Idea engines: Unifying innovation & obsolescence from markets & genetic evolution to science
Edward D. Lee, Christopher P. Kempes, and Geoffrey B. West

Innovation and obsolescence describe dynamics of ever-churning and adapting social and biological systems, concepts that encompass field-specific formulations. We formalize the connection with a reduced model of the dynamics of the “space of the possible” (e.g., technologies, mutations, theories) to which agents (e.g., firms, organisms, scientists) couple as they grow, die, and replicate. We predict three regimes: The space is finite, ever growing, or a Schumpeterian dystopia in which obsolescence drives the system to collapse. We reveal a critical boundary at which the space of the possible fluctuates dramatically in size, displaying recurrent periods of minimal and of veritable diversity. When the space is finite, corresponding to physically realizable systems, we find surprising structure. This structure predicts a taxonomy for the densities of agents near and away from the innovative frontier that we compare with real-world data, consistent with “follow-the-leader” dynamics in firm cost efficiency and biological evolution, whereas scientific progress reflects consensus that waits on old ideas to go obsolete. Our minimal model derived from first principles aligns with empirical distributions of firm productivity, COVID diversity, and citation rates for scientific publications. Our contribution is to map proxies of innovativeness from empirical examples, implying a follow-the-leader dynamic in firm cost efficiency and biological evolution, whereas scientific progress reflects consensus that waits on old ideas to go obsolete. Our theory introduces a fresh and empirically testable framework for unifying innovation and obsolescence across fields.

Understanding the dynamics and structure of innovation and obsolescence has been a subject of considerable interest across many domains ranging from business, economics, and technology to evolutionary biology, medicine, and the physical sciences. The forces of innovation and obsolescence define classical capitalist markets, summarized in Schumpeter’s iconic term “creative destruction,” based on the idea that new methods of production survive by eliminating existing ones (1). This is echoed in Spencer’s description of evolution as the “survival of the fittest” and in Spielrein’s “destruction as the cause of coming into being” for psychological development (2, 3). In the natural sciences, we have the more charitable adage from Newton that we build “on the shoulders of giants.” Each of these aphorisms implies that the new destroys or eclipses the old. A key point is that innovation of one thing often causes the obsolescence of another. A second key point is that agents, such as firms, organisms, or scientists, are themselves creating technologies, behaviors, or capacities that they could adopt from the (potentially changing) set of the possible, while disregarding the irrelevant. In the substantial literature, each area has developed its own particular formalization of the problem that has masked their fundamental similarities. Furthermore, there have been few attempts to formulate the problem of innovation and obsolescence in an analytic framework that is quantitative, predictive, and testable. Indeed, one of the major obstacles for validating models of innovation and obsolescence is that they make predictions that are difficult to test empirically—not to mention across examples as diverse as economics, diseases, and science. One of our contributions is to map proxies of innovativeness from empirical examples to model predictions of the density of agents near the innovative frontier using examples of firm productivity, the emergence of widespread clades in viral mutations, and citation rates that represent the wave of attention across the innovation front. These concrete examples align surprisingly well with our theory derived from first principles, capture essential features of shared dynamics, and thus connect diverse systems within a unified theory.

We picture innovation and obsolescence as taking place in a space of possibilities, an idea lattice, in which agents live and that is itself constantly churning as shown in Fig. 1A. Here, each vertex \( x \) represents an “idea” in which one or many agents are invested akin to Kauffman’s space of the possible (4). In the case of markets, a natural agent would be the firm and an innovation could be a method of manufacture or practice that improves profit margins. In the case of viral families, an agent may correspond to a

**Significance**

We introduce a mathematical model to connect innovation and obsolescence across diverse systems like economics, biology, and science. Taking the “space of the possible” (e.g., potential technologies, mutations, theories), we model how it grows from innovation and decays with obsolescence. We identify three scenarios: a finite, an ever-expanding, or a collapsing space—the latter a Schumpeterian dystopia where innovations fail to outpace obsolescence. A key finding is a boundary where this space fluctuates dramatically in size, showing periods of low and high diversity. The model aligns with real-world data, consistent with “follow-the-leader” dynamics in business and biology but scientific progress that waits on the obsolescence of old ideas. This work provides a fresh, testable framework, offering insights across multiple fields.
We approximate this by binning the obsolescence front to the left at $B$ mutations (6, 7); agents live in this lattice, and some sites will them with industry sector codes or with function for genetic where agents can be neatly assigned to sets such as by clustering grainning simplification is akin to choosing a scale of analysis, with related ideas, denoting either material similarity, shared science, agents might be author combinations, and ideas might structures such as trees (19, 20), extinction is a natural result of being outcompeted. (18); and in biological evolution binary way from one to another (18); and in biological evolution, obsolescence may depend on environmental shifts (22). Thus, a general theory must encompass different relationships between innovation and obsolescence of which creative destruction and complementarity are special cases.

A secondary distinction between fields is the choice of what is being innovated (the idea lattice) and who is doing the innovating (the agent). For example, a site on the idea lattice may represent a product offered or a manufacturing method used by a firm; here, we are taking “innovation” in the Schumpeterian sense as an economically successful idea, in contrast to an “invention” which may not be, a distinction that has been extended to biology (24). Similarly, the composition of the lattice could be widespread mutations in a population (as opposed to all enumerable ones) or a time-ordered list of topics that have emerged in the scientific literature. The definition of the unit of innovation has been an especially vexing problem in biology, where innovations span multiple scales from the genotypic, phenotypic, behavioral, to the ecological (24, 25). Such distinctions are essential in terms of mechanism and dynamics. Furthermore, the space of the possible is shaped by exogenous forces such as by resource constraints on organism metabolism (26, 27) or physical constraints on the distribution of new mutations (28, 29). By focusing on the elementary tension between innovation and obsolescence, we aim to integrate the variety of mappings and constraints, the details of which could be represented as mathematical relations between a general set of dynamical parameters.

A Generalized Model from First Principles

As concrete examples for our abstracted model, we consider a few empirical contexts that provide inspiration. When taking the aforementioned example of markets and agents as firms and innovations as a method of manufacture or practice that improves profit margins, a lattice site could be marked by productivity or cost efficiency, given a particular industry. For viral families such as SARS-CoV-2, we take an agent as a particular strain, which is genotypically distinct from others. In science, agents might be author combinations, and ideas might be conceptual connections. In our model, ideas share an edge with related ideas, denoting either material similarity, shared inputs and skills, or common ancestry (5). Clustering of similar ideas allows us to compress related items into a single site, which leads to the linear lattice approximation in Fig. 1B. This coarse-graining simplification is akin to choosing a scale of analysis, where agents can be neatly assigned to sets such as by clustering them with industry sector codes or with function for genetic mutations (6, 7); agents live in this lattice, and some sites will be more densely occupied than others such as agents i and j in Fig. 1B. Then, we define an obsolescence front at which ideas go defunct (vertices removed) and an innovation front at which agents drive ahead into the “adjacent possible” (vertices added) (4).

Our abstraction makes clear that a primary distinction between fields is the relationship assumed between innovation and obsolescence. In Schumpeter’s creative destruction, the relationship is one of conservation, where productive innovation comes at the expense of an existing mode of production. A realization of this is the study of economic competition as new methods of production are innovated (9, 10). It follows that firms on a line of productivity margins (11, 12), and obsolescence occurs endogenously either from desuetude or unsustainable profit margins (13). Across many areas obsolescence is the complement of innovation: In the study of scientific progress (14–16), reduced citation rates imply that articles are being forgotten or becoming increasingly irrelevant; in social change (17), norms switch in a binary way from one to another (18); and in biological evolution (19, 20), extinction is a natural result of being outcompeted.

Other examples, however, reveal more complicated relationships between innovation and obsolescence. In markets, obsolescence may be driven by external research developments funded by government programs (e.g., GPS, Internet, mRNA vaccines) or by products from technologically advanced neighbors. Some innovations open many more possibilities than they close (21). The network structure connecting innovations to one another may drive cascades of either innovation or obsolescence (22).

As concrete examples for our abstracted model, we consider a few empirical contexts that provide inspiration. When taking the aforementioned example of markets and agents as firms and innovations as a method of manufacture or practice that improves profit margins, a lattice site could be marked by productivity or cost efficiency, given a particular industry. For viral families such as SARS-CoV-2, we take an agent as a particular strain, which is genotypically distinct from others. In science, agents might be author combinations, and ideas might be conceptual connections. In our model, ideas share an edge with related ideas, denoting either material similarity, shared inputs and skills, or common ancestry (5). Clustering of similar ideas allows us to compress related items into a single site, which leads to the linear lattice approximation in Fig. 1B. This coarse-graining simplification is akin to choosing a scale of analysis, where agents can be neatly assigned to sets such as by clustering them with industry sector codes or with function for genetic mutations (6, 7); agents live in this lattice, and some sites will be more densely occupied than others such as agents i and j in Fig. 1B. Then, we define an obsolescence front at which ideas go defunct (vertices removed) and an innovation front at which agents drive ahead into the “adjacent possible” (vertices added) (4).

Our abstraction makes clear that a primary distinction between fields is the relationship assumed between innovation and obsolescence. In Schumpeter’s creative destruction, the relationship is one of conservation, where productive innovation comes at the expense of an existing mode of production. A realization of this is the study of economic competition as new methods of production are innovated (9, 10). It follows that firms on a line of productivity margins (11, 12), and obsolescence occurs endogenously either from desuetude or unsustainable profit margins (13). Across many areas obsolescence is the complement of innovation: In the study of scientific progress (14–16), reduced citation rates imply that articles are being forgotten or becoming increasingly irrelevant; in social change (17), norms switch in a binary way from one to another (18); and in biological evolution (19, 20), extinction is a natural result of being outcompeted.

*We do not require in this formulation that the space of the possible be “prestatable” (8). We only require that at every point in time, the lattice is built out into the adjacent possible in some typical way characterized by the dimension.

A Generalized Model from First Principles

As concrete examples for our abstracted model, we consider a few empirical contexts that provide inspiration. When taking the aforementioned example of markets and agents as firms and innovations as a method of manufacture or practice that improves profit margins, a lattice site could be marked by productivity or cost efficiency, given a particular industry. For viral families such as SARS-CoV-2, we take an agent as a particular strain, which is genotypically distinct from others. In science, agents might be author combinations, and ideas might be conceptual connections. In our model, ideas share an edge with related ideas, denoting either material similarity, shared inputs and skills, or common ancestry (5). Clustering of similar ideas allows us to compress related items into a single site, which leads to the linear lattice approximation in Fig. 1B. This coarse-graining simplification is akin to choosing a scale of analysis, where agents can be neatly assigned to sets such as by clustering them with industry sector codes or with function for genetic mutations (6, 7); agents live in this lattice, and some sites will be more densely occupied than others such as agents i and j in Fig. 1B. Then, we define an obsolescence front at which ideas go defunct (vertices removed) and an innovation front at which agents drive ahead into the “adjacent possible” (vertices added) (4).

Our abstraction makes clear that a primary distinction between fields is the relationship assumed between innovation and obsolescence. In Schumpeter’s creative destruction, the relationship is one of conservation, where productive innovation comes at the expense of an existing mode of production. A realization of this is the study of economic competition as new methods of production are innovated (9, 10). It follows that firms on a line of productivity margins (11, 12), and obsolescence occurs endogenously either from desuetude or unsustainable profit margins (13). Across many areas obsolescence is the complement of innovation: In the study of scientific progress (14–16), reduced citation rates imply that articles are being forgotten or becoming increasingly irrelevant; in social change (17), norms switch in a binary way from one to another (18); and in biological evolution (19, 20), extinction is a natural result of being outcompeted.

*We do not require in this formulation that the space of the possible be “prestatable” (8). We only require that at every point in time, the lattice is built out into the adjacent possible in some typical way characterized by the dimension.
system with death rate $r_d$. In the special case of $r = r_d$, an agent could be its own descendant, mimicking agents that hop from one lattice site to the next. Our treatment considers these as two basic independent parameters whose interrelationship maps to a wide range of scenarios.

At the innovation front, agents encounter the additional difficulty of inventing what is possible before occupying it. In the case where agents innovate independently, the rate of successful innovation is proportional to how often they seek to expand, $r$, their innovativeness, $I$, and the number of agents at the innovation front, $n(x = 0, t)$, thereby summarizing the complex process of discovery in terms of a mean rate (30, 31) (SI Appendix, A).

Finally, we incorporate obsolescence by assuming that the oldest idea goes obsolete with a rate of $r_o$, simultaneously eliminating all agents on the site; below, we consider a more general case where this constraint is relaxed. Consequently, the rate of change of the lattice length $L(t)$ is the difference between the rates of innovation and obsolescence:

$$L(t) = rIn(0, t) - r_o.$$ \[1\]

A minimum allowed length of $L = 1$ corresponds to when the two fronts coalesce. When the lattice length is stable $L(t) = 0$, the system looks like an “innovation train” moving into the innovation frontier with innovativeness directly related to obsolescence with

$$n(0, t) = r_o(I)^{-1}.$$ \[2\]

To simplify the mathematical treatment, we imagine “sitting” on the train and fixing the coordinate system such that the innovation front is at the origin $x = 0$ with its movement represented by the train tracks moving past us to the right (this reverses the coordinate system from left to right from what is depicted in Fig. 1). Putting these together, the rate of increase in the number of agents $n(x, t)$ at lattice site $x$ at time $t$ is

$$\dot{n}(x, t) = \frac{G}{L(t)} + r[n(x + 1, t) - r_d n(x, t) - r_o n(0, t)[n(x, t) - n(x - 1, t)]] - rIn(0, t)n(x, t).$$ \[3\]

The first term is the rate at which new firms enter the system, the second the rate at which they replicate by mimetic innovation (32), the third the rate at which they leave the system, and the last term the effective shift from the motion of the innovation front. We confirm the accuracy of the solution in Eq. 9 with numerical calculation including simulations that are further detailed in SI Appendix, B. These results summarize the mean-field dynamics of a parsimonious system of agents growing, dying, and innovating in a one-dimensional space.

Below, we show how these equations can be straightforwardly generalized to include more complex dynamics such as cooperative innovation, higher-dimensional graphs, inverted obsolescence-driven innovation, and delayed obsolescence.

Creative Destruction, Runaway Innovation, & Collapse

Eq. 1 predicts three regimes of idea graph dynamics: i) Innovation and obsolescence are roughly balanced, leading to a typical size of the idea space; ii) innovation outpaces obsolescence and the system grows indefinitely, providing an unbounded “marketplace” for exploitation; iii) obsolescence outpaces innovation and the system collapses to only a few, transient ideas. In order to demonstrate the three regimes, we provide three examples of a stochastic automaton simulation following the dynamics specified in Eqs. 1 and 3, and illustrated in Fig. 2 (SI Appendix, B). When the rate of obsolescence $r_o$ is sufficiently small (blue line), we are in the regime of runaway innovation. As we increase $r_o$, we pass through a regime of steady lattice length (orange line) to a small lattice that repeatedly collapses to its minimum size $L = 1$ (green line). As a point of departure for analysis, we start with a stationary configuration in which each innovation extinguishes one old idea, the assumption underlying Schumpeter’s original formulation of creative destruction (1).

The stationary condition is a fundamentally collective property.\(^1\) In Eq. 2, global stability means that the number of agents on the leading edge is proportional to the rate at which ideas go obsolete, i.e., faster obsolescence means more highly innovative agents to sustain rapid progress. Alternatively, faster innovation leads to a drop in the number of innovative agents because fewer agents keep innovations apace. Under stochastic variation, this observation means that a temporary increase in the number of innovative agents $n(0, t)$ above steady state drives the innovation front ahead quickly, consequently reducing the number of innovative agents. This reduced number then slows the innovation front down and allows for agents to flow in from the existing lattice, resulting in oscillations around steady-state occupancy (Fig. 2B). Thus, the age of an idea near the innovation

---

\(^{1}\)In the balanced state, each lattice site has a typical lifetime, which is consistent with a mean-field formulation of Van Valen’s Red Queen hypothesis. This predicts that every taxon has a constant rate of extinction regardless of age (33, 34), even though some evidence suggests that the rate of extinction may change over time (35).
front is naturally related to the number of agents because newer sites have fewer agents.‡

Following the logic that older ideas will have accumulated more agents, we expect the typical number of agents to be minimal near the innovation front and to increase as we move toward older ideas. This describes what in physics is referred to as a "pseudogap," namely a drop in the density of excited states, around \( x = 0 \). If this is assumed to be approximately linear at stationarity, i.e., \( \Delta n(x = 0) \approx r \Delta n \), then \( n(1) - n(0) \approx n(0) - n(-1) \).

Once, by definition, the site at \( x = -1 \) is unoccupied, i.e., \( n(-1) = 0 \), this gives \( n(1) \approx 2n(0) \). Setting \( x = 0 \) in Eq. 3, we can then solve for lattice length at stationarity:

\[
L = GIrr^\frac{1}{2} (r_o + r_d - 2r)^{-1}. \tag{4}
\]

While Eq. 4 is only an approximation to the exact nontrivial solution, which manifests surprisingly complex variation (SI Appendix, A and Fig. 5A.3), it does provide a rather revealing starting point for capturing the essential characteristics of the resulting idea space.

Two pivotal points are indicated by Eq. 4: i) where obsolescence is perpetually outpaced by innovation and \( L \rightarrow \infty \); and ii) where obsolescence outpaces innovation and \( L \rightarrow 1 \) signaling the collapse of the system. Formally, the first can occur when \( G, I, \) or \( r \) become infinite, or when \( r_o = 0 \), none of which is realistic because they require infinitely fast rates or perpetual suppression of obsolescence. On the other hand, the singularity at \( r_o + r_d - 2r = 0 \) requires only balancing \( r_o/r \) and \( r_d/r \) such that

\[
r_o/r \leq 2 - r_d/r. \tag{5}
\]

We delineate this region in red in Fig. 3. Eq. 5 indicates that the typical number of times an agent replicates before it dies is a crucial order parameter and that agent-level properties can drive unbounded growth of the idea lattice, leading to runaway innovation (the first pivotal point).

The second pivotal point is reached when the innovation front number falls below a self-sustaining threshold, \( n(0, t) < (r/r_o)I \).

From Eq. 1, growth becomes negative, \( \dot{L}(t) < 0 \), driving the system to collapse to its minimum length \( L \sim 1 \). From Eq. 4, we can solve for the corresponding cutoff obsolescence rate as:

\[
r_o/r \sim 1 - r_d/2r + \sqrt{(1 - r_d/2r)^2 + GI/r}. \tag{6}
\]

In the limit \( GI/r \rightarrow 0 \), the boundary between collapse and growth shrinks to a line and eliminates the region of stability (white region in Fig. 3A). Consequently, we expect to find a sharp transition between collapse and growth, and the system can display long time scales and large fluctuations (37) (SI Appendix, D). Thus, Eqs. 5 and 6 elucidate the three regimes of lattice dynamics: a balanced, a runaway, and a collapsed regime.

### Extensions

This picture holds for several key generalizations. We discuss cooperative innovation, higher-dimensional graphs, inverted obsolescence-driven innovation (a reversed picture where obsolescence furthers system progress), and delayed obsolescence.

(i) Cooperative innovation implies that the front velocity scales nonlinearly with the number of agents as \( rIn(0) \). The default value of \( a = 1 \), as in Eq. 3, corresponds to agents innovating independently of one another. Here superlinearity, \( a > 1 \), signals cooperation, whereas sublinearity, \( a < 1 \), implies competition. From Eq. 2, these differences can be mapped back to Eq. 3 by the transformation \( I^a \rightarrow I \) and \( (r_o/r)^a \rightarrow r_o/r \).

(ii) For tree graphs, shown in SI Appendix, Fig. 5C.8, each sequential site branches into \( Q - 1 \) additional branches, one of which must be chosen by the next new agent. If branches are equally likely to be chosen, the replication term in Eq. 3 acquires an additional factor \( (Q - 1)^{-1} \) such that the number of agents systematically decreases toward the innovation front. This picture makes clear the importance of the relative dimensions of agent replication and the idea space. When next-generation agents do not fill all of the available space, then agents eventually occupy a small fraction of the idea space. Nevertheless, the dimensional depletion effect does not fundamentally alter the dynamics. If we rescale \( r \rightarrow (Q - 1)r \) and \( I \rightarrow I/(Q - 1) \) and give each branch the same entry rate as in the linear case (as if each branch were a replica of the linear lattice), then we again recover Eq. 3. As argued in SI Appendix, C, higher Euclidean dimensions can also be approximated by the linear model. Thus, important classes of dynamical or structural generalizations do not appreciably alter the basic model.

(iii) Obsolescence-driven innovation is the antithesis of forward-looking innovative dynamics. By reversing the direction of the \( x \)-axis in Fig. 1 and setting the innovation front at

---

‡This result agrees with the intuition, for example, that firms expand into “adjacent markets” because there is less competition (36).

---

**Fig. 3.** Model phase space. (A) Basic model predicts regimes in which the size of the space of the possible, or the lattice length, stabilizes (white), grows indefinitely (red), or collapses (blue). Parameter values are set to a rescaled growth rate \( GI/r = 1 \) and innovativeness \( I = 1 \). (Inset) Empirical examples located in the phase space include metal stamping firms from Fig. 4D (brown triangle), Indian firms from Fig. 4G (small/large green stars for small/large firms), patent citation curves from Fig. 4F (least/most as small/large purple circles), and Physical Review B citation curves from Fig. 4J (most citations as black X, fewest falls outside and to the right of plotted range). Model extensions to (B) anticooperative innovation, \( a = 1/2 \), and (C) cooperative innovation, \( a = 3/2 \). See SI Appendix, Fig. SC.9 for an example on a tree graph.
the origin, agents now replicate toward obsolescence. For stable pseudogaps, the introduction of a new idea propels every agent toward the innovative front, making the entire system more innovative. As a result, innovation is driven at a rate proportional to the number of agents on the verge of extinction, which is still given by Eq. 2. Agents occupying newer ideas tend to beget agents on older ideas, which then drive themselves to extinction by eventually increasing occupancy at the obsolescence front. This is evocative of Kuhnian scientific progress, where invalidation of old ideas permits new ideas, new ideas stimulate revision of existing topics, and the system as a whole progresses (44).

(iv) Delayed obsolescence occurs when lattice sites slowly and randomly decay after obsolescence such that they can disappear differently from order of appearance. We take the defunct tail to consist of innovations that no longer contribute to the process of innovation but linger around (e.g., horse and buggy, pseudogenes, professors emeritus).

We formalize this as a decay process that is instigated as soon as the obsolescence front reaches a lattice site. The decay process gives a distribution of lifetime, $\rho(t)$, such that the probability that the lattice site is still alive after time $t$ is $p_{\text{alive}}(t) = \int_0^\infty \rho(t) \, dt$ and the probability that it is gone after time $t$ is $p_{\text{extinct}}(t) = 1 - p_{\text{alive}}(t)$. Since the obsolescence front marches ahead at rate $r_o$, each following site is on average $1/r_o$ younger than the previously obsolescent site. If the decay process gives each site a constant rate of disappearing $\lambda$, then this corresponds to an exponential distribution. If the site becomes more stable or less stable over time, then it is described by the Weibull distribution, $\rho(r_o) = \lambda k (r_o)^{k-1} e^{-(r_o)^k}$. When $k = 1$, we recover the exponential case where $\lambda$ is the decay rate.

When $0 < k < 1$, we have long-lived vestigial innovations and when $k > 1$ we do not.

We find that the typical number of vestigial innovations is approximately (SI Appendix, C.4)

$$\langle n \rangle \approx \frac{\Gamma(2/k + 1)}{2\lambda^2},$$

where $\Gamma$ is the Euler gamma function. The average tail is approximately

$$\langle b \rangle \approx \frac{\infty}{k=0} \sum b e^{-(\lambda br_o)^k} e^{-\Gamma(k^{-1}, (\lambda br_o)^k)r_o(\lambda k)^{-1}},$$

expressed in terms of the upper incomplete Euler gamma function with exponent and lower integration bound as arguments, respectively. Both $\langle n \rangle$ and $\langle b \rangle$ diverge in the limits $k \to 0$ and $\lambda \to 0$. This means that as we approach the two end cases, the tail grows to infinity. Numerical calculations suggest that the tail grows exponentially fast as $\sim e^{1/k}$ when $k \to 0$, or as a power law singularity $\sim \lambda^{-\kappa}$ with positive exponent $\kappa$ when $\lambda \to 0$. Otherwise, there is a short, finite tail, whose composition turns over.

Innovation Distributions: Model vs. Data

The fact that Eq. 3 can be transformed to account for structural, dimensional, and reversed extensions of the model implies that its resulting pseudogap form predicts a common taxonomy. The taxonomy is highly constrained, collapsing onto a few characteristic forms for the density of agents near the adjacent possible, or the shape of the pseudogap $\rho(x)$ (SI Appendix, C.4).
Fig. SA.3). These are revealed in our approximate solution to Eqs. 1 and 3 at steady state of the form

\[ n(x) = Ae^{λ−} + Be^{λ+} + C, \]  

for constants \( A, B, \) and \( C \) and the characteristic eigenvalues \( λ− \) and \( λ+ \) from the homogeneous solution. Eq. 9 predicts three shapes: an exponential rise to a skewed peak (black curve in Fig. 2C); a slow, sublinear increase that dives exponentially to zero; and a wide, linear pseudogap (SI Appendix, Fig. SA.3). To compare with our predicted forms for the innovative adjacent possible are several empirical examples of which we plot in Fig. 4 from markets, genetic evolution, and science in terms of occupancy number vs. proxies for their distance from the innovation front. Given that the pseudogap taxonomy reflects a balance between a collapsed lattice (such as when \( G \rightarrow 0 \)) or runaway growth (when \( r_0 \rightarrow 0 \)), the empirical taxonomy reflects interplay between our model parameters that we explore.

The several examples in Fig. 4 include firms, viral genetic evolution, and technology and science, putting together examples that have so far been considered independently of one another in the literature. In each of the panels and with the dashed black lines, we show that one functional form—the predicted characteristic exponential rise to a skewed peak with a decaying tail—aligns surprisingly well with the data. As we discuss in further detail below, we must determine the appropriate axis along which to map the idea lattice coordinate. For firms, we use histograms along an economic proxy for innovativeness, although the outcome depends on whether or not one aggregates industries (panels A and D). In genetic evolution, a natural measure of innovativeness is in terms of the number of base-pair mutations that distinguish a particular viral strain. Then, plotting the number of unique strains per number of mutations leads to a histogram relative to the innovative frontier, which is the most genetically distant strain from the origin (panel B). Finally, in science and technology, we do not explicitly map the idea lattice. Instead, we follow in the footsteps of previous work to consider papers and patents as innovative combinations of ideas (30, 44). This implies that papers are typically somewhere on the frontier, and so the citation rate serves as a proxy for the density of agents near the frontier (panels C and G). That the citation rate obtained over separate literatures and subliteratures is meaningful is surprising, but this is one of the principal discoveries from the science of science: the “universal” citation dynamics that arch across fields of research (14, 45, 46). We go through each example in more detail below to explain how we unify them within our model.

Economics provides the classic example of Schumpeterian innovation measured in terms of cost efficiency, or the ability of a firm to extract profit from a fixed amount of investment (11–13, 47). This imposes a seemingly natural one-dimensional axis for ordering firms (the agents), where more innovative firms decrease cost per unit output (the idea lattice coordinate). Fig. 4B shows an example of the distribution of labor costs per value added for the US metal stamp industry in 1958 and 1963 from Iwai’s classic work (11, 12). The fewest number of firms are the most and the least cost-efficient, although the distribution is skewed to the right because firms with higher costs survive. This characteristic form also appears in recent distributions of Indian firms in Fig. 4G, but when plotted against cost efficiency, which is essentially the reversed x-axis (38) (see Materials and Methods for details about fitting procedure). In this case, larger firms as measured by total assets are typically less cost efficient than smaller firms—although at odds with the canonical interpretation of efficiency and innovation, it would be consistent with the narrative that disruptive technologies are initially inefficient to develop (e.g., Tesla). While this seems different from the first example, Iwai’s histograms are specific to a single industry, whereas Indian firms are aggregated across industries. More generally, innovativeness is a multifaceted concept with many definitions in the economics literature (48). Our dynamical formulation provides a derivation of the shape of economic distributions as a function of innovative distance and thus a way of resolving discrepancies by inferring the innovative frontier from data.

As an example of biological, or Darwinian, innovation, we consider the tree of SARS-CoV-2 clades measured from the GISAID repository of sequences (39, 40). The set of possible innovations, measured by base pair mutations, is a phylogenetic tree with a branching ratio \( Q \approx 2.3 \) per unit phylogenetic branch length (Materials and Methods). We take the number of base pair mutations (the lattice coordinate) from the first known strain hCoV-19/Wuhan/Hu-1/2019 to order mutants (the agents) in innovative order. Then, our model presents a recursive relation (SI Appendix, Eq. S1) that determines how the number of strains with \( k \rightarrow 2 \) and \( k \rightarrow 1 \) mutations sets the number of strains with \( k \) mutations, the prediction of which is substantially different from a random mutation null model (SI Appendix, F). As shown in Fig. 4E, the resulting occupancy plot again shows the same aforementioned characteristic form, which our model fits closely over nearly the entire period of observation.

As confirmation of what our model leaves out, we fail to capture the sharp, temporary increases when variants such as Delta emerge, indicating an important deviation from model predictions: These punctuated, unpredictable, and singular innovations violate our model assumptions—Einsteinian innovations represent a discontinuous shift in parameters (17). We highlight an example in Fig. 4H, where only the number of Delta clades are shown as a function of mutations from root. A similar Einsteinian innovation is revealed in long-term evolution experiments, where the sudden emergence of a mutator strain increases the mutation rate in an E. Coli population, indicated by the discontinuity in accumulated mutations in Fig. 4I (43). The punctuated changes in innovation rates reveal a meta-dynamic in the space of parameters notably specific to genetic innovation.

Science and technology, in contrast, builds on an edifice. New technological and scientific ideas must be tethered to the past and are often judged by their consistency with established knowledge, theory, and pedagogy (49). It is only when existing frameworks have been proved insufficient that a new idea can flourish. This suggests a reversed dynamics, where innovation is driven by the obsolescence of old ideas.

Scientific and patent citation rates support this picture as we show in Fig. 4 C and G. Citation rates peak to maximal prominence quickly then slowly fall out of favor with age, a horizontally mirrored version of the previous examples. In our formulation, we take the citation profile to be proportional to the occupancy function, which measures the wave of agents moving across a graph of sequential papers that mark the progression to new ideas (44). Thus, the agent is the set of authors and the lattice coordinate is time. Since papers contribute different levels of innovation attributed to some intrinsic fitness (14, 45, 50), we bin them based on citation count. Furthermore, we account for citation inflation (Materials and Methods). We show yearly variation in normalized citations received by scientific papers in Physical Review B published in 1980 in Fig. 4C and for patent applications filed in 1990 in electrical and electronics
Distance to Runaway Innovation

A natural way to compare the systems is to measure their distance $\Delta$ from the boundary of runaway innovation as an indicator of innovativeness. We define the distance in terms of the change in the rescaled parameters $r_o/r$ and $r_d/r$ required to reach the threshold defined in Eq. 4 or visually the distance from the red region in the examples plotted in the Fig. 3 A. Inset. It is only determined by the rescaled rates, allowing us to compare systems in a common way.

Reassuringly, we find similar distances between the dynamics of SARS-CoV-2 evolution in Europe and North America as we show in Fig. 5. This aligns with our expectations that the processes determining viral evolution were not dramatically different between the two regions. In contrast, small and large nonfinancial Indian firms differ, and this difference extends far beyond the variation in the parameters measured across the ten closest fits. The minuscule $\Delta$ for small Indian firms suggests that they, unlike large ones, lie at the boundary of creative destruction, a critical boundary at which we predict larger fluctuations in population cost efficiency and size.

When comparing patents and PRB, we find that documents with more citations tend to lie closer to runaway innovation. This difference is especially pronounced for PRB when comparing those with fewer than ten citations versus those with at least ten (see SI Appendix, Fig. SF.14 for all citation classes). A much smaller disparity appears for patents. This observation raises questions about the mechanisms that explain the disparity. As has been noted, a scientific citation is not just an indicator of innovativeness because it can refer to established pedagogy, corrections, substantiating evidence, etc., in contrast with how patents focus on elements that highlight novelty (54, 55). In other words, differences in the content, rules, and norms are in the microscopic and mechanistic equations behind our reduced, mean-field formulation, which systematically connects innovative variation with a few essential parameters.

Discussion

Constituents of social and biological systems constantly undergo turnover, expanding, exploiting, or reducing the space of realized capabilities in a sequence of innovations that eventually renders the innovator obsolete. In any particular system, the exact details of this process may be modeled by competition (13), innovative risks (56), resource constraints (57), limited attention (53, 58), and strategy (59), among other factors. Furthermore, we know that the process of innovation displays combinatorial dynamics (21, 30, 31, 60–62), depends on the topology and dimension of the adjacent possible (44, 62, 63), and is influenced by agent interaction (31, 64). While important, the plethora of factors obscures the fact that the same effective dynamics couple agents with the lattice on which they live (Fig. 1). The power of our generalization originates from incorporating these processes into mean rates that highlight three fundamental innovative regimes (Fig. 3) common across dynamical, dimensional, and structural extensions of the basic model.

Our focus on the elementary tension between innovation and obsolescence is an aspect that is repeatedly brought up in the literature and so even more surprisingly absent in many quantitative frameworks of innovation. Perhaps the most widely recognized example of this is Schumpeter’s “creative destruction.” The major dichotomy in the biological literature is the opposition between the “survival” of the fittest versus the “arrival” of the fittest (65). The influential Kuhnian formulation of scientific advance relies on the elimination of existing paradigms, the scientific revolution, to permit scientific progress on the large scale (49). Looking further afield, we see parallels in the synthesis of opposing ideas in Hegelian dialecticism. We have developed a framework for implementing and generalizing the cyclical concept with a formal, quantitative, and empirical model for innovation and obsolescence dynamics.

As a demonstration of our quantitative predictions, we focus on the region of parameter space corresponding to creative destruction, where each innovative advance is matched by obsolescence. There we predict a surprisingly rich taxonomy for agent density around the innovative frontier (SI Appendix, Fig. SA.4). In the region of parameter space where replication rate dominates over death, or $r_d/r < 1$, we find a skewed density characterized by a peak opposite a decaying tail that extends to the innovation front. The shape aligns with empirical examples on firm productivity, genetic evolution, and scientific citation (Fig. 4), an empirical similarity that has not to our knowledge been noted. Thus, our framework provides a way of unifying technologies in Fig. 4G. To test the predictive power of our model, we fit to only the first quarter of the duration shown, or the first decade after publication for PRB and the first 5 y after publication for patents. Our model fits the beginning of the citation curves remarkably well and captures the generic shape of the tails, but tends to underpredict their longer-term behavior. Possible explanations for the deviations are unconsidered effects such as bimodal memory (51) and debated “runaway” events (50, 52, 53). In addition, citation rates are biased in more recent years because citations from the newest papers and patents are missing. This bias is stronger for patents because we do not have information about applications that are pending review, a process that typically takes several years. Yet remarkably, innovation and obsolescence dynamics lead to a first-principles, predictive explanation for citation curves.

These examples demonstrate how our approach can serve as a baseline for isolating additional processes beyond basic innovation and obsolescence dynamics, especially given how it aligns with distributions of social and biological agents from the innovative to the obsolete.

Fig. 5. Distance to runaway innovation for best fit model. Numbers following patents and Physical Review B (PRB) indicate the range of citation counts as in Fig. 4. Error bars represent variability in the fit as the minimum and maximum of the five parameter sets with smallest fit residuals.

As a demonstration of our quantitative predictions, we focus on the region of parameter space corresponding to creative destruction, where each innovative advance is matched by obsolescence. There we predict a surprisingly rich taxonomy for agent density around the innovative frontier (SI Appendix, Fig. SA.4). In the region of parameter space where replication rate dominates over death, or $r_d/r < 1$, we find a skewed density characterized by a peak opposite a decaying tail that extends to the innovation front. The shape aligns with empirical examples on firm productivity, genetic evolution, and scientific citation (Fig. 4), an empirical similarity that has not to our knowledge been noted. Thus, our framework provides a way of unifying

---

5Accounting for such effects could drive down the tail by increasing the normalization factor (Materials and Methods).
phenomena that have so far been studied separately with a quantitative theory.

As one consequence, we locate the examples within the same parameter space and compare them systematically with one another. For instance, our fit parameters indicate that the systems are distinct in terms of the ratio of death to replication rate \( r_d/r \), where a value close to unity can correspond to a dynamic of agents hopping from lattice site to the next. We measure for PRB citations consistently \( r_d/r > 0.9 \), consistent with the hopping dynamic as if scientific teams are advancing in the direction of the innovation front. Nevertheless, the measured obsolescence rate is much larger, \( r_o/r \approx 4/3 \) for all above the least cited papers, which suggests that authors do not in the long run keep up with the publication frontier. This is consistent with evidence that scientists tend to cite a relatively static canon over their career (66). Similarly, we measure for metal stamping firms \( r_d/r = 0.95 \), suggesting a tendency to maintain or improve cost efficiency over time in the same firms. In contrast, the other examples indicate many replication events per site such as for Indian firms \( (r_d/r \approx 1/4) \), SARS-CoV-2 \( (r_d/r \approx 1/20) \), and patents \( (r_d/r < 10^{-2}) \). This suggests that these systems are dominated by newcomers exploiting past innovations. As a second point of comparison between systems, we define a distance to runaway innovation (Fig. 5). We find comparable values for SARS-CoV-2 mutants in both North America and Europe, but we find that small, nonfinancial Indian firms are closer to the critical boundary than large ones and the same for patents compared to PRB articles with a comparable number of citations. It remains an open question how these reflect mechanistic differences for pushing these systems toward or away from runaway innovation. In the context of debates about unleashing innovative forces in science and the economy, such metrics might be a tentative step toward deciding how systems might be modified to promote certain collective outcomes.

Of particular relevance is the fact that the examples are all relatively close to the boundary of runaway innovation. To locate them at the boundary, we first use the simple fact of diversity in the population, or that the systems are not in the collapsed regime. Furthermore, it is impossible for finite systems to generate infinite growth. This pins us to the steady state or a transient to it. Yet, this leaves open the possibility of finding systems outside steady state or displaying transitions between regimes. As a relevant example, a new mutator in the Long Term Evolutionary Experiment discontinuously breaks the rate at which the population cycles through new mutants (Fig. 4 F). Such unpredictable Einsteinian revolutions could reflect dynamics in the parameters of our model that may result from endogenous developments or in other cases as exogenous changes in environmental conditions (67–69). Further work may explain whether or not proximity to runaway innovation is representative of natural systems because of a general principle (e.g., evolution favors innovative diversity) or of interest (e.g., researchers study vibrant systems), questions that our analysis raises.

Importantly, our framework is not meant to be a unique explanation of the patterns that we find but a complementary perspective on system dynamics in the mathematical language of innovation and obsolescence. As one alternative, Assembly Theory describes how spaces of objects give rise to new objects. It could add detail to how innovation pathways are constructed beyond the simplifying assumptions made here (62). While we focus on mean-field rates, the respective systems considered in Fig. 4 build on the complex dynamics of firm competition and the emergence of new technologies (70), complicated equations for viral evolutionary dynamics (71), and particularities of scientific citation across fields and measures thereof (16, 50). The failures of our baseline model in fitting the precise shapes of firm productivity or the tails of the citation rates indicate missing features that might be incorporated into more detailed extensions of our model, and a new perspective may inspire new modeling insights.

Given these caveats, a natural question that arises is whether our framework could be used to promote innovative economies, inhibit viral evolution, or shape scientific progress. Our starting principles lead to several relevant insights. First, some system parameters may be more opportune than others in forcing a transition in idea lattice dynamics. It may be counterintuitive that boosting growth or innovativeness or reducing obsolescence are not the most forceful ways of maintaining a diverse set of ideas but balancing death and obsolescence rates is (Eq. 5). A second intriguing prediction is that transitions from Schum-peterian dystopia and runaway innovation can be sudden (SI Appendix, D). The fluctuations near the critical point highlight an opening for dynamical classification of systems through rate parameters or signaling when systems are on the verge of collapse (72), and it may reflect endogenous dynamics that drive large-fluctuations in biodiversity (23, 73), economic growth (74), or scientific decline (57). This puts forth the possibility of a comparative meta-dynamics, where we envision tracking systems in the innovative-obsolescence space with finer-scale dynamical data. Our model organizes these hypotheses and opens up a framework for thinking about the forces of innovation and obsolescence. After all, an engine may burst from too much fuel or putter out from too little, so likewise, an idea engine needs to be fine-tuned.

**Materials and Methods**

**Data.** We test our model with several datasets. Here, we describe the sources of the datasets and how we calculated the values that we show in Fig. 4.

The distributions of firm cost efficiency in Fig. 4 are digitized from figure 1 of ref. 11 by Iwai, and the densities in Fig. 4B are digitized from the third panel of figure 3 from ref. 38 by Jangili. In Iwai’s plot, we are showing the cost of labor relative to the cost of the product such that a cost of zero implies that the value of the sale is profit and other capital costs. Iwai also notes that a similar histogram appears across industries at the time of his investigation citing Sato’s 1975 publication. Iwai’s interpretation of the figure aligns with the classic interpretation of firm innovativeness in that more innovative firms are the ones lowering their cost ratios (13). On the other hand, Jangili’s distribution derives from an estimate of the relative cost efficiency of firms using a technique called stochastic frontier analysis (75). In short, this technique involves estimating the maximal cost efficiency that firms could hope to obtain from a given set of measures about firms (e.g., size, age, liquidity, leverage, capital-labor ratio, etc.) that convey information about the costs firms incur. When cost efficiency is unity, firms have reached maximal possible efficiency. Jangili’s work shows that the shape of firm distribution is relatively consistent over long periods of time, and the distribution does not drift toward perfect cost efficiency of one. Jangili’s SI Appendix, Figs. SA.1 and SA.2 are of particular interest, which show nearly the same inferred distribution of firm cost efficiency over 20 y. Another example of this characteristic shape is for European financial firms (76).

The classic assumption is that more efficient methods of production are more innovative because firms compete to improve their profit margins (11, 36). The level of aggregation, however, is important for comparing firms. As we point out in the main text, the level of aggregation is one major point of contrast between Iwai’s and Jangili’s analysis. As another counterpoint to the classic assumption, we note that it is at odds with the observation that disruptive firms, at first, are cost inefficient such as Tesla, which had been producing vehicles at a loss for many years (77, 78). In contrast, we do not start with the assumption that higher
cost efficiency is more innovative, but instead have a distribution for which we seek the most reasonable mapping of the given economic variable to our model. We emphasize that our measure of innovativeness takes into account the shape of firm density, not their absolute values of cost efficiency. In this sense, we identify the appropriate axes for innovativeness in metal stamping vs. Indian firms.

As is detailed in ref. 11, Iwais obtains the data from Satō (1975) that were originally obtained from the U.S. Department of Commerce. As is detailed in ref. 38, Jangili samples 11,410 nonfinancial, public, Indian firms between the years 1995 and 2014 listed in the PROWESS database maintained by the Centre for Monitoring Indian Economy, but the data are not publicly available. Small firms and large firms are distinguished by either belonging to the lowest or highest quartile of total assets.

We obtain the SARS-CoV-2 clade data from the Nextstrain project downloaded on August 10, 2022. We use the inferred phylogenetic trees based in the GISAID sequence repository that contains millions of global samples of SARS-CoV-2 strains. We focus on the European and North American subset since we expect these to be particularly well sampled though generally, it is nigh impossible to track all circulating strains. After mapping the imputed phylogenetic tree from Nextstrain into distance marked by base pair mutations, we calculate the average number of branches into which any unit length of the tree divides. In other words, a nonbranching unit contributes a branch $Q - 1 = 1$, one that leads to two clades $Q = 1 - 2 = 2$, etc., where $Q$ is defined as the branching ratio. Taking the average number of outgoing branches, we then find the branching ratio $Q = 2.317$ for North America and $Q = 2.344$ for Europe. Using the calculated value, we normalize the number of unique individual clades as by $(Q - 1)^y$, where $y$ is the number of mutations from the original detected strain in order to calculate the typical number of strains per branch. This will lead to an exponential decay if the number of detected strains remains constant as is approximately the case in the data far from the root strain hCoV-19/Wuhan/ Hu-1/2019. Note that in our model fits, we find second-order terms to be important, which are not captured by a simple exponential decay as would be predicted by a simple random mutation null model (SI Appendix, F).

The scientific article and patent citation data in panels $F$ and $I$ come from the American Physical Society’s repository for Physical Review B (PRB) and PatentsView (a repository for the US Patent and Trademark Office) (79), respectively. For PRB citations, we consider 1,369 papers published in 1980 and only citations within the universe of PRB papers. We first bin these papers by total cumulative citations to the papers till 2020 as a measure of fitness. While we wish to obtain a measure of interest in a paper over time, we also know that not only is the number of annual publications increasing over time but also the length of bibliographies. To obtain a citation rate that accounts for the changes exogenous to our model, we normalize the citation counts for each paper published in 1980 by the typical number of citations made by every paper published in each following year, i.e., we count effective citations relative to the typical number of citations made per paper per year.

On the other hand, we separate US patents by the six technology classes that are identified in ref. 42. In Fig. 4, we consider the 21,896 patents filed in category 4 (electrical & electronic) from 1990. We normalize the citation rates, following convention, by the number of patents filed each year within the focus technology category, a normalization factor that grows with the number of potential citation recipients within the same technology category (50, 51). This procedure again accounts for change in the population of citing patents (assumed to the proportional to the density of agents) that would not be captured in our simple model.

To map citations to our model, we take the citations within a publication as a proxy for the agent, author combinations. The picture is that there is an underlying idea lattice in science, which is tiled by the set of published articles. Under the assumption that a new scientific paper is at the frontier as long it accumulates some threshold number of citations (and because of its contribution however incremental), a citation made reveals where in the space the authors sit at that moment in time; papers have many citations, so authors reveal a superposition of location across the entire idea lattice. Counting each citation separately for each author combination, we obtain a density of attention for each site in the idea lattice per year. An additional assumption that we make is that the flow at each lattice site is approximated by the one-dimensional model. This approximation is justified by the argument in SI Appendix, C showing that the linear model will work for systems of higher dimension after the dynamics have been evolving for some time.

In future work extending on a more mechanistic model, it will be important to distinguish between authors and author combinations and types of citations by using the context of the citation (55). These effects will change how we weight each of the citations. Indeed, one major effect that we already use to weight the citations is the exponential increase in papers. Overall, the other effects are subtleties that are important but do not seem sufficiently so to obscure the characteristic shape of our density curves in the data.

The observed citation rates—because they initially peak then slowly decay—are consistent with the hypothesis that scientific ideas are driven by the extinction of obsolete ones. In both cases of citations we consider, the overall shapes align with our obsolescence-driven formulation.

Our analysis implies that the majority of agents in the scientific system are closer to the innovative edge, whereas most agents are closer to the obsolescence front in the other examples. This means that science as a whole is relatively innovative. This is a result of the fact that driving an idea obsolete in the stationary case effectively makes every agent more innovative. Perversely, the same dynamic drives out the most innovative agents from the system whose exit reflects the fact that scientific experts are not particularly good at valuing novel ideas (80). Firms and viral evolution, on the other hand, are forward-looking because the front is only driven ahead by most innovative agents. That means that fewer agents are near the innovative edge.

Fitting the Model. We fit model parameters by scanning through parameter space consisting of rescaled rates for growth $G/r$, obsolescence $r_0/r$, death $d_G/r$. Innovativeness $I$ does not need to be rescaled because it is unitless. At each test parameter combination, we find the unit conversion for the lattice coordinate $a$ and the density $b$ that minimizes the linear (e.g., firm densities) or logarithmic (e.g., genetic and citation curves) squared distance between the data and the model as visible in Fig. 4. We calculate the model density using the procedure described in SI Appendix, B.4 and use the steady state lattice length $L$ calculated from SI Appendix, Eq. S24.

In order to calculate the distance, we must decide on how to align model lattice coordinates with data coordinates. It is not always clear whether the leftmost or the rightmost data coordinates correspond to innovation and obsolescence fronts. This problem is especially vexing in the context of firms, where one can have in principle any real-valued positive or negative cost efficiency or productivity (such as after accounting for subsidies or other external costs). In viral mutations, this question arises with establishing the "origin" from which the virus descends, which in principle could be traced back to the origins of life. For citations, the definition of the origin is seemingly straightforward since the innovation front cannot precede the publication of the paper or patent, but again, there are some practical considerations that muddy the boundary. Papers might have been cited in the first year of publication, while in preprint form, or by the authors before publication in any form. In all of these cases, we take the simplest mapping from the innovation or obsolescence front to the data.

With firms, we take the leftmost point to correspond to the innovation front but allow the distance minimization procedure to choose the optimal location of the obsolescence front given by the scales $a$ and $b$. For SARS-CoV-2 mutations, we assume that the origin is the first detected hCoV-19/Wuhan/Hu-1/2019 strain. Since citations can occur as soon as a publication or patent filing appears, it would be natural to set the innovation front at the first year. We do so for patents. With the PRB citations, however, we find the optimal fit with this assumption returns unusually large parameters that are difficult to handle with our grid search algorithm. Instead, by assuming that the innovation front corresponds to the year following publication such that the first year has density of zero in our model, we find similarly good fits except that the parameters are well bounded. Thus, we rely on this more controlled procedure.

A secondary question is how to map the spacing of our discrete lattice to the units of the data. Since the solution corresponds to discrete values of a
To approximate the model occupancy densities, we rely on the second-order calculation of lattice length described in SI Appendix, A and calculate the density with the flow mean-field calculation in SI Appendix, B.

Beyond the steady state densities, our model also makes temporal predictions about the evolution of agent density (SI Appendix, B.4), but this is an aspect we do not focus on because of limitations on fine-grained temporal data.

**Data, Materials, and Software Availability.** The code for processing the data, calculating the model, and producing all the plots as described below is available at https://github/elotropmetro/innovation (81).

**ACKNOWLEDGMENTS.** E.D.L thanks Frank Neffeke, Anjali Bhatt, Alan Kwan, and the Complexity Science Hub Vienna (CSH) theory group including Yuan Pham, Jan Korbel, Stefan Thurner, Ernesto Ortega, and Rudi Hanel (who suggested the term “innovation train”) for inspiring discussions at various stages of ideation. E.D.L also thanks Vito Servedio and Snehal Shekatkar for discussion on obsolescence. E.D.L acknowledges funding from Bundesministerium für Bildung, Wissenschaft und Forschung (BMBWF), Hochschulraum-Strukturmittel (HRSF) 2016 (Complexity Science Hub Vienna). We thank the American Physical Society for access to journal citation data. We thank Jeff Barrick for help with the E. Coli phylogenetic data and Amélie Desvar-Larrive for COVID-19 references.

1. J. Schumpeter, *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*, Economics Third World Studies (Transaction Books, 1983).
2. H. Spencer, “The principle of evolution” in Nineteenth Century Online (NCCO): Science, Technology, and Medicine: 1780–1925 (Williams and Norgate, No. Bd. 1, 1864).
3. S. Spielrein, *Destaction as the cause of coming into being*. J. Anal. Psychol. 39, 155–186 (1994).
4. S. A. Kaufmann, Investigations (Oxford University Press, Oxford, 2000).
5. D. J. Teese, R. Rumsfeld, G. Doi, S. Winter, *Understanding corporate coherence*. J. Econ. Behav. Org. 23 (1), 1–30 (1994).
6. C. A. Hidalgo, B. Klinger, A. L. Barabási, R. Hausmann, The household space conditions the development of nations. Science 317 (5837), 482–487 (2007).
7. F. Neffeke, M. Henning, R. Bouchama, How do regions diversify over time? Industry relatedness and the development of new growth paths in regions Econ. Geogr. 87, 237–265 (2011).
8. S. A. Kaufman, *Forword: Stable and non-prestigious fitness landscapes* in Recent Advances in the Theory and Application of Fitness Landscapes, Emergence in Complexity and Computation, H. Richter, A. Engelbrecht, Eds. (Springer, Heidelberg, 2014), vol. 6, pp. vii–vii.
9. R. Franke, Wave trains, innovation noise, and long waves. J. Econ. Behav. Org. 45, 49–68 (2001).
10. R. Andersgård, F. Nardin, M. Ricottelli, Innovation waves, self-organized criticality and technological convergence. J. Econ. Behav. Org. 61, 710–728 (2006).
11. K. Imai, *Schumpeterian dynamics*. J. Econ. Behav. Org. 5, 159–190 (1984).
12. K. Imai, Schumpeterian dynamics Part II. J. Econ. Behav. Org. 5, 321–351 (1984).
13. P. Aghion, P. Howitt, “A model of growth through creative destruction” (National Bureau of Economic Research Cambridge, MA, Tech. Rep. 322, 1990).
14. S. Valverde, R. V. Solé, M. A. Bedau, Packard, Topology and evolution of technology innovation networks. Phys. Rev. X 76, 054118 (2007).
15. S. Valverde, R. V. Solé, Punctuated equilibrium in the large-scale evolution of programming languages. J. R. Soc. Interface. 12, 20150249 (2015).
16. S. Fortunato et al., *Science of science*. Science 339, eaao1858 (2015).
17. O. Kolody, N. Creanza, M. W. Feldman, Game-changing innovations: How culture can change the parameters of its own evolution and induce abrupt cultural shifts. PLoS Comput. Biol. 12, e1005302 (2016).
18. R. Amato, L. Lacasa, A. Díaz-Guilera, A. Baronchelli, The dynamics of norm change in the cultural development of new growth paths in regions Econ. Geogr. 87, 237–265 (2011).
19. E. D. L. thanks Frank Neffeke, Anjali Bhatt, Alan Kwan, and Rudi Hanel (who suggested the term “innovation train”) for inspiring discussions at various stages of ideation. E.D.L also thanks Vito Servedio and Snehal Shekatkar for discussion on obsolescence. E.D.L acknowledges funding from Bundesministerium für Bildung, Wissenschaft und Forschung (BMBWF), Hochschulraum-Strukturmittel (HRSF) 2016 (Complexity Science Hub Vienna). We thank the American Physical Society for access to journal citation data. We thank Jeff Barrick for help with the E. Coli phylogenetic data and Amélie Desvar-Larrive for COVID-19 references.
60. F. Tria, V. Loreto, V. D. P. Servedio, S. H. Strogatz, The dynamics of correlated novelties. *Sci. Rep.* **4**, 5890 (2015).
61. T. M. A. Fink, M. Reeves, How much can we influence the rate of innovation? *Sci. Adv.* **5**, eaat6107 (2019).
62. A. Sharma et al., Assembly theory explains and quantifies selection and evolution. *Nature* **622**, 321–328 (2023).
63. S. A. Kauffman, E. D. Weinberger, The NK model of rugged fitness landscapes and its application to maturation of the immune response. *J. Theor. Biol.* **141**, 211–245 (1989).
64. E. Ubaldi, R. Burioni, V. Loreto, F. Tria, Emergence and evolution of social networks through exploration of the Adjacent Possible space. *Commun. Phys.* **4**, 28 (2021).
65. W. Fontana, L. W. Buss, "The arrival of the fittest": Toward a theory of biological organization. *Bull. Math. Biol.* **56**, 1–64 (1994).
66. J. S. G. Chu, J. A. Evans, Slowed canonical progress in large fields of science. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2021636118 (2021).
67. U. Bhat, C. P. Kempes, J. D. Yeakel, Scaling the risk landscape drives optimal life-history strategies and the evolution of grazing. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 1580–1586 (2020).
68. J. J. Saarinen et al., Patterns of maximum body size evolution in Cenozoic land mammals: Eco-evolutionary processes and abiotic forcing. *Proc. R. Soc. B.* **281**, 20132049 (2014).
69. F. A. Smith, J. L. Betancourt, J. H. Brown, Evolution of body size in the woodrat over the past 25,000 years of climate change. *Science* **270**, 2012–2014 (1995).
70. P. Aghion, C. Antonin, S. Bunel, *The Power of Creative Destruction: Economic Upheaval and the Wealth of Nations* (Harvard University Press, 2021).
71. M. Nowak, R. A. May, *Viral Dynamics: Mathematical Principles of Immunology and Virology* (Oxford University Press, Oxford, 2000).
72. G. Silverberg, B. Verspagen, "An evolutionary model of long term cyclical variations of catching up and falling behind. IIAA Working Paper" (WP-95-009, IIAA, Luxenburg, Austria, 1995).
73. R. A. Rohde, R. A. Muller, Cycles in fossil diversity. *Nature* **434**, 208–210 (2005).
74. G. Silverberg, B. Verspagen, An evolutionary model of long term cyclical variations of catching up and falling behind. *Catching Up Falling Behind* **24** (1995).
75. S. Kumbhakar, C. A. K. Lovell, *Stochastic Frontier Analysis* (Cambridge University Press, Cambridge, 2004).
76. D. Badunenko, S. C. Kumbhakar, A. Lozano-Vivas, Achieving a sustainable cost-efficient business model in banking: The case of European commercial banks. *Eur. J. Operat. Res.* **293**, 773–785 (2021).
77. S. Alvarez, Tesla’s vehicle manufacturing cost in 2017 was $84k per car – it has since dropped to $36k (2022).
78. Tesla, Q4 and FY2021 Update, Tesla Financial Results (2022).
79. PatentsView, PatentsView (2022). https://patentsview.org.
80. K. J. Boudreau, E. C. Guinan, K. R. Lakhani, C. Riedl, Looking across and looking beyond the knowledge Frontier: Intellectual distance, novelty, and resource allocation in science. *Manage. Sci.* **62**, 2765–2783 (2016).
81. E. Lee, Idea engines. GitHub. https://github.com/eltrompetero/innovation. Deposited 17 January 2024.