] The List of Popes, <http://www.newadvent.org/cathen/12272b.htm> (1911), in *The Catholic Encyclopedia*. New York: Robert Appleton Company. Retrieved online October 31, 2025.
 [13] B. E. Kendall, C. J. Briggs, W. W. Murdoch, P. Turchin, S. P. Ellner, E. McCauley, R. M. Nisbet, and S. N. Wood, Why do populations cycle? a synthesis of statistical and mechanistic modeling approaches, *Ecology* **80**, 1789 (1999).
 [14] P. Turchin, T. E. Currie, E. A. L. Turner, and S. Gavrillets, War, space, and the evolution of old world complex societies, *Proceedings of the National Academy of Sciences* **110**, 16384 (2013).
 [15] S. Roman, S. Bullock, and M. Brede, Coupled Societies are More Robust Against Collapse: A Hypothetical Look at Easter Island, *Ecological Economics* **132**, 264 (2017).
 [16] S. Roman, Historical dynamics of the Chinese dynasties, *Heliyon* **7**, e07293 (2021).
 [17] P. Turchin, H. Whitehouse, S. Gavrillets, D. Hoyer, P. François, J. S. Bennett, K. C. Feeney, P. Peregrine, G. Feinman, A. Korotayev, N. Kradin, J. Levine, J. Reddish, E. Cioni, R. Wacziarg, G. Mendel-Gleason, and M. Benam, Disentangling the evolutionary drivers of social complexity: A comprehensive test of hypotheses, *Science Advances* **8**, eabn3517 (2022).
 [18] Roman, S. and Bertolotti, F., Global history, the emergence of chaos and inducing sustainability in networks of socio-ecological systems, *PLOS ONE* **18**, e0293391 (2023).
 [19] A. Piras and F. Bertolotti, How risk preferences shape city-state success: An agent-based model of resource management, in *WOA 2024: 25th Workshop “From Objects to Agents”* (Forte di Bard, Aosta, Italy, 2024).
 [20] F. Diaz-Diaz, P. Bartesaghi, and E. Estrada, Mathematical modeling of local balance in signed networks and its appli-

- cations to global international analysis, *Journal of Applied Mathematics and Computing* **70**, 6195 (2024).
- [21] L. Hébert-Dufresne, N. W. Landry, J. Lovato, J. St-Onge, J.-G. Young, M.-É. Couture-Ménard, S. Bernatchez, C. Choquette, and A. A. Cohen, Governance as a complex, networked, democratic, satisfiability problem, *npj Complexity* **2**, 14 (2025).
- [22] D. M. Abrams and S. H. Strogatz, Modelling the dynamics of language death, *Nature* **424**, 900 (2003).
- [23] P. Rosillo-Rodes, M. San Miguel, and D. Sánchez, Modeling language ideologies for the dynamics of languages in contact, *Chaos* **33**, 113117 (2023).
- [24] R. Alexander Bentley and Matthew W. Hahn and Stephen J. Shennan, Random drift and culture change, *Proceedings of the Royal Society of London. Series B: Biological Sciences* **271**, 1443 (2004).
- [25] M. G. Newberry and J. B. Plotkin, Measuring frequency-dependent selection in culture, *Nature Human Behaviour* **6**, 1048 (2022).
- [26] M. Baratto, I. Casanovas, I. Decostanzi, H. M. Borges, S. M. Alcalá, I. Stanzani, A. Antonioni, I. Iacopini, M. R. Fiorentin, and E. Valdano, The relationship between episcopal genealogy and ideology in the roman catholic church, arXiv preprint arXiv:2506.22108 (2025).
- [27] A. Antonioni, M. R. Fiorentin, and E. Valdano, Complex totopapa: predicting the successor to pope Francis, arXiv preprint arXiv:2505.01553 (2025).
- [28] Egor Lappo and Noah A. Rosenberg, Quo nomine vis vocari? A random-copying model explains the temporal sequence of papal names (2026), arXiv preprint arXiv:2605.07028, arXiv:2605.07028 [q-bio.PE].
- [29] N. L. Johnson and S. Kotz, *Urn Models and Their Application: An Approach to Modern Discrete Probability Theory* (Wiley, 1977) p. 402.
- [30] R. Pemantle, A survey of random processes with reinforcement, *Probability Surveys* **4**, 1 (2007).
- [31] H. Mahmoud, *Pólya urn models* (Chapman and Hall/CRC, 2008).
- [32] J. Bernoulli, *Ars conjectandi, opus posthumum: accedit tractatus de seriebus infinitis, et epistola Gallice scripta de ludo pilæ reticularis* (Impensis Thurnisiorum Fratrum, 1713).
- [33] K. Pearson, X. contributions to the mathematical theory of evolution.—ii. skew variation in homogeneous material, *Philosophical Transactions of the Royal Society of London.(A.)* , 343 (1895).
- [34] W. Feller, *An introduction to probability theory and its applications*, Vol. 2 (John Wiley & Sons, 1991).
- [35] F. Eggenberger and G. Pólya, Über die statistik verketeter vorgänge, *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik* **3**, 279 (1923).
- [36] B. Friedman, A simple urn model, *Communications on Pure and Applied Mathematics* **2**, 59 (1949).
- [37] D. Blackwell and J. B. MacQueen, Ferguson distributions via Pólya urn schemes, *The Annals of Statistics* **1**, 353 (1973).
- [38] B. M. Hill, D. Lane, and W. Sudderth, A strong law for some generalized urn processes, *The Annals of Probability* , 214 (1980).
- [39] P. Muliere, A. M. Paganoni, and P. Secchi, A randomly reinforced urn, *Journal of Statistical Planning and Inference* **136**, 1853 (2006).
- [40] S. Fortini, S. Petrone, and H. Sariev, Predictive constructions based on measure-valued Pólya urn processes, *Mathematics* **9**, 2845 (2021).
- [41] H. Sariev, S. Fortini, and S. Petrone, Infinite-color randomly reinforced urns with dominant colors, *Bernoulli* **29**, 132 (2023).
- [42] F. M. Hoppe, Pólya-like urns and the Ewens' sampling formula, *Journal of Mathematical Biology* **20**, 91 (1984).
- [43] D. J. Aldous, Exchangeability and related topics, in *École d'Été de Probabilités de Saint-Flour XIII—1983* (Springer, 2006) pp. 1–198.
- [44] F. Tria, V. Loreto, V. D. P. Servedio, and S. H. Strogatz, The dynamics of correlated novelties, *Scientific Reports* **4**, 5890 (2014).
- [45] R. K. Merton, The Matthew effect in science: The reward and communication systems of science are considered, *Science* **159**, 56 (1968).
- [46] G. U. Yule, A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, *Philosophical transactions of the Royal Society of London - B* **213**, 21 (1925).
- [47] G. K. Zipf, *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology* (Addison-Wesley Press, Cambridge, MA, 1949).
- [48] D. J. de Solla Price, A general theory of bibliometric and other cumulative advantage processes, *Journal of the American Society for Information Science* **27**, 292 (1976).
- [49] S. Redner, How popular is your paper? an empirical study of the citation distribution, *The European Physical Journal B-Condensed Matter and Complex Systems* **4**, 131 (1998).
- [50] X. Gabaix, Zipf's law and the growth of cities, *American Economic Review* **89**, 129 (1999).
- [51] M. Batty, The size, scale, and shape of cities, *Science* **319**, 769 (2008).
- [52] Tim Lewens, Stochasticity in cultural evolution: a revolution yet to happen, *History and Philosophy of the Life Sciences* **40**, 9 (2018).
- [53] H. A. Simon, On a class of skew distribution functions, *Biometrika* **42**, 425 (1955).
- [54] D. Zanette and M. Montemurro, Dynamics of Text Generation with Realistic Zipf's Distribution, *Journal of Quantitative Linguistics* **12**, 29 (2005).
- [55] P. Rosillo-Rodes, J. W. Zimmerman, L. Hébert-Dufresne and P. S. Dodds, Simon's model does not produce Zipf's law: The fundamental rich-get-richer mechanism for any power-law size ranking (2026), arXiv preprint arXiv:2604.13184, arXiv:2604.13184 [physics.soc-ph].
- [56] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, *Science* **286**, 509 (1999).
- [57] J. Gomez-Gardenes and Y. Moreno, Local versus global knowledge in the barabási-albert scale-free network model, *Physical Review E* **69**, 037103 (2004).
- [58] Wikipedia contributors, List of pharaohs — Wikipedia, the free encyclopedia (2026).
- [59] Wikipedia contributors, List of Roman emperors — Wikipedia, the free encyclopedia (2026).
- [60] Wikipedia contributors, List of Russian monarchs — Wikipedia, the free encyclopedia (2026).
- [61] OrthodoxWiki contributors, List of Patriarchs of Constantinople, OrthodoxWiki (2026).
- [62] Wikipedia contributors, List of heads of state of Spain — Wikipedia, the free encyclopedia (2026).
- [63] Wikipedia contributors, List of monarchs of Denmark — Wikipedia, the free encyclopedia (2026).

- [64] The Editors of Encyclopaedia Britannica, Kings and Queens of Britain, Encyclopaedia Britannica (2024).
- [65] G. Bianconi and A.-L. Barabási, Bose-Einstein condensation in complex networks, *Physical Review Letters* **86**, 5632 (2001).
- [66] L. Hébert-Dufresne, J. Lovato, G. Burgio, J. P. Gleeson, S. Redner, and P. L. Krapivsky, Self-reinforcing cascades: A spreading model for beliefs or products of varying intensity or quality, *Physical Review Letters* **135**, 087401 (2025).
- [67] O. Kolodny, N. Creanza and M. W. Feldman, Game-Changing Innovations: How Culture Can Change the Parameters of Its Own Evolution and Induce Abrupt Cultural Shifts, *PLOS Computational Biology* **12**, e1005302 (2016).
- [68] D. Centola, J. Becker, D. Brackbill and A. Baronchelli, Experimental evidence for tipping points in social convention, *Science* **360**, 1116 (2018).
- [69] Shakshi Singhal and Adarsh Anand and Ompal Singh, Studying dynamic market size-based adoption modeling & product diffusion under stochastic environment, *Technological Forecasting & Social Change* **161**, 120285 (2020).
- [70] J. Raaijmakers, The social meaning of names, in *Writing Names in Medieval Sacred Spaces*, Utrecht Studies in Medieval Literacy, pp. 299–307.

METHODS

Name standardization

We standardize the name data by isolating the primary component of authority in each dynastic tradition. In the case of the Roman emperors, we isolate the *cognomen*. In the case of monarchs, we restrict the analysis to the primary regnal proper name (e.g., King George VI is considered as George). We limit the analysis of papal names to the chosen proper name (e.g., Jorge Bergoglio is Francis). In the case of the pharaohs, we discard the first four titular names, isolating the *nomen*, as it is the case of the Ottoman sultans, where we analyze the core regnal name, omitting all introductory and concluding titular formulas.

Prestige-driven innovation-reinforcement model

We model the temporal evolution of papal names as a stochastic process combining innovation and prestige-driven reinforcement. At each election time t , the set of names that have already appeared is denoted by $A(t)$, with $M(t) = |A(t)|$. Each active name $i \in A(t)$ is characterized by its abundance $n_i(t)$, defined as the number of times the name has been used, and by its accumulated prestige $x_i(t)$. The total prestige in the system is

$$X(t) = \sum_{j \in A(t)} x_j(t). \quad (3)$$

At election t , a new name is introduced with probability $\rho(t)$. With complementary probability $1 - \rho(t)$, an already active name is reused. Conditional on reuse, name i is selected with probability proportional to its accumulated prestige. Therefore, the unconditional probability that an already active name $i \in A(t)$ is chosen is

$$p_i(t) = [1 - \rho(t)] \frac{x_i(t)}{X(t)}. \quad (4)$$

The probability that the election introduces a new name is

$$p_{\text{new}}(t) = \rho(t). \quad (5)$$

Whenever a name is selected, its prestige increases by a random increment K , independently drawn from a prestige-increment distribution $P(K)$, with mean $\mu = \mathbb{E}[K]$ and variance $\sigma^2 = \text{Var}(K)$. If an already active name i is reused, then

$$n_i(t+1) = n_i(t) + 1, \quad x_i(t+1) = x_i(t) + K(t). \quad (6)$$

If innovation occurs, a new name $m \notin A(t)$ enters the system with

$$n_m(t+1) = 1, \quad x_m(t+1) = K(t). \quad (7)$$

Thus, innovation does not create empty types: each new name enters with one observed use and an initial prestige sampled from the same distribution that governs later prestige increments.

The total number of name elections is

$$N(t) = \sum_{i \in A(t)} n_i(t), \quad (8)$$

and grows as $N(t) = N_0 + t$. The expected number of distinct names follows

$$M(t) = M_0 + \int_0^t \rho(s) ds. \quad (9)$$

Prestige distribution

In the RIP model, each reign contributes a positive prestige increment K to the name under which it occurs. We model these increments as independent random variables drawn from a truncated power-law distribution,

$$P(K) = C_{\alpha, \kappa} K^{-\alpha}, \quad 1 \leq K \leq \kappa, \quad (10)$$

where $\alpha > 0$ is the power-law exponent, $\kappa \geq 1$ is the upper cutoff, and $C_{\alpha, \kappa}$ is the normalization constant. The exponent α controls the heaviness of the tail, whereas κ fixes the largest possible prestige increment. Thus, smaller values of α or larger values of κ both increase the potential role of rare but highly prestigious reigns.

We can connect the parameters of the truncated power law distribution with the mean and variance as follows. The mean prestige increment is

$$\mu(\alpha, \kappa) = \mathbb{E}[K] = \frac{1 - \alpha \kappa^{2-\alpha} - 1}{2 - \alpha \kappa^{1-\alpha} - 1}, \quad (11)$$

and the variance of prestige increments

$$\sigma^2(\alpha, \kappa) = \frac{1 - \alpha \kappa^{3-\alpha} - 1}{3 - \alpha \kappa^{1-\alpha} - 1} - \left[\frac{1 - \alpha \kappa^{2-\alpha} - 1}{2 - \alpha \kappa^{1-\alpha} - 1} \right]^2. \quad (12)$$

Note that in the special cases $\alpha = 1, 2, 3$ the corresponding expressions can be obtained as logarithmic limits.

Fitting the prestige distribution

Because each naming tradition requires different magnitudes of rare events, the prestige distribution must be tailored to each dynastic time series. To achieve this, we evaluate an (α, κ) parameter grid across 100 simulations, identifying the optimal (α, κ) point that minimizes the RMSE between the average simulated variance and the empirical variance. For the innovation rate, we compute a smoothed $\rho(t)$ using a centered rolling window of length 3. The results of this optimization process are presented in Table 1.

TABLE I. Optimal fit parameters for $P(K)$ in Eq. (10) obtained through RMSE minimization.

Tradition	κ	α
Spanish monarchs	105	4.98
Danish monarchs	374	3.90
English monarchs	70	1.82
Russian tsars	1	-
Roman emperors	585	2.71
Pharaohs	244	1.31
Holy Roman emperors	1	-
Constantinople patriarchs	330	4.22
Ottoman sultans	1	-
Popes	209	3.07

Quantifying the empirical variance

To quantify the heterogeneity of name abundances, we consider the empirical variance over the active repertoire,

$$\bar{V}_n(t) = \frac{\bar{Q}(t)}{M(t)} - \left[\frac{N(t)}{M(t)} \right]^2, \quad (13)$$

where we have introduced the second abundance moment

$$\bar{Q}(t) = \sum_{i \in A(t)} n_i^2(t). \quad (14)$$

We approximate the expected evolution of this quantity through a continuous-time second-moment closure. We now aim to compute the theoretical variance

$$\bar{V}_n(t) = \frac{\bar{Q}(t)}{M(t)} - \left[\frac{N(t)}{M(t)} \right]^2. \quad (15)$$

Besides $\bar{Q}(t)$, we define the abundance–prestige mixed moment

$$\bar{Y}(t) = \sum_{i \in A(t)} n_i(t) x_i(t), \quad (16)$$

and the prestige second moment

$$\bar{W}(t) = \sum_{i \in A(t)} x_i^2(t). \quad (17)$$

As derived in the Supplementary Note 2, for an arbitrary innovation function $\rho(t)$, the formal solution in Eq. (13) can be written explicitly by computing Eq. (9) and the second moments as follows:

$$\bar{W}(t) = e^{2G_1(t)} \left[W_0 + (\mu^2 + \sigma^2) \int_0^t e^{-2G_1(s)} ds \right], \quad (18)$$

$$\bar{Y}(t) = e^{G_1(t)} \quad (19)$$

$$\left[Y_0 + \int_0^t e^{-G_1(s)} \left(\mu + [1 - \rho(s)] \frac{\bar{W}(s)}{\mu N(s)} \right) ds \right],$$

$$\bar{Q}(t) = Q_0 + t + \frac{2}{\mu} \int_0^t [1 - \rho(s)] \frac{\bar{Y}(s)}{N(s)} ds, \quad (20)$$

being $G_1(t) = \int_0^t \frac{1 - \rho(u)}{N(u)} du$.

RIP vs. RGR in the limit scenario of no innovation

To get an analytical insight into the excess variance induced by the heterogeneous prestige distribution, let us consider the closed-repertoire limit, $\rho(t) = 0$, for which the number of active names remains fixed, $M(t) = M_0$, while the total number of choices grows as $N(t) = N_0 + t$.

For the prestige-driven model, the empirical abundance variance scales as (see the full derivation in Supplementary Note 2)

$$\bar{V}_{n,\text{RIP}}(t) \sim \left[\frac{1}{M_0} \left(\frac{W_0}{\mu^2 N_0^2} + \frac{1}{N_0} \left[1 + \frac{\sigma^2}{\mu^2} \right] \right) - \frac{1}{M_0^2} \right] (N_0 + t)^2, \quad (21)$$

where $W_0 = \sum_{i=1}^{M_0} x_i^2(0)$, $\mu = \mathbb{E}[K]$, and $\sigma^2 = \text{Var}(K)$. For the pure preferential-attachment model, in which reused names are selected proportionally to their current abundance, the corresponding leading-order expression is

$$\bar{V}_{n,\text{RGR}}(t) \sim \left[\frac{1}{M_0} \left(\frac{Q_0}{N_0^2} + \frac{1}{N_0} \right) - \frac{1}{M_0^2} \right] (N_0 + t)^2, \quad (22)$$

with $Q_0 = \sum_{i=1}^{M_0} n_i^2(0)$. Thus, in the absence of innovation, both reinforcement mechanisms generate quadratic growth of the abundance variance, but heterogeneous prestige increments increase the prefactor through the dimensionless ratio σ^2/μ^2 .

If both processes are initialized consistently, $x_i(0) = \mu n_i(0)$, so that $W_0 = \mu^2 Q_0$, Eqs. (21)–(22) yield Eq. (2).

SUPPLEMENTARY NOTE 1: ONOMASTICS

The role of prestige in each dataset

Name selection varies significantly across ruling traditions. In several monarchies, rulers traditionally reigned under their primary baptismal name. In other royal systems monarchs selected from a composite of multiple baptismal names, as exemplified by King George VI of England, whose full baptismal name was Albert Frederick Arthur George. In contrast, ancient Egyptian pharaohs systematically adopted regnal names upon ascension. In the case of popes, name change did not establish itself as an institutional trend until approximately the year 900 with the election of John X (with only two prior pontiffs changing their names).

Due to these operational differences, the symbolic prestige of names acts differently in each ruling system. We operationalize prestige as the probability of a name being selected based on the historical legacy and achievements of its prior holders. Within dynastic, hereditary frameworks, the assignment of a given name to an heir reflects a naming tradition in a given family or lineage. However, when an emperor or pope consciously adopts a specific name upon installation, it serves as an exercise in political marketing, drawing upon the accumulated prestige of that lineage to project institutional continuity, stability, and power.

Papal onomastics

The bishops of Rome have occupied a special place in the naming culture over the last two millennia. In the Middle Ages, a name was never a neutral label. Naming was a social act that inscribed an individual into kinship networks, local traditions, and patterns of patronage, and it was often a religious act that invoked the protection or imitation of a saint [70]. By the seventh century in Western Europe, local customs largely determined naming patterns, yet a gradual convergence toward a more universal repertoire can be detected, even as communities clung to their own onomastic habits. At the same time, demographic growth and new administrative demands drove a shift from single names to hereditary surnames, while the stock of given names shrank in what has been described as an onomastic deflation [70]. Before Francis, the papal names had been repeated for more than a millennium.

The papacy's own naming history divides neatly into two long stages. Across roughly the first millennium, almost every pope retained the name he bore at the moment of election [3]. Romans, Greeks, Africans, Jews, and others ascended the Holy See and brought with them names rooted in different linguistic and cultural worlds. The early papal onomasticon thus mirrors the diversity and relative fluidity of late antique and early medieval Christian society.

The transition from this first stage of stability to a second one governed by papal self-naming was triggered by an intricate political history. Between the mid-ninth and mid-eleventh centuries, Rome and its church were buffeted by internal factionalism, intervention from Frankish and German emperors, competition with Constantinople, and threats from Muslim powers [4]. In this context, an "anomalous" interlocking series of Popes John and Benedict emerged between the ninth and eleventh centuries, during which the two names reached exceptionally high rates of reuse [4]. Their recurrence tracked the influence of powerful Roman lineages such as the Crescentii and the counts of Tusculum and showed how strongly emperors were steering papal affairs at different moments [4, 5]. Several of these Johns and Benedicts were murdered, deposed, accused of simony or immorality, or installed and removed by secular patrons.

It is in the wake of this crisis that papal onomastic innovation abruptly disappears, starting with John X in 914. Gradually, papal self-naming becomes established. What had been an occasional expedient (e.g., Mercurius taking the name John II in 533 to avoid a pagan theonym) turned into a custom endowed with theological and political meaning. The reforming German popes installed by Henry III in the mid-eleventh century chose new names that linked them to ancient martyrs and revered bishops of Rome even as they signaled a break with the scandalous recent past [4]. By changing their names, popes could also imitate the biblical pattern of Peter and Paul, dramatize the spiritual transformation implied in their election, and underline the separation of their office from secular family ties and worldly careers, in harmony with the broader reform agenda that sought clerical celibacy and autonomy [3].

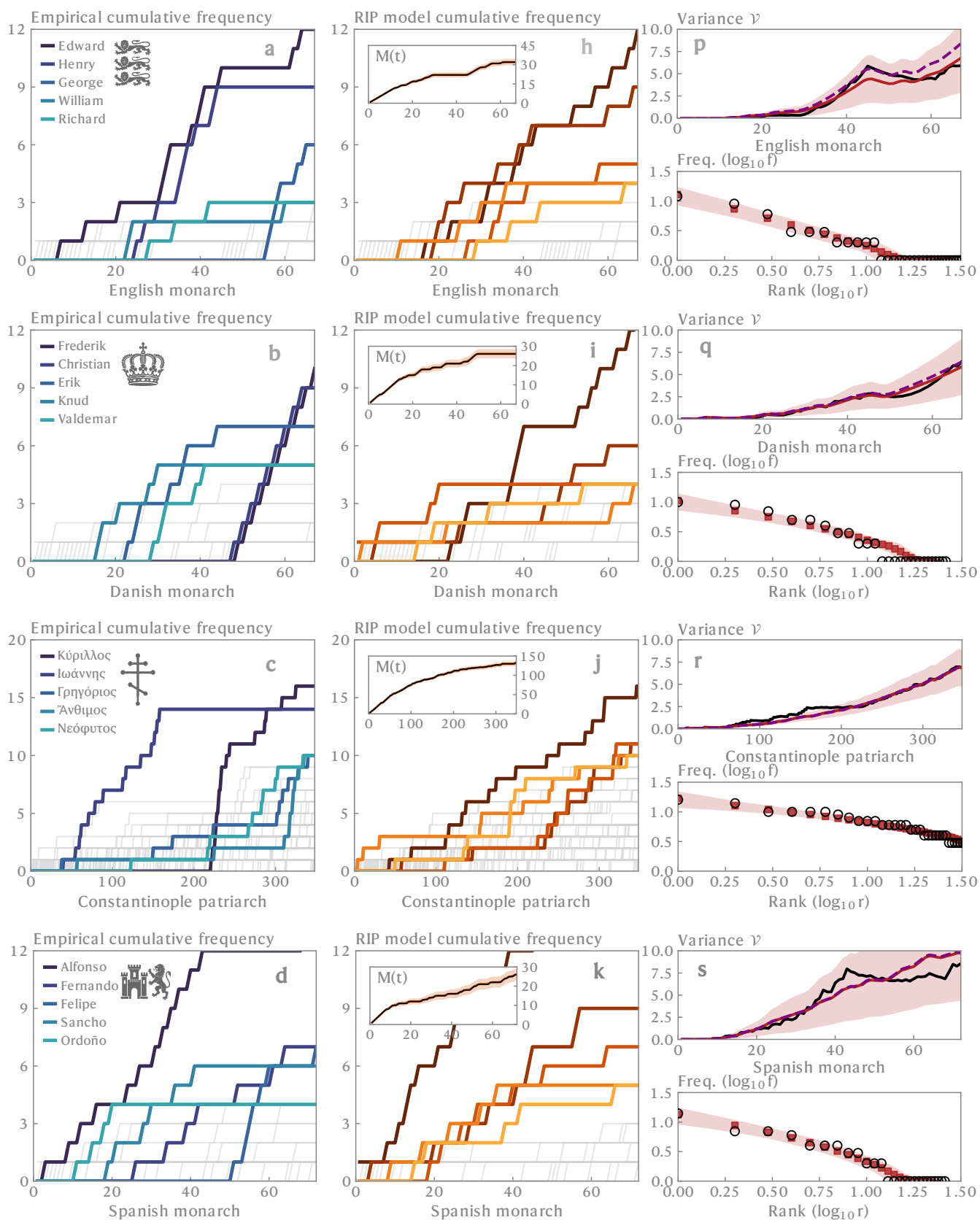
Once introduced, the new practice quickly crystallized into an onomastic tradition with its own internal logic. Name choice allowed a newly elected pope to weave himself into a chain of predecessors and saints while projecting a future program for his pontificate [4]; now, popes are telling stories when choosing a name. High medieval popes

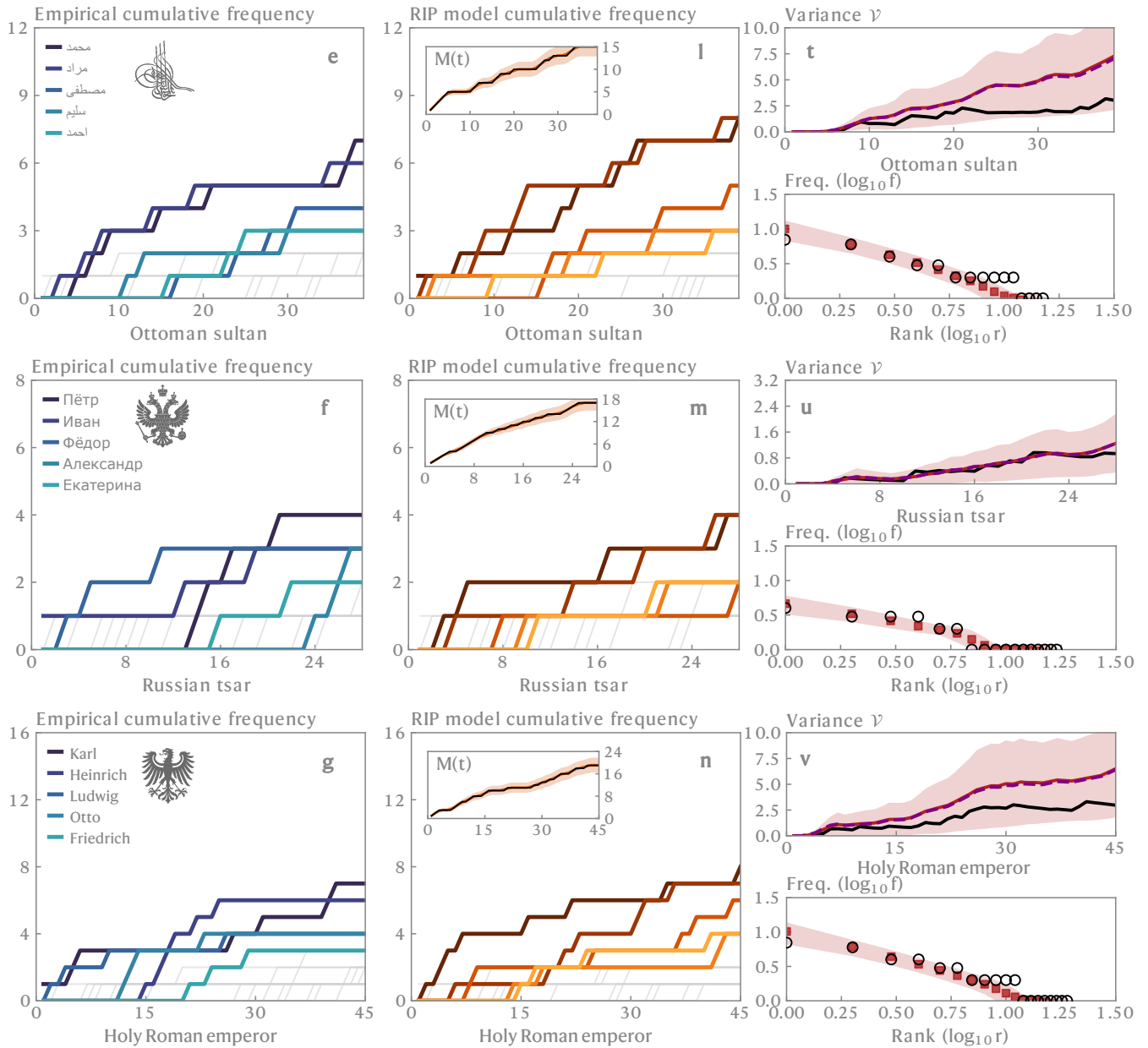
overwhelmingly selected names of early bishops of Rome. In later centuries, the pattern shifted from emulating the distant patristic past to honoring more recent predecessors, family connections, religious orders, or the symbolic content of a name. In all these cases, the underlying criterion remained the same: names carried accumulated prestige, and choosing one meant tapping into the symbolic capital built by earlier pontiffs or revered figures. For example, Pius II famously drew on Virgil's epithet "pius Aeneas" when selecting his own name; John XXIII resurrected a neglected, even tarnished, name for reasons ranging from family homage to local patronage and modest expectations of a short pontificate; and John Paul I crafted an unprecedented double name to signal continuity with both John XXIII's pastoral warmth and Paul VI's administrative skill [3, 4]. More recent popes, such as Benedict XVI, have explicitly explained their choices as bids for peace, reform, or alignment with particular spiritual traditions, making the act of naming itself a first narrative gesture of their pontificates [4].

The long pattern of reusing established papal names ended with the election of Pope Francis in 2013. He was the first Jesuit pope, the first from Latin America, and the first pope named Francis. His choice introduced a new phase in papal naming, one that has continued with the recent election of Pope Leo XIV.

Onomastics of powerful dynasties

In Supplementary Fig. 1 we expand the analysis of Fig. 1 to the remaining dynasties.





Supplementary Fig. 1. **a-g** Number of Danish monarchs, English monarchs, patriarchs of Constantinople, Spanish monarchs, SRGE Holy Roman emperors, Ottoman sultans and Russian tsars, respectively, bearing each name across chronologically ordered reigns. **h-n** Output of a single run of the RIP model. The insets show the evolution of the number of distinct names ($M(t)$ in bottom rows) over time. **p-v** Top: Variance of name frequencies in the data compared with 1000 simulations of the RIP model and with the analytical expression from Eqs. (15)-(20). Bottom: rank-frequency distribution of names. The plots show the frequency f_i of each name i as a function of its rank r_i , comparing the empirical data with the mean outcome of 1000 simulations of the RIP model $\pm 1\sigma$.

SUPPLEMENTARY NOTE 2: BIC AND AIC FOR ASSESSING THE GOODNESS OF THE FIT

The Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are statistical tools used for model selection that balance how well a model fits the data with its complexity to prevent overfitting. AIC focuses primarily on minimizing information loss and optimizing predictive performance on new datasets, which can sometimes lead it to choose overly complex models in large samples. Conversely, BIC introduces a stricter penalty that depends on the total sample size, favoring more parsimonious, simpler models and demonstrating statistical consistency by asymptotically selecting the true data-generating model if it is among the choices.

When we only have RMSE data available instead of a direct log-likelihood value, AIC and BIC can still be computed under the assumption that the model's residuals are independent and identically normally distributed. In this case, the maximized log-likelihood simplifies into a direct function of the sample size and the log-RMSE. We have then

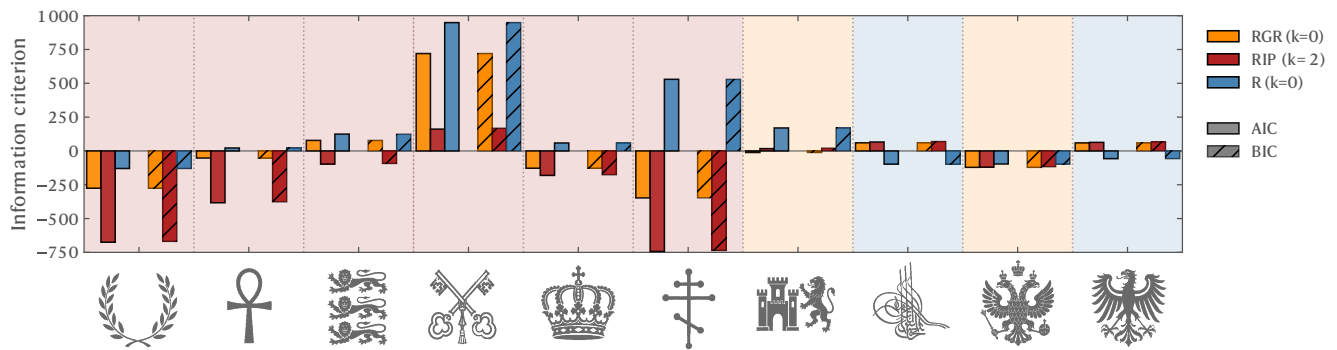
$$\text{AIC} = 2n \ln(\text{RMSE}) + 2k$$

and

$$\text{BIC} = 2n \ln(\text{RMSE}) + k \ln(n),$$

where n represents the number of observations and k denotes the total number of parameters. Therefore, the lowest the BIC or AIC, the better.

In Supplementary Fig. 2 we show the performance of RIP, rich-get-richer and random models for each dataset.



Supplementary Fig. 2. **Goodness of the fit for prestige, rich-get-richer and random models** in terms of two additional metrics, the Bayesian and the Akaike Information Criteria. The horizontal labels correspond to the symbols shown in Fig. 2.

SUPPLEMENTARY NOTE 3: THE REINFORCEMENT AND INNOVATION THROUGH PRESTIGE (RIP) MODEL

Reinforcement and Innovation through Prestige (RIP) model

We consider a stochastic model for the temporal evolution of names. At each name election, either a new name is introduced or one of the names that have already appeared in the historical record is reused.

Let $A(t)$ denote the set of active names before election t , and let $M(t) = |A(t)|$ be the number of distinct names observed up to that point. Each active name $i \in A(t)$ is characterized by two variables: its abundance $n_i(t)$, defined as the number of times the name has been used, and its accumulated prestige $x_i(t)$. The total prestige in the system is

$$X(t) = \sum_{j \in A(t)} x_j(t). \quad (\text{S.1})$$

At election t , innovation occurs with probability $\rho(t) \in [0, 1]$. If innovation occurs, a new name $m \notin A(t)$ is introduced. If innovation does not occur, one of the already active names is selected with probability proportional to its accumulated prestige. Therefore, for an already active name $i \in A(t)$, the unconditional probability of being chosen at election t is

$$p_i(t) = [1 - \rho(t)] \frac{x_i(t)}{X(t)}. \quad (\text{S.2})$$

The probability that election t introduces a new name is

$$p_{\text{new}}(t) = \rho(t). \quad (\text{S.3})$$

We define the innovation indicator J_t which takes value $J_t = 1$ if election introduces a new name and $J_t = 0$ otherwise, with $\mathbb{E}[J_t] = \rho(t)$. We also define the reinforcement indicator $Z_t = 1 - J_t$, so that $Z_t = 1$ when election t reuses an already existing name, and $Z_t = 0$ when election t introduces a new name.

Whenever a name is selected, its prestige increases by a random increment K , independently drawn from a prestige-increment distribution $P(K)$. We denote its mean, variance and second moment by $\mu = \mathbb{E}[K] > 0$, $\sigma^2 = \text{Var}(K)$ and $m_2 = \mathbb{E}[K^2] = \mu^2 + \sigma^2$.

Therefore, if an already active name $i \in A(t)$ is selected at election t , then its abundance and prestige evolve as

$$n_i(t+1) = n_i(t) + 1, \quad (\text{S.4})$$

$$x_i(t+1) = x_i(t) + K(t), \quad (\text{S.5})$$

where $K(t)$ is drawn from $P(K)$. All other active names $j \neq i$ remain unchanged during that election.

If innovation occurs, a new name $m \notin A(t)$ enters the system. Its abundance and prestige are initialized as

$$n_m(t+1) = 1, \quad (\text{S.6})$$

$$x_m(t+1) = K(t), \quad (\text{S.7})$$

where, again, $K(t)$ is drawn from $P(K)$. Thus, innovation does not introduce an empty type: a new name enters the system with one observed use and with an initial prestige drawn from the same prestige-increment distribution that governs reinforcement events. After innovation, the active set is updated as $A(t+1) = A(t) \cup \{m\}$.

The model therefore combines two mechanisms. The function $\rho(t)$ controls the opening of the name repertoire through the introduction of new names. The prestige variables $x_i(t)$ control the reinforcement dynamics among already active names. Large values of $\rho(t)$ favor innovation and repertoire expansion, whereas small values of $\rho(t)$ favor the reuse of established names through prestige-driven reinforcement.

Growth of the repertoire and variance of name abundances

We now derive the expected growth of the number of distinct names and the time evolution of the variance of name abundances in the RIP model. Let

$$N(t) = \sum_{i \in A(t)} n_i(t) \quad (\text{S.8})$$

be the total number of name elections recorded up to time t . Since exactly one name is assigned at each election,

$$N(t) = N_0 + t, \quad (\text{S.9})$$

where N_0 is the number of elections already present in the initial condition. The number of active names evolves according to

$$M(t) = M_0 + \sum_{s=0}^{t-1} J_s. \quad (\text{S.10})$$

Since $\mathbb{E}[J_s] = \rho(s)$, we obtain $\mathbb{E}[M(t)] = M_0 + \sum_{s=0}^{t-1} \rho(s)$. In a continuous-time approximation, this becomes

$$M(t) = M_0 + \int_0^t \rho(s) ds. \quad (\text{S.11})$$

We next consider the variance of the abundance distribution over the set of active names. We define the empirical variance

$$\mathcal{V}(t) = \frac{1}{M(t)} \sum_{i \in A(t)} \left[n_i(t) - \frac{N(t)}{M(t)} \right]^2. \quad (\text{S.12})$$

Introducing the second moment

$$Q(t) = \sum_{i \in A(t)} n_i^2(t), \quad (\text{S.13})$$

the variance can be written as

$$\mathcal{V}(t) = \frac{\overline{Q}(t)}{M(t)} - \left[\frac{N(t)}{M(t)} \right]^2. \quad (\text{S.14})$$

Therefore, the problem reduces to obtaining the evolution of $Q(t)$.

Determining the evolution of $Q(t)$

If election t is an innovation event, a new name enters with abundance $n_m(t+1) = 1$. Hence,

$$\Delta Q(t)|_{\text{innovation}} = 1. \quad (\text{S.15})$$

If election t is a reinforcement event and an already active name i is selected, then

$$n_i(t+1) = n_i(t) + 1, \quad (\text{S.16})$$

and therefore

$$\Delta Q(t)|_{\text{reinforcement}} = [n_i(t) + 1]^2 - n_i^2(t) = 2n_i(t) + 1. \quad (\text{S.17})$$

Using the prestige-driven selection probability in Eq. (S.2), we obtain the exact conditional evolution given all the previous information about the state of the system F_t ,

$$\mathbb{E}[\Delta Q(t) | \mathcal{F}_t] = \rho(t) + [1 - \rho(t)] \sum_{i \in A(t)} \frac{x_i(t)}{X(t)} [2n_i(t) + 1]. \quad (\text{S.18})$$

We can introduce the abundance-prestige second moment

$$Y(t) = \sum_{i \in A(t)} n_i(t)x_i(t), \quad (\text{S.19})$$

and since $\sum_{i \in A(t)} x_i(t)/X(t) = 1$, Eq. (S.18) simplifies to

$$\mathbb{E}[\Delta Q(t) | \mathcal{F}_t] = 1 + 2[1 - \rho(t)] \frac{Y(t)}{X(t)}. \quad (\text{S.20})$$

Equation (S.20) is exact, but it is not closed because the second moment of abundances depends on the evolution of the prestige and the correlation between abundance and prestige.

To close the calculation, we derive the evolution of $X(t)$ and $Y(t)$. The total prestige increases by one prestige increment at every election, either because a new name enters with initial prestige K , or because an already active name receives an additional prestige increment. Thus,

$$\mathbb{E}[\Delta X(t) | \mathcal{F}_t] = \mu, \quad (\text{S.21})$$

and, at the level of expectations, $\bar{X}(t) \simeq X_0 + \mu t$. If the initial prestige is proportional to the initial number of elections, this can be written as

$$\bar{X}(t) \simeq \mu N(t). \quad (\text{S.22})$$

Then, let's derive $Y(t)$. If innovation occurs, a new name enters with $n_m(t+1) = 1$ and $x_m(t+1) = K(t)$, and hence

$$\Delta Y(t)|_{\text{innovation}} = K(t). \quad (\text{S.23})$$

If an already active name i is selected, then

$$n_i(t+1) = n_i(t) + 1, \quad x_i(t+1) = x_i(t) + K(t). \quad (\text{S.24})$$

Therefore, the contribution of name i to $Y(t)$ changes as

$$\Delta Y(t)|_{\text{reinforcement}} = [n_i(t) + 1][x_i(t) + K(t)] - n_i(t)x_i(t) \quad (\text{S.25})$$

$$= x_i(t) + K(t)n_i(t) + K(t). \quad (\text{S.26})$$

Averaging over the prestige increment and over the selected name gives

$$\mathbb{E}[\Delta Y(t) | \mathcal{F}_t] = \rho(t)\mu + [1 - \rho(t)] \sum_{i \in A(t)} \frac{x_i(t)}{X(t)} [x_i(t) + \mu n_i(t) + \mu] \quad (\text{S.27})$$

$$= \mu + [1 - \rho(t)] \left[\frac{Y(t)}{X(t)} + \mu \frac{Y(t)}{X(t)} \right], \quad (\text{S.28})$$

where we have introduced the second moment of prestige,

$$W(t) = \sum_{i \in A(t)} x_i^2(t). \quad (\text{S.29})$$

Again, Eq. (S.28) is not closed, and depends on the evolution of $W(t)$. If innovation occurs, a new name with prestige K is introduced, and therefore

$$\Delta W(t) = K^2. \quad (\text{S.30})$$

If an already active name i is selected, its prestige changes as

$$x_i(t+1) = x_i(t) + K(t), \quad (\text{S.31})$$

so that

$$\Delta W(t) = [x_i(t) + K(t)]^2 - x_i^2(t) = 2K(t)x_i(t) + K^2(t). \quad (\text{S.32})$$

Conditioned on a reinforcement event, the probability that the selected active name is i is $x_i(t)/X(t)$. Equivalently, the unconditional probability of selecting name i is $[1 - \rho(t)]x_i(t)/X(t)$, consistently with Eq. (S.2). Averaging over the prestige increment and over the prestige-driven choice of the selected name gives

$$\mathbb{E}[\Delta W(t) | \mathcal{F}_t] = \rho(t)m_2 + [1 - \rho(t)] \sum_{i \in A(t)} \frac{x_i(t)}{X(t)} [2\mu x_i(t) + m_2] \quad (\text{S.33})$$

$$= m_2 + 2\mu[1 - \rho(t)] \frac{W(t)}{X(t)}, \quad (\text{S.34})$$

where we have used again that the prestige-weighted probabilities sum to one. Note that the variance of prestige increments does not enter directly in the update of the abundances $n_i(t)$. Instead, it modifies the dispersion of accumulated prestige through $W(t)$. This affects the abundance–prestige coupling $Y(t)$, which in turn controls the growth of $Q(t)$ and therefore the variance of name abundances.

In a continuous-time second-moment closure, we replace random quantities by their expected values and recall that $X(t) \simeq \mu N(t)$. We finally obtain the closed system

$$\frac{dM}{dt} = \rho(t), \quad (\text{S.35})$$

$$\frac{d\bar{W}}{dt} = \mu^2 + \sigma^2 + 2(1 - \rho(t)) \frac{\bar{W}(t)}{N(t)}, \quad (\text{S.36})$$

$$\frac{d\bar{Y}}{dt} = \mu + (1 - \rho(t)) \left[\frac{\bar{W}(t)}{\mu N(t)} + \frac{\bar{Y}(t)}{N(t)} \right], \quad (\text{S.37})$$

$$\frac{d\bar{Q}}{dt} = 1 + 2(1 - \rho(t)) \frac{\bar{Y}(t)}{\mu N(t)}. \quad (\text{S.38})$$

The expected empirical variance is then approximated by

$$\bar{V}(t) \simeq \frac{\bar{Q}(t)}{M(t)} - \left[\frac{N(t)}{M(t)} \right]^2. \quad (\text{S.39})$$

This provides the desired general expression for arbitrary innovation function $\rho(t)$. Equations (S.35)–(S.38) can be integrated numerically for any prescribed $\rho(t)$, and Eq. (S.39) gives the corresponding predicted variance of the abundance distribution.

For completeness, the formal solution for arbitrary $\rho(t)$ can also be written explicitly. Let

$$G_1(t) = \int_0^t \frac{1 - \rho(u)}{N(u)} du, \quad G_2(t) = 2G_1(t). \quad (\text{S.40})$$

Then

$$M(t) = M_0 + \int_0^t \rho(s) ds, \quad (\text{S.41})$$

$$\bar{W}(t) = e^{G_2(t)} \left[W_0 + (\mu^2 + \sigma^2) \int_0^t e^{-G_2(s)} ds \right], \quad (\text{S.42})$$

$$\bar{Y}(t) = e^{G_1(t)} \left[Y_0 + \int_0^t e^{-G_1(s)} \left(\mu + (1 - \rho(s)) \frac{\bar{W}(s)}{\mu N(s)} \right) ds \right], \quad (\text{S.43})$$

$$\bar{Q}(t) = Q_0 + t + \frac{2}{\mu} \int_0^t (1 - \rho(s)) \frac{\bar{Y}(s)}{N(s)} ds. \quad (\text{S.44})$$

Substituting these expressions into Eq. (S.39) gives the abundance variance for a general innovation function $\rho(t)$.

Limit scenario of no innovation

In the non-innovation case, $\rho(t) = 0$, the number of names is fixed, and therefore the evolution of the active names reads $M(t) = M_0$. We use Eq. (S.9), recalling that $\rho(t) = 0$. Since $N(t) = N_0 + t$, we have $dN/dt = 1$, and we can equivalently use N as the time variable. Therefore, Eq. (S.36) can be solved explicitly as

$$\bar{W}(t) = N^2(t) \left[\frac{W_0}{N_0^2} + (\mu^2 + \sigma^2) \left(\frac{1}{N_0} - \frac{1}{N(t)} \right) \right] \quad (\text{S.45})$$

$$= B_W N^2(t) - (\mu^2 + \sigma^2) N(t), \quad (\text{S.46})$$

being

$$B_W = \frac{W_0}{N_0^2} + \frac{\mu^2 + \sigma^2}{N_0}. \quad (\text{S.47})$$

Hence, the leading-order behavior is

$$\bar{W}(t) \sim B_W N^2(t), \quad (\text{S.48})$$

and shows explicitly that the heterogeneity of the prestige increments modifies the leading growth of the accumulated-prestige second moment. We now propagate this leading contribution through the moment hierarchy. Again, we can equivalently use N as the time variable. This way, Eq. (S.37) becomes

$$\frac{d\bar{Y}}{dN} - \frac{\bar{Y}}{N} = \mu + \frac{\bar{W}(N)}{\mu N}. \quad (\text{S.49})$$

Using Eq. (S.46) and integrating, we obtain

$$\bar{Y}(t) = N(t) \left\{ \frac{Y_0}{N_0} + \frac{B_W}{\mu} [N(t) - N_0] - \frac{\sigma^2}{\mu} \log \left[\frac{N(t)}{N_0} \right] \right\}. \quad (\text{S.50})$$

The dominant term is thus

$$\bar{Y}(t) \sim \frac{B_W}{\mu} N^2(t). \quad (\text{S.51})$$

We now use the leading contribution in Eq. (S.51) to solve the differential equation for the abundance second moment, Eq. (S.38). After retaining the dominant linear contribution, we obtain

$$\bar{Q}(t) \sim \frac{B_W}{\mu^2} N^2(t). \quad (\text{S.52})$$

Finally, since the repertoire is closed, $M(t) = M_0$, and using Eqs. (S.9)-(S.47)-(S.52) the approximated empirical abundance variance in Eq. (S.39) is

$$\bar{\mathcal{V}}(t) \sim \left[\frac{1}{M_0} \left(\frac{W_0}{\mu^2 N_0^2} + \frac{1}{N_0} \left[1 + \frac{\sigma^2}{\mu^2} \right] \right) - \frac{1}{M_0^2} \right] (N_0 + t)^2, \quad (\text{S.53})$$

which for large t yields a quadratic asymptotic behavior $\bar{\mathcal{V}}(t) \sim t^2$. Importantly, this equation shows explicitly how the prestige-increment variance enters the predicted abundance variance. The quadratic scaling comes from the absence of innovation, whereas the prefactor is increased by the dimensionless prestige heterogeneity σ^2/μ^2 .

Limit scenario of full innovation

If every election introduces a new name, then there are no reinforcement events. If the initial condition consists only of names that have appeared once, then $N_0 = M_0$ and $Q_0 = M_0$. Consequently,

$$N(t) = M_0 + t, \quad M(t) = M_0 + t, \quad Q(t) = M_0 + t, \quad (\text{S.54})$$

and therefore when we substitute in Eq. (S.39) we obtain that

$$\mathcal{V}(t) = 0. \quad (\text{S.55})$$

In this full-innovation limit, the prestige distribution still affects the accumulated prestige variables $x_i(t)$, but it has no effect on the abundance variance because prestige never enters the selection of an already active name.

Limit scenario of constant innovation in the large-times limit

For constant innovation probability, $\rho(t) = \rho$, the not only the number of elections grows asymptotically as $N(t) \sim t$, but also the number of active names. Integrating Eq. (S.11), we find that $M(t) \sim \rho t$.

Since both $N(t)$ and $M(t)$ grow linearly with time, $Q(t)$ must also scale with t in order that the variance in Eq. (S.39) approaches a finite asymptotic value and does not diverge. This implies that $d\bar{Q}/dt$ must approach a constant.

We introduce the rescaled moments

$$U(t) = \frac{\bar{W}(t)}{\mu^2}, \quad V(t) = \frac{\bar{Y}(t)}{\mu}. \quad (\text{S.56})$$

For constant ρ , their equations are

$$\frac{dU}{dt} = 1 + \frac{\sigma^2}{\mu^2} + 2(1 - \rho) \frac{U}{N}, \quad (\text{S.57})$$

$$\frac{dV}{dt} = 1 + (1 - \rho) \left[\frac{U}{N} + \frac{V}{N} \right], \quad (\text{S.58})$$

so that Eq. (S.39) becomes

$$\frac{d\bar{Q}}{dt} = 1 + 2(1 - \rho) \frac{V}{N}. \quad (\text{S.59})$$

Again using $dN/dt = 1$, we look for linear asymptotic scalings $U(N) \sim c_U N$, $V(N) \sim c_V N$, $\bar{Q}(N) \sim c_Q N$. Substituting $U(N) \sim c_U N$ into Eq. (S.57) yields

$$c_U = \frac{1 + \sigma^2/\mu^2}{2\rho - 1}. \quad (\text{S.60})$$

Next, substituting $U(N) \sim c_U N$ and $V(N) \sim c_V N$ into Eq. (S.58), we get

$$c_V = \frac{1}{\rho} \left[1 + (1 - \rho) \frac{1 + \sigma^2/\mu^2}{2\rho - 1} \right], \quad (\text{S.61})$$

and using Eq. (S.59),

$$c_Q = 1 + \frac{2(1 - \rho)}{\rho} \left[1 + (1 - \rho) \frac{1 + \sigma^2/\mu^2}{2\rho - 1} \right]. \quad (\text{S.62})$$

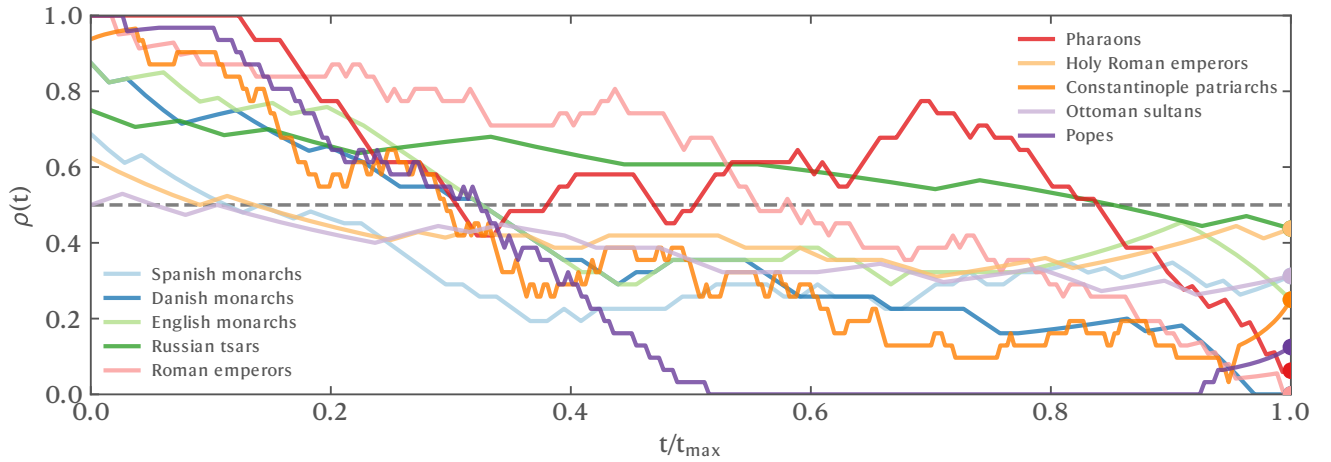
Substituting all the rescaling in Eq. (S.39), we find that the asymptotic abundance variance is

$$\bar{V}(\infty) = \frac{c_Q}{\rho} - \frac{1}{\rho^2} \quad (\text{S.63})$$

$$= \frac{1 - \rho}{\rho^2} \left[1 + \frac{2(1 - \rho)}{2\rho - 1} \left(1 + \frac{\sigma^2}{\mu^2} \right) \right]. \quad (\text{S.64})$$

The denominator $2\rho - 1$ shows that this finite asymptotic expression is well defined only for $\rho > 1/2$. For $\rho \leq \rho_c = 1/2$, reinforcement dominates strongly enough that the second moments no longer grow linearly with $N(t)$; consequently, $\bar{Q}(t)/M(t)$ diverges and the abundance variance does not approach a finite large-time limit.

Supplementary Fig. 3 shows that the empirical innovation rates are not constant across the historical sequences. Instead, after an initial innovation-dominated regime, $\rho(t)$ decreases in all datasets, reflecting the progressive exhaustion or closure of the available name repertoire. Nonetheless, this decay places the long-time dynamics in an effective low-innovation regime (below the aforementioned $\rho_c = 1/2$), where reinforcement is no longer diluted by the arrival of new names. Therefore, the continued growth of the variance trajectories observed in Fig. 1g–i and Supplementary Fig. 1 is the expected outcome of the empirical innovation profiles.



Supplementary Fig. 3. **Empirical innovation profiles across dynastic naming traditions.** Time-dependent innovation rate $\rho(t)$ for the ten historical datasets, plotted as a function of normalized time t/t_{\max} . The innovation rate is estimated from the empirical sequence of name choices using a centered rolling window of length 30 for visual clarity.

Comparison with rich-get-richer and random sampling

We now compare the abundance variance generated by the prestige-driven model with that generated by pure rich-get-richer. In the pure rich-get-richer model, established names are selected with probability

$$p_i^{\text{RGR}}(t) = [1 - \rho(t)] \frac{n_i(t)}{N(t)}, \quad (\text{S.65})$$

which does not depend on the prestige. As a consequence, the second moment of the abundance distribution now obeys the closed equation

$$\frac{d\bar{Q}_{\text{RGR}}}{dt} = 1 + 2(1 - \rho(t)) \frac{\bar{Q}_{\text{RGR}}(t)}{N(t)}, \quad (\text{S.66})$$

Note that pure rich-get-richer is recovered when prestige is exactly proportional to abundance, $x_i(t) = \mu n_i(t)$. In that case, the rescaled moments $U(t) = V(t) = \bar{Q}_{\text{RGR}}(t)$, and the prestige-driven equation reduces to the rich-get-richer equation.

To compare the two models, assume the same initial condition and take $x_i(0) = \mu n_i(0)$, so that $U(0) = V(0) = \bar{Q}_{\text{RGR}}(0)$. We define the differences $\Delta U(t) = U(t) - \bar{Q}_{\text{RGR}}(t)$, $\Delta V(t) = V(t) - \bar{Q}_{\text{RGR}}(t)$, and $\Delta Q(t) = \bar{Q}_{\text{RIP}}(t) - \bar{Q}_{\text{RGR}}(t)$. Subtracting the rich-get-richer equation from the prestige-driven closure yields

$$\frac{d\Delta U}{dt} = \frac{\sigma^2}{\mu^2} + 2(1 - \rho(t)) \frac{\Delta U(t)}{N(t)}, \quad (\text{S.67})$$

$$\frac{d\Delta V}{dt} = (1 - \rho(t)) \frac{\Delta U(t) + \Delta V(t)}{N(t)}, \quad (\text{S.68})$$

$$\frac{d\Delta Q}{dt} = 2(1 - \rho(t)) \frac{\Delta V(t)}{N(t)}. \quad (\text{S.69})$$

Since $1 - \rho(t) \geq 0$, $N(t) > 0$, $\sigma^2/\mu^2 \geq 0$, and the initial differences vanish, these equations imply that $\Delta Q(t) \geq 0$. Therefore, $\bar{Q}_{\text{RIP}}(t) \geq \bar{Q}_{\text{RGR}}(t)$. Since both models have the same $N(t)$ and the same expected number of names $M(t)$, the difference in abundance variance is

$$\bar{V}_{\text{RIP}}(t) - \bar{V}_{\text{RGR}}(t) \simeq \frac{\bar{Q}_{\text{RIP}}(t) - \bar{Q}_{\text{RGR}}(t)}{M(t)}. \quad (\text{S.70})$$

Hence,

$$\bar{V}_{\text{RIP}}(t) \geq \bar{V}_{\text{RGR}}(t). \quad (\text{S.71})$$

The inequality becomes an equality if prestige increments are deterministic, $\sigma^2 = 0$, provided that $x_i(0) = \mu n_i(0)$. In that case, prestige remains proportional to abundance and the prestige-driven model collapses exactly onto pure rich-get-richer.

Finally, we compare with a null model in which name reuse is not preferential. Innovation is kept unchanged: with probability $\rho(t)$, a new name enters with abundance one. However, when no innovation occurs, an already active name is selected uniformly at random. Thus, established names are selected with probability

$$p_i^{\text{R}}(t) = [1 - \rho(t)] \frac{1}{M(t)}, \quad (\text{S.72})$$

which does not depend on the prestige nor in the abundances. Therefore, now the second moment of the abundance distribution obeys the closed equation

$$\frac{d\bar{Q}_{\text{R}}}{dt} = 1 + 2(1 - \rho(t)) \frac{N(t)}{M(t)}. \quad (\text{S.73})$$

This should be compared with pure rich-get-richer in Eq. (S.66). The difference between the two mechanisms is clear. Random reuse is controlled only by the mean abundance per active name, $N(t)/M(t)$, whereas rich-get-richer is controlled by the abundance-weighted mean, $\bar{Q}_{\text{RGR}}(t)/N(t)$. Since the abundance variance is non-negative,

$$\frac{\bar{Q}_{\text{RGR}}(t)}{M(t)} - \left[\frac{N(t)}{M(t)} \right]^2 \geq 0, \quad (\text{S.74})$$

we have $\bar{Q}_{\text{RGR}}(t)/N(t) \geq N(t)/M(t)$. Therefore, $d\bar{Q}_{\text{RGR}}/dt \geq d\bar{Q}_{\text{R}}/dt$. With the same initial condition, this implies $\bar{Q}_{\text{RGR}}(t) \geq \bar{Q}_{\text{R}}(t)$. Since both models have the same $N(t)$ and the same expected number of active names $M(t)$, the corresponding abundance variances satisfy

$$\bar{V}_{\text{RGR}}(t) \geq \bar{V}_{\text{R}}(t). \quad (\text{S.75})$$

Combining this inequality with the comparison between prestige-driven reinforcement and pure rich-get-richer, we obtain the hierarchy

$$\bar{V}_{\text{RIP}}(t) \geq \bar{V}_{\text{RGR}}(t) \geq \bar{V}_{\text{R}}(t). \quad (\text{S.76})$$

The first inequality follows from the additional heterogeneity induced by random prestige increments, whereas the second follows from the reinforcement of already abundant names in rich-get-richer. Random reuse lacks both amplification mechanisms and therefore produces the smallest abundance variance within this comparison.

Asymptotic Limits in the Closed-Repertoire Regime (No Innovation)

We consider again the closed-repertoire limit where the discovery of novel elements is suppressed, locking the total number of unique names at a fixed baseline, i.e. $M(t) = M_0$.

RGR model: In the baseline pure rich-get-richer model, provided $\rho(t) = 0$, Eq. (S.66) can be solved explicitly using the integrating factor as

$$\bar{Q}_{\text{RGR}}(t) = \left[\frac{Q_0}{N_0^2} + \frac{1}{N_0} - \frac{1}{N(t)} \right] N^2(t). \quad (\text{S.77})$$

As chronological time advances to the asymptotic limit ($t \rightarrow \infty$, implying $N(t) \rightarrow \infty$), the transient $1/N(t)$ term vanishes, leaving the leading-order scaling behavior:

$$\bar{Q}_{\text{RGR}}(t) \sim \left[\frac{Q_0}{N_0^2} + \frac{1}{N_0} \right] N^2(t). \quad (\text{S.78})$$

Then, substituting in Eq. (S.39) and recalling that $M(t) = M_0$ and $N(t) = N_0 + t$ gives:

$$\bar{V}_{\text{RGR}}(t) \sim \left[\frac{1}{M_0} \left(\frac{Q_0}{N_0^2} + \frac{1}{N_0} \right) - \frac{1}{M_0^2} \right] (N_0 + t)^2. \quad (\text{S.79})$$

Random model: In the random reuse model, names are selected uniformly from the closed pool without historical feedback. Provided $\rho(t) = 0$, Eq. (S.73) can be solved explicitly as

$$\bar{Q}_{\text{R}}(t) = Q_0 + [N(t) - N_0] + \frac{N^2(t) - N_0^2}{M_0}. \quad (\text{S.80})$$

Then, substituting in Eq. (S.39) and recalling that $M(t) = M_0$ and $N(t) - N_0 = t$ gives

$$\bar{V}_{\text{R}}(t) = \left(\frac{Q_0}{M_0} - \frac{N_0^2}{M_0^2} \right) + \frac{N(t) - N_0}{M_0} \quad (\text{S.81})$$

$$= \mathcal{V}(0) + \frac{t}{M_0} \quad (\text{S.82})$$

Substituting $N(t) - N_0 = t$ simplifies the continuous-closure random reuse variance to a strictly linear relationship

$$\bar{V}_{\text{R}}(t) = \mathcal{V}(0) + \frac{t}{M_0}. \quad (\text{S.83})$$

The analytical expressions derived within this continuous-time closure framework in Eqs. (S.53)-(S.79)-(S.83) confirm the macroscopic asymptotic hierarchy

$$\bar{V}_{\text{R}}(t) = \mathcal{O}(t), \quad \bar{V}_{\text{RGR}}(t) = \mathcal{O}(t^2), \quad \bar{V}_{\text{RIP}}(t) = \mathcal{O}(t^2). \quad (\text{S.84})$$

Clearly, the random model has a lower variance than the other two, as the growth depends linearly on t . Moreover, we can subtract the two asymptotic variances given by Eqs. (S.53)-(S.79) to isolate the excess heterogeneity generated by our model

$$\bar{V}_{\text{RIP}}(t) - \bar{V}_{\text{RGR}}(t) \sim \frac{1}{M_0 N_0} \frac{\sigma^2}{\mu^2} N^2(t). \quad (\text{S.85})$$

This formula demonstrates that while both mechanisms drive a quadratic accumulation of variance, prestige-driven reinforcement yields a strictly larger prefactor whenever the internal weights are updated via heterogeneous increments ($\sigma^2 > 0$).

We note that solving the process using a continuous-moment closure introduces minor finite-size discretization differences relative to an exact discrete-time calculation. For example, in the pure rich-get-richer model, the true discrete-time Pólya urn prefactor replaces the continuous-closure term:

$$\left[\frac{Q_0 + N_0}{N_0^2} \right] \rightarrow \left[\frac{Q_0 + N_0}{N_0(N_0 + 1)} \right]. \quad (\text{S.86})$$

Similarly, for the random reuse model, the exact discrete-time calculation yields a linear slope of $(M_0 - 1)/M_0^2$ instead of the continuous-closure slope of $1/M_0$. These variations are strictly finite-size discretization corrections that vanish as $N_0 \rightarrow \infty$; they leave the underlying scaling laws and the asymptotic hierarchy qualitatively unchanged.

Limit scenario of full innovation

If every election introduces a new name, then there are no reinforcement events. If the initial condition consists only of names that have appeared once, then $N_0 = M_0$ and $Q_0 = M_0$ regardless of the model. Consequently, in all cases

$$\bar{V}(t) = 0. \quad (\text{S.87})$$

Limit scenario of constant innovation in the large-times limit

We now consider $\rho(t) = \rho$ in the $t \rightarrow \infty$ limit, where we recall that as $N(t) \sim t$ and $M(t) \sim \rho t$, $N(t)/M(t) \rightarrow 1/\rho$.

RGR Model: For the macroscopic collective abundance variance to converge toward a stable, non-divergent finite ceiling at long times, the second moment $\bar{Q}_{\text{RGR}}(N)$ in Eq. (S.66) must scale linearly with the total system size N as $\bar{Q}_{\text{RGR}}(N) \sim c_{\text{RGR}} N$. Substituting the ansatz in Eq. (S.66) we are able to obtain that

$$\bar{V}_{\text{RGR}}(\infty) = \frac{c_{\text{RGR}}}{\rho} - \frac{1}{\rho^2} \quad (\text{S.88})$$

$$= \frac{1 - \rho}{\rho^2(2\rho - 1)}. \quad (\text{S.89})$$

As in the prestige model, when the entry rate of novel elements drops below this critical threshold ($\rho \leq 1/2$), the ongoing influx of innovation is no longer sufficient to dilute the competitive compounding advantages of early-mover elements. In this weak-innovation regime, the raw second moment $\bar{Q}_{\text{RGR}}(t)$ grows strictly faster than linearly with respect to $N(t)$ (scaling as $\mathcal{O}(t \ln t)$ at $\rho = 1/2$ and as $\mathcal{O}(t^{2-2\rho})$ for $\rho < 1/2$). As a result, the collective abundance variance diverges indefinitely, preventing the system from ever settling into a stable, stationary structural equilibrium.

Random Model: In the random case, we can directly substitute the stationary state limit $N(t)/M(t) \rightarrow 1/\rho$ directly into Eq. (S.73) and integrate the equation to obtain yields the leading order asymptotic rate of change:

$$\bar{Q}_{\text{R}}(t) \sim \frac{2 - \rho}{\rho} t. \quad (\text{S.90})$$

Substituting in Eq. (S.66) and taking the $t \rightarrow \infty$ we are able to obtain that

$$\bar{V}_{\text{R}}(\infty) = \frac{1 - \rho}{\rho^2}. \quad (\text{S.91})$$

In this case, the steady-state solution is unconditionally stable and structurally bounded for any active innovation rate $\rho > 0$.

Synthesizing these results for the strong-innovation regime ($\rho > 1/2$) where all the variances are defined, we recover that

$$\bar{\mathcal{V}}_{\text{RIP}}(\infty) \geq \bar{\mathcal{V}}_{\text{RGR}}(\infty) \geq \bar{\mathcal{V}}_{\text{R}}(\infty). \quad (\text{S.92})$$

The first inequality ($\bar{\mathcal{V}}_{\text{RIP}}(\infty) \geq \bar{\mathcal{V}}_{\text{RGR}}(\infty)$) isolates the additional macroscopic heterogeneity injected by the stochastic updates of prestige increments (quantified by the squared coefficient of variation σ^2/μ^2). This inequality collapses to a perfect identity if and only if prestige increments are purely deterministic ($\sigma^2 = 0$). The second inequality ($\bar{\mathcal{V}}_{\text{RGR}}(\infty) \geq \bar{\mathcal{V}}_{\text{R}}(\infty)$) emerges because history-dependent empirical feedback loops actively amplify early random sampling fluctuations, whereas random reuse continually dampens fluctuations back toward a uniform baseline.

SUPPLEMENTARY NOTE 4: ISOMORPHIC NETWORK MODEL

The RIP model can be represented as a growing network whose connected components correspond to names. This duality offers a topological interpretation of naming history.

The network formulation

Each node represents one ruler, pontificate, or reign, and is endowed at birth with an intrinsic prestige value $\eta_v = K_v$, drawn independently from the prestige-increment distribution $P(K)$. At time t , let $C_i(t)$ denote the connected component associated with name i . The abundance and accumulated prestige of name i are then defined as

$$n_i(t) = |C_i(t)|, \quad x_i(t) = \sum_{v \in C_i(t)} \eta_v, \quad (\text{S.93})$$

and the total prestige in the system is

$$X(t) = \sum_i x_i(t) = \sum_v \eta_v. \quad (\text{S.94})$$

At each time step, a new node v_t enters the system with intrinsic prestige $\eta_{v_t} = K$, where $K \sim P(K)$. With probability $\rho(t)$, the new node does not attach to any previous node. It therefore becomes the root of a new connected component, corresponding to the introduction of a new name:

$$n_{\text{new}}(t+1) = 1, \quad x_{\text{new}}(t+1) = K. \quad (\text{S.95})$$

With complementary probability $1 - \rho(t)$, the new node attaches to one previous node u . The target node is selected with probability proportional to its intrinsic prestige,

$$\Pi_u(t) = \frac{\eta_u}{\sum_w \eta_w} = \frac{\eta_u}{X(t)}. \quad (\text{S.96})$$

If $u \in C_i(t)$, the new node joins component C_i . Therefore,

$$n_i(t+1) = n_i(t) + 1, \quad x_i(t+1) = x_i(t) + K. \quad (\text{S.97})$$

Coarse-graining: from nodes to names

The probability that a reused name is i is the probability that the new node attaches to any node belonging to component $C_i(t)$. Hence,

$$P_i(t) = [1 - \rho(t)] \sum_{u \in C_i(t)} \Pi_u(t) = [1 - \rho(t)] \sum_{u \in C_i(t)} \frac{\eta_u}{X(t)}. \quad (\text{S.98})$$

Using the definition of component prestige,

$$\sum_{u \in C_i(t)} \eta_u = x_i(t), \quad (\text{S.99})$$

we obtain

$$P_i(t) = [1 - \rho(t)] \frac{x_i(t)}{X(t)}, \quad (\text{S.100})$$

which is exactly the reuse probability of the RIP model. Similarly, the probability of creating a new component is

$$P_{\text{new}}(t) = \rho(t). \quad (\text{S.101})$$

Thus, after coarse-graining connected components into names, the growing network is pathwise equivalent to the RIP urn process.

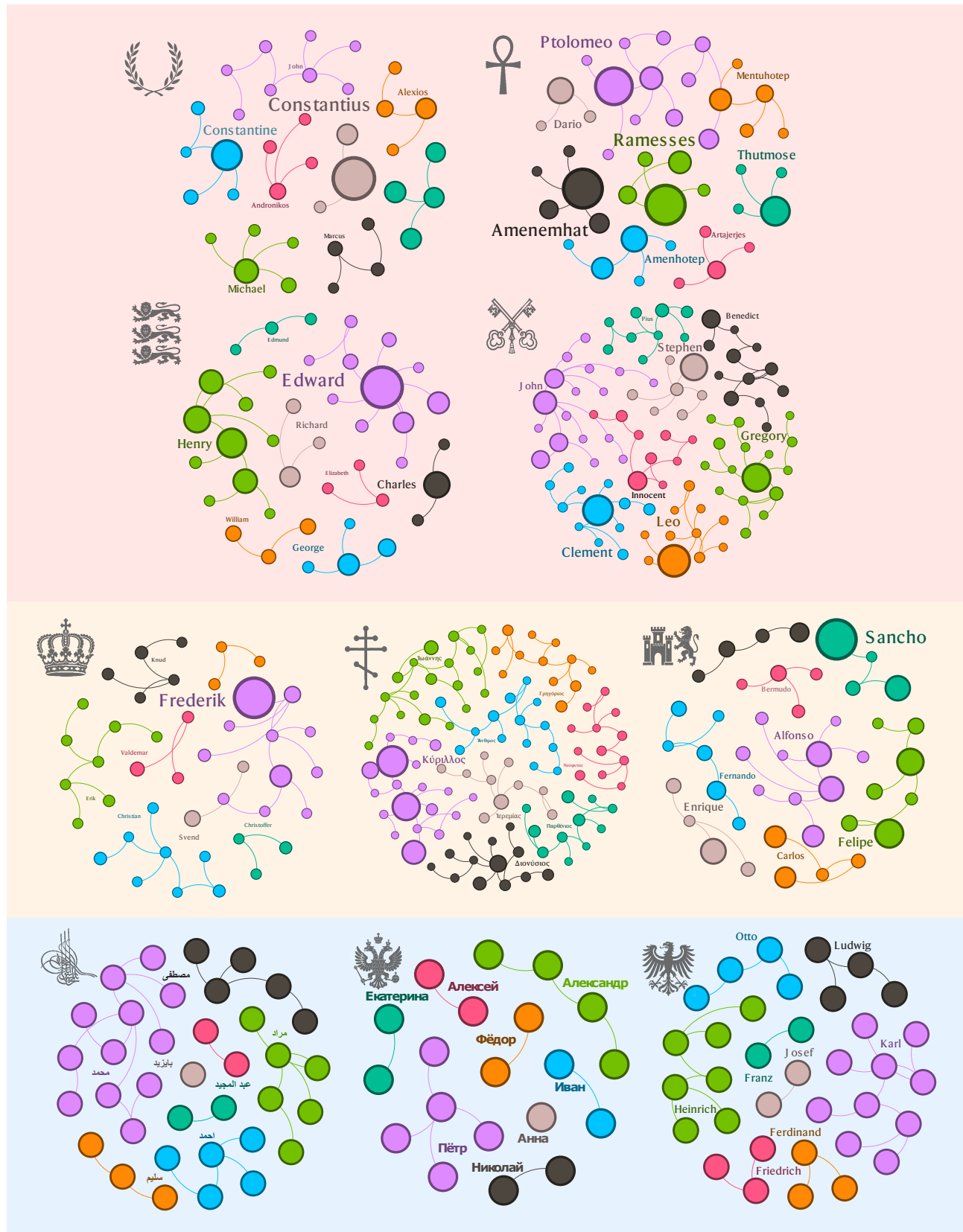
Relation to standard network growth models

This representation differs from classical preferential-attachment models. In the Barabási–Albert model [56], attachment is driven by degree, whereas here the target node is selected according to its intrinsic prestige and independently of its degree. A node with small degree can therefore attract later rulers if its prestige is sufficiently large. The model also differs from the Bianconi–Barabási model [65], where attachment depends on a product of degree and fitness. Here, the microscopic attachment rule is purely fitness-driven.

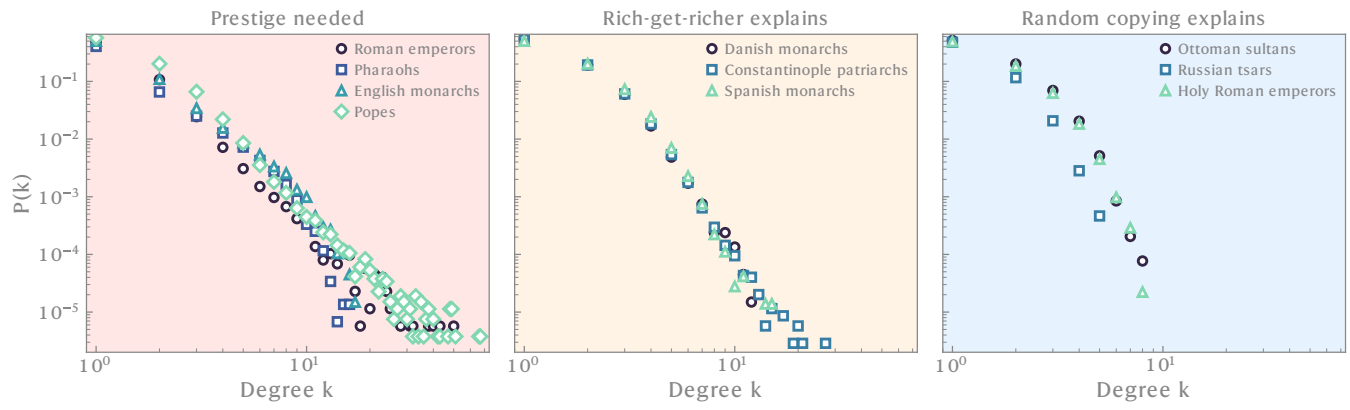
coarse-graining: names with larger accumulated prestige contain more total fitness and are therefore more likely to attract future nodes. Innovation corresponds to the birth of new connected components. The RIP model is therefore equivalent to a growing fitness forest with component birth, where component sizes are name abundances and component weights are accumulated name prestige.

Realizations of the isomorphic growing network model

In Supplementary Fig. 4 we show network visualizations corresponding to one realization of the RIP model for each ruling dynasty. The parameters of the prestige-increment distribution are those reported in Table 1. Nodes represent individual rulers or reigns, connected components represent names, and node size is proportional to individual prestige. In Supplementary Fig. 5 we further quantify these network structures through their degree distributions $P(k)$. The datasets for which prestige is required display broader degree distributions, indicating the emergence of hubs generated by rare high-prestige reigns. In contrast, datasets well explained by random sampling dynamics show steeper degree profiles, consistent with all nodes having the same prestige.



Supplementary Fig. 4. **Visual representation of the growing network model isomorphic to the RIP model.** Nodes represent individual rulers. Edges connect a new ruler to the specific predecessor selected with probability proportional to its individual prestige, rather than to the most recent usage or the name's founder. Consequently, distinct connected components emerge naturally as lineages of names, where the size of each component corresponds to the name's total frequency.



Supplementary Fig. 5. **Degree distributions of the network representation of the RIP model.** Degree distributions $P(k)$ of the RIP growing network representation for the ten dynastic naming traditions, grouped according to the mechanism identified in Fig. 2 of the main text. For each dataset, the distribution is averaged over 10^3 independent realizations using the empirical innovation function $\rho(t)$ and the fitted prestige-increment distribution $P(K)$.