

# Dynamics of rapid innovation

T. M. A. Fink<sup>\*†</sup>, M. Reeves<sup>‡</sup>, R. Palma<sup>‡</sup> and R. S. Farr<sup>†</sup>

<sup>†</sup>London Institute for, Mayfair, London W1K 2XF, UK, lims.ac.uk

<sup>\*</sup>Centre National de la Recherche Scientifique, Paris, France

<sup>‡</sup>BCG Henderson Institute, The Boston Consulting Group, New York, USA

**Abstract.** We introduce a model of innovation in which products are composed of components and new components are adopted one at a time. We show that the number of products we can make now gives a distorted view of the number we can make in the future: the more complex a product is, the more it gets under-represented. From this complexity discount we derive a strategy for increasing the rate of innovation by choosing components on the basis of long-term growth rather than just short-term gain. We test our model on data from language, gastronomy and technology and predict the best strategy for innovating in each.

Innovation is the process by which organizations improve and adapt to changes in the environment. Innovation is to organizations what evolution is to organisms [1]. Institutions that fail to innovate fall behind their competitors and succumb to environmental changes. The need to improve and adapt is amplified by a flat economy because companies and governments must innovate their way to growth rather than merely participate in it.

The rate of innovation for any organization is determined by the choices it makes. Firm managers, research leaders and policymakers regularly face difficult innovation decisions. Their choices can have long-term impacts, because the outputs of companies, universities and governments are determined by the skills and materials they have access to.

Despite the importance of innovation, how to innovate remains elusive [2]. There is a perennial tension between a managerial school—which sees innovation as a predictable process [3]—and a visionary school—which attributes innovation to extraordinary individuals and serendipity [4]. At the same time, research studies of innovation can mostly be divided into qualitative investigations of real data [5] or quantitative studies of theoretical models [6–8]. Connecting theory to application is hard because the global insights from theoretical models are typically lost in the noise of real systems, which necessarily comprise a large design space, only a fraction of which are of value.

In this paper, we give a mathematical foundation for innovation which serves as a common framework for its prescriptive and serendipitous aspects. We test our model on real data from language, gastronomy and technology and use it to make predictions about these domains. We begin from first principles—making no assumptions about the domain or design matrices—by considering how many designs we can make from a set of building blocks, and how this grows as we introduce new species of blocks to the set. In particular, we show that by choosing to adopt the right kinds of blocks, we can increase the growth rate of our design space, and identify a spectrum of strategies for balancing short-term gains and long-term growth.

## Our model

Suppose that we possess a number of distinct components, which we can combine in different ways to make products. A component can be a material object, like a transistor, or a skill, like 3D printing. We have more than enough of each component for our needs, so we do not have to worry about running out. Any subset of our components can be combined, but a combination either is, or is not, a potential product, according to some universal recipe book of products. Suppose further that there are a total of  $N$  possible components in ‘God’s own cupboard’, but that at a given time we only possess  $n$  of these  $N$  building blocks. Figure 1 shows examples of products and components.

Let  $p(n)$  be the number of products we can make from our  $n$  components [16]; as  $n$  approaches  $N$ ,  $p(n)$  approaches  $p(N)$ , the number of products in the universal recipe book. The size  $s$  of a product is the number of distinct components it is made of;

multiple occurrences of a component count once, e.g., ‘banana’ (Figure 1) has size  $s = 3$ , not 6. To be able to make a product of size  $s$ , we must possess all  $s$  of its components. The number of makeable products of size  $s$  is  $p(n, s)$ , so that summing  $p(n, s)$  over  $s$  gives  $p(n)$ . For example, from the letters a–d we can make  $p(4) = 9$  words: *a, ad, add, baa, bad, cab, cad, dab* and *dad*. Broken down by size, we find  $p_1 = 1$ ,  $p_2 = 4$  and  $p_3 = 4$ .

## Rate of innovation

The size of our product space  $p(n)$  depends on our choice of the  $n$  components we possess. In this sense innovation is the acquisition of components that enable the creation of new products. The innovation *rate* is the growth rate of our product space.

Before trying to increase the innovation rate, we calculate it for components adopted in an arbitrary order. We do so by averaging over all choices of our  $n$  components from the  $N$  possible, which gives the expected number of makeable products  $\bar{p}(n)$ . The number of combinations of  $s$  distinct components is  $\binom{n}{s}$ ; of these,  $p(N)$  are products. Thus the probability that an arbitrary combination of  $s$  components is a product is  $p(N)/\binom{N}{s}$ . We prove (Supp. Info. A) that the expected value of this quantity is invariant over all stages of the innovation process:

$$\bar{p}(n, s)/\binom{n}{s} = \bar{p}(n', s)/\binom{n'}{s}, \quad (1)$$

where we have  $n$  components at stage  $n$ , so that the ultimate size of our space need never be considered. When the number of components is big compared to the product size ( $n, n' \gg s$ ), eq. (1) can be written  $\bar{p}(n, s)/\bar{p}(n', s) \simeq (n/n')^s$ . This turns out to be very useful for calculating innovation rates.

The invariance of  $\bar{p}(n, s)/\binom{n}{s}$  tells us that the number of products we can make at stage  $n$  gives a distorted view of the number we can make at stage  $n'$ : the more complex a product is, the more it gets under-represented for  $n' > n$ , and over-represented for  $n' < n$ . This *complexity discount*,  $\binom{n'}{s}/\binom{n}{s}$ , arises because we are much more likely to possess the few components in a simple product, than the many components in a complex product. For example, if a child knows only half the letters in the alphabet, he is unlikely to write either ‘banana’ or ‘orange’. But he’s a lot more likely to write the former,  $\binom{13}{3}/\binom{26}{3} = 11\%$ , than the



FIG. 1: Products and components from language, gastronomy and technology. Left. Words and the letters used to make them. Right. Recipes and the ingredients used to make them. Bottom. Software products and the development tools used to make them.

latter,  $\binom{13}{6}/\binom{26}{6} = 0.7\%$ : ‘banana’, made of three different components, is a simpler word than ‘orange’, made of six.

Suppose we know the number and size of products we can make at stage  $n$ , but not some other stage  $n'$ . We can predict the size of our product space at  $n'$  from information we have at  $n$  by correcting for the complexity discount. Summing eq. (1) over  $s$  (Supp. Info C) shows us how:

$$p(n') \simeq \sum_{s=1}^n p(n, s) \binom{n'}{s} / \binom{n}{s}. \quad (2)$$

For example, from the size distribution alone of the  $p(6) = 45$  words makeable from the letters a–f, we can estimate the total number of English words,  $\ln p(26)$ , to within 4%.

A striking consequence of eq. (2) (Supp. Info. C) is that the expected size of our product space  $\bar{p}(n)$  depends only on the number and size of products, regardless of which components are used to make each one. If we assume a specific size distribution, we can calculate  $p(n')$  explicitly. We evaluate three typical distributions, all with identical mean size  $\bar{s}$ : constant, binomial, and Poisson; Figure 2, bottom shows examples of each. We find (Supp. Info. D) that the number of products we can make at stage  $n'$  can be expressed solely in terms of  $n'/n$  and mean size  $\bar{s}$  as shown in Figure 2, top. For example, for  $\bar{s} = 10$  and  $n'/n = 3$ , for a Poisson distribution of size the product space is 460 times bigger than that for a binomial distribution, which is 18 times that for constant size. Our space of products grows much faster for distributions with many simple products, even though the distributions all have the same mean. This is a consequence of the complexity discount: the number of both simple and complex products we can make grows super-linearly with  $n$ , but it grows faster for complex products than for simple ones.

### Strategy for increasing the rate of innovation

We can increase the innovation rate by using information from the unfolding innovation process to guide the choices we make. To see how, consider that adopting a new component offers two benefits: it is the missing link for products for which we *already have* all the other components; and it is the missing link for products for which we *will have* all the other components. To quantify these benefits, we characterise components in two

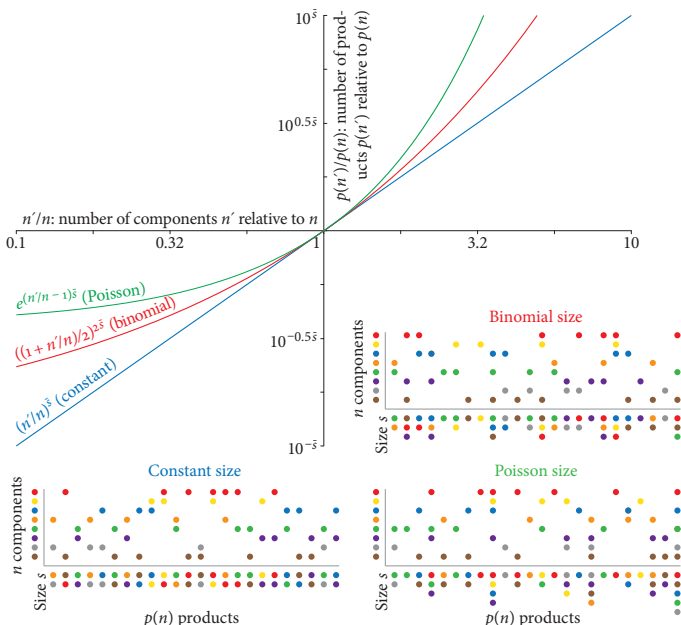


FIG. 2: Top. Universal rates of innovation for three distributions of product size, all with the same mean size  $\bar{s}$ . The distribution alone determines the expected innovation rate: those with many simple products yield the fastest innovation. Bottom. Examples of 24 products made from 8 components for each of the distributions, all with  $\bar{s} = 2$ .

ways. The *prevalence*  $p_\alpha(n)$  of some component  $\alpha$  is the number of products it appears in at stage  $n$ . The *valence*  $\bar{s}_\alpha(n)$  of  $\alpha$  is the average size of the products it appears in at stage  $n$ . For highly valent components to be useful, we need a lot of the other components they co-appear with. On the other hand, components that belong to simpler products are less reliant on others and more likely to boost our product space straightaway.

The products containing  $\alpha$  can be grouped together by their size  $s$ . Let  $p_\alpha(n, s)$  be the number of products we can make of size  $s$  that contain  $\alpha$ . The probability that an arbitrary combination of  $s$  components is a product and contains  $\alpha$  is  $p_\alpha(N)/\binom{N}{s}$ . We prove (Supp. Info. B) that the expected value of this quantity is invariant over all stages of the innovation process:

$$\bar{p}_\alpha(n, s)/\binom{n}{s} = \bar{p}_\alpha(n', s)/\binom{n'}{s}. \quad (3)$$

The invariance of  $\bar{p}_\alpha(n, s)/\binom{n}{s}$  tells us that the prevalence of a component at stage  $n$  gives a distorted view of its prevalence at stage  $n'$ ; the more valent a component is, the more it gets under-represented for  $n' > n$ , and over-represented for  $n' < n$ . Again this is a consequence of the complexity discount. Summing eq. (3) over  $s$  gives an estimate of the prevalence of  $\alpha$  at  $n'$  from information we have at  $n$ :

$$p_\alpha(n') \simeq \sum_{s=1}^N p_\alpha(n, s) \binom{n'}{s} / \binom{n}{s}. \quad (4)$$

Eq. (4) suggests a spectrum of strategies to alter the rate of innovation, contingent on how far into the future we are willing to focus. An *impatient* strategy considers what a potential new component can do for us *now*; it orders components by the current prevalence  $p_\alpha(n)$ . A *far-sighted* strategy, on the other hand, considers what a potential new component can do for us *later*; the further into the future we focus, the more it orders components by valence  $\bar{s}_\alpha(n)$ . In between these two extremes, a successful innovator will balance prevalence and valence in choosing the components he adopts, according to his resources and expected return on investment.

### Language, gastronomy and technology

To test our model, we analysed data from three domains: language, gastronomy and technology (Figure 1, Supp. Info. F). In language, the products are the 39,915 common English words and the components are the 26 letters in the alphabet. In gastronomy, the products are 56,498 recipes from the databases

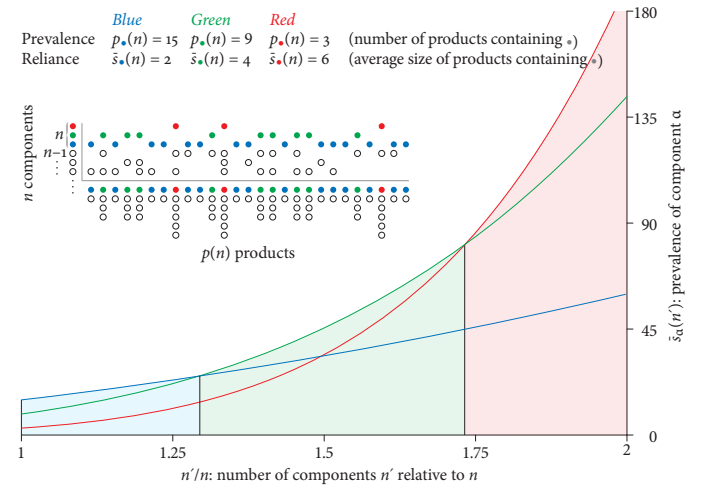


FIG. 3: Suppose we possess  $n - 1$  components, and must choose as the  $n$ th blue, green or red. We can test each option, one at a time, by checking what new products we can make with it and our  $n - 1$  other components. Blue maximises the size of our product space immediately after adopting the  $n$ th component ( $n'/n = 1$ ). But depending on how far we look ahead, other options are optimal, as eq. (4) shows.

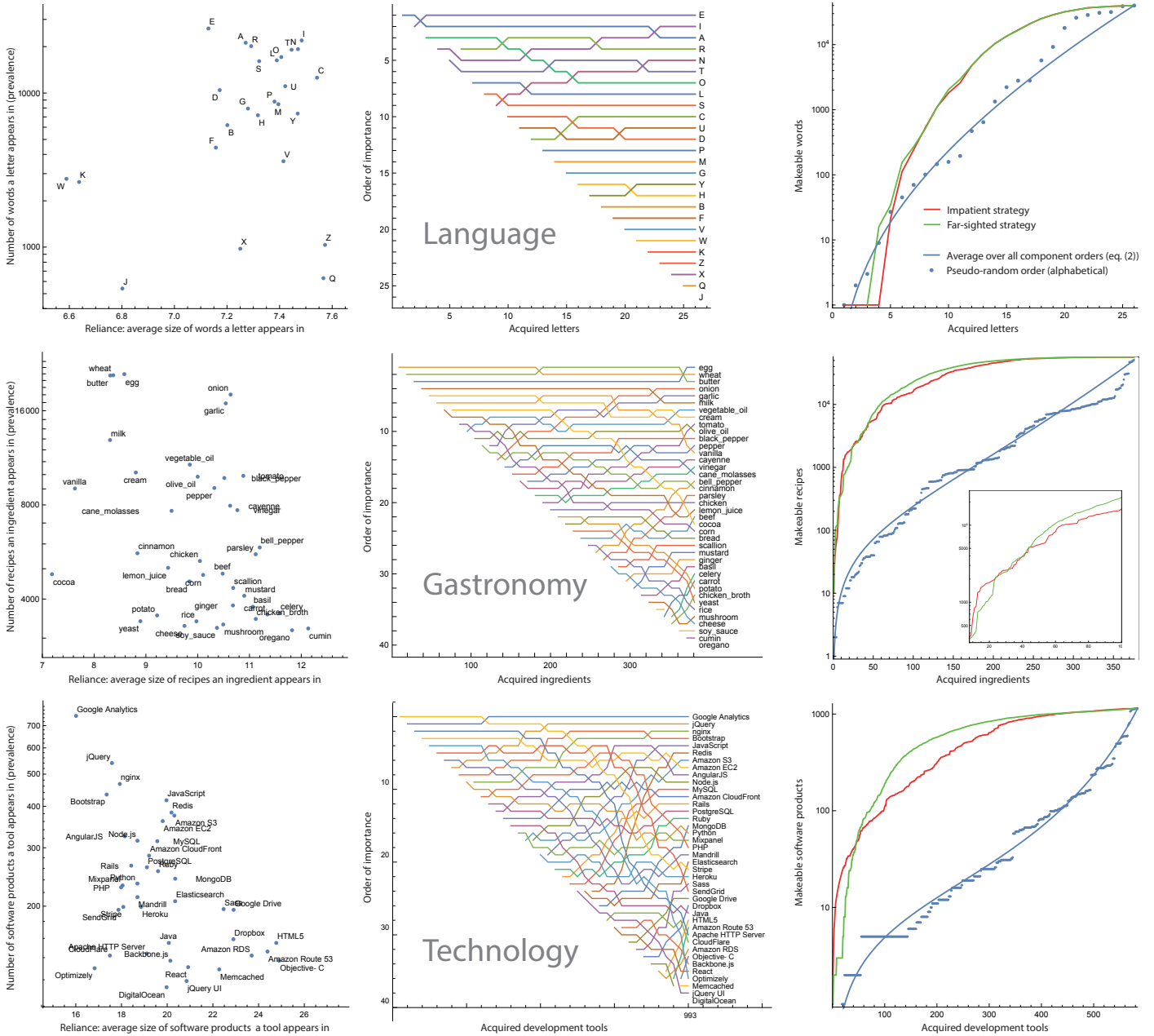


FIG. 4: Top to bottom: we analysed data from language (39,915 words made from 26 letters); gastronomy (56,498 recipes made from 381 ingredients); and technology (1158 software products made from 993 development tools). Left column: scatter plots of component prevalence versus valence (for gastronomy and technology, we only show the top 40). Center column: the relative prevalence of the components at each stage of the innovation process. Right column: the size of the product space for a pseudo-random (alphabetical) component ordering (blue dots) and averaged over all possible orderings, as given by eq. (2) (blue lines); the impatient strategy of choosing the most prevalent components at time  $n$  ranked by  $p(n)$  (red lines); and the far-sighted strategy of choosing the most prevalent at time  $n$  ranked by  $p(N)$  (green lines).

allrecipes.com, epicurious.com, and menupan.com [13] and the components are 381 ingredients. In technology, the products are 1158 software products catalogued by StackShare.io and the components are 993 development tools used to make them.

For each component in each domain, we calculated the prevalence  $p_\alpha(N)$  and valence  $\bar{s}_\alpha(N)$ , which we show as scatter plots in Figure 4, left. For clarity, we only show the top 40 in gastronomy and technology. The three domains show different levels of variation in valence of the data shown: in language, it ranges from 6.6 (w) to 7.6 (z); in gastronomy, from 7.2 (cocoa) to 12.2 (cumin); in technology, from 16 (Google Analytics) to 25 (Objective C). If we pick two arbitrary points in each scatter plot, the typical slope of the line joining them is 3.9, 0.38 and 0.16 for the three domains. These slopes are important because the key parameter in determining the crossover point of the value

of two components  $\alpha$  and  $\beta$  (Sup. Info. F) is

$$z(n) = -\frac{\ln p_\beta(n) - \ln p_\alpha(n)}{\bar{s}_\beta(n) - \bar{s}_\alpha(n)};$$

see Figure 3 for examples of crossovers for constant valence. This is just the negative of the slope of the line joining them. For the roughly Poisson distributed valence in the three domains, the crossovers for language will occur 92% before or after the  $x$  domain shown; language  $xx\%$  outside the domain, and technology  $xx\%$ . We verify this below by checking the actual crossovers.

In Figure 4, center, we show ‘bumps charts’ for the different components in each domain: the relative prevalence at each stage of the innovation process. Crossings in these rankings mean that the order of prevalence of components at one stage is not the same as the order at another stage. There are few crossings in language, a moderate number in gastronomy and many in technology. Consider the top 13, 20 and 20 compo-

nents halfway across the bumps charts. How do they compare to the top 13, 20 and 20 components at the right of the bumps charts? For language, the 13 components (e i a r n t o l s c u d versus e a r i o t n l s d u c) are 100% identical, with some swaps in ordering; for gastronomy, the top components overlap 80%; for technology, 40%. Crucially, it is the presence of crossovers which make possible a far-sighted strategy. In other words, crossovers permit an advantaged innovation strategy which is contingent on the stage of the game and information from the unfolding innovation process.

In Figure 4, right, we show the number of makeable products for a random component discovery order (blue dots), and our prediction in eq. (2) for random discovery order (blue line). We also show the impatient strategy (red line) and the far-sighted strategy (green line). For language, these two strategies cannot be very different, because there are so few crossings in the bumps charts. For gastronomy, the curves diverge moderately: impatience has a two-fold advantage over far-sightedness at first, but then the roles reverse and far-sightedness has a two-fold advantage. For technology, the curves diverge significantly: impatience has a five-fold advantage, but again the roles reverse and far-sightedness wins by a factor of five.

## Discussion

Key to this work is that the most important components when a system is small or less mature may not be the same as when it is big or more mature. This spectrum of how much priorities depend on maturity characterizes any system in which constituent building blocks can be combined in a multitude of ways. As we have shown, the most valuable letters for playing Scrabble (seven letters) are essentially the same as in everyday English (26 letters); the most valuable ingredients in a small kitchen (10 ingredients) are somewhat different from those in a large one (40 ingredients); the most valuable development skills for a small software company (experience with 20 tools) are significantly different from those for a more mature one (80 tools). Failure to recognise that an organization's priorities can depend on its size or maturity is a source of disagreement and confusion in determining which innovation choices it should make [1].

In the *Theory of Economic Development*, the economist Joseph Schumpeter discriminates between invention and innovation. An invention is the introduction of something new. An innovation is the introduction of something new which because of its diverse value has significant uptake. Thus only some inventions go on to become innovations. This may happen straightaway, when the invention immediately enables the creation of new products; or it may be delayed, when uptake of the invention depends on additional events. Our model naturally captures this distinction: The introduction of tomato as a Western ingredient is an innovation, because of its use in many recipes. The introduction of Google Maps as a development tool is an innovation, because of its use in many software products.

Writing about the *The Three Princes of Serendip*, the 18th-century historian Horace Walpole records that the princes "were always making discoveries, by accidents and sagacity, of things they were not in quest of". Serendipity is the fortunate development of events. "One thing leads to another" is a familiar sentiment to many [14], and many organizations deploying far-sighted innovation, such as Google and Zappos, stress the importance of serendipity. Our model helps us see why. Components adopted by far-sighted innovators are of little immediate benefit. However, as the innovation process unfolds and the far-sighted strategy pays off, the results will seem serendipitous, because a number of previously low value components become invaluable. Thus, what appears as serendipity is not happenstance but the confluence of components previously adopted according to a far-sighted innovation strategy.

This work offers several lessons for organisations looking to innovate faster. (i) Recognise the value of simple designs. On av-

erage the likelihood of being able to make a product decreases super-exponentially with the number of components it contains. Organizations operating in a domain with even a small fraction of simple products will have access to a bigger repertoire of products, faster, than their competitors. This gives credence to start-ups searching for product-market fit in the form of a 'minimum viable product'. (ii) Determine if your priorities depend on your maturity. If not, then an impatient and far-sighted strategy are one-and-the-same, and imitating more successful mature organisations can be an appropriate strategy. If so, then a strategic balance of impatience and far-sightedness is required, as follows. (iii) Correct for the complexity discount. On the basis of equation (4), employ a far-sighted strategy in making innovation decisions to the extent permitted by your resource constraints. This will give slow growth at first, but a disproportionate windfall later on. (iv) Engineer a disruption of the innovation space. Although we didn't study disruption here, our model can be naturally extended to incorporate it. At its heart, disruption is simplification. Organisations looking to up-end their space can either attempt to make specific components obsolete—as when Warby Parker disintermediated much of the eyewear value chain. Or they can merge several components into a single component module to make products simpler—as infrastructure-as-a-service companies have done by replacing the need to own, configure, and maintain physical computers.

- [1] D. H. Erwin and D. C. Krakauer, 'Insights into innovation', *Science* **304**, 1117 (2004).
- [2] M. Reeves, K. Haanaes and J. Sinha, *Your Strategy Needs a Strategy* (Harvard Business Review Press, 2015).
- [3] P. Drucker, 'The discipline of innovation', *Harvard Business Review* **8**, 1 (2002).
- [4] W. Isaacson, *The Innovators: How a Group of Hackers, Geniuses, and Geeks Created the Digital Revolution*, (2014).
- [5] C. Weiss et al., 'Adoption of a high-impact innovation in a homogeneous population', *Phys Rev X*, **4**, 041008 (2014).
- [6] J. McNerney, D. Farmer, S. Redner, J. Trancik, 'Role of design complexity in technology improvement', *Proc Natl Acad Sci*, **108**, 9008 (2011).
- [7] V. Sood et al., 'Interacting branching process as a simple model of innovation', *Phys. Rev. Lett.*, **105**, 178701 (2010).
- [8] T. Guede, A. Dobrinevski and J.-P. Bouchaud, 'Explore or exploit? A generic model and an exactly solvable case', *Phys. Rev. Lett.*, **112**, 050602 (2014).
- [9] R. Van Noorden, 'Physicists make 'weather forecasts' for economies', *Nature*, **1038**, 16963 (2015).
- [10] A. Tacchella et al., 'A new metric for countries' fitness and products' complexity', *Scientific Reports*, **2**, 723 (2012).
- [11] J. A. Schumpeter, *The Theory of Economic Development*, R. Opie, Transl. (Harvard Univ. Press, Cambridge, MA, 1926).
- [12] A. Wagner, 'Robustness and evolvability: a paradox resolved', *Proc Roy Soc B* **91**, 275 (2008).
- [13] Y.-Y. Ahn, S. E. Ahnert, J. P. Bagrow and A.-L. Barabási, 'Flavor network and the principles of food pairing', *Scientific Reports* **1**, 196 (2011).
- [14] F. Triana, V. Loreto, V. D. P. Servedio and S. H. Strogatz, 'The dynamics of correlated novelties', *Scientific Reports* **4**, 5890 (2014).
- [15] We make no assumptions about the values associated with specific products, which will depend on the market environment. But we can be sure that maximising the number of products is a proxy for maximising any reasonable property of them. A similar proxy is used in evolutionary models, where evolvability is defined to be the number of new phenotypes in the adjacent possible (1-mutation boundary) of a given phenotype [12]. Some of the new phenotypes will have higher fitness, some lower, depending on the environment and the fitness of the phenotype in hand.

## Supplementary information (online only)

### A. Proof of products invariant

In proving this invariant and the one below, we posit the existence of  $N$ , the total number of components ‘in God’s own cupboard’, but make no reference to this inaccessible quantity in our actual results.

Let  $\mathcal{N}$  be the set of  $N$  possible components, let  $n$  be a subset of  $n$  components chosen from  $\mathcal{N}$ , and let  $s$  be a subset of  $s$  components chosen from  $n$ . The number of products of size  $s$  we can make from  $n$  components is

$$p(n, s) = \sum_{s \subseteq n} \text{prod}(s),$$

where  $\text{prod}(s)$  take the value 1 if the combination of components  $s$  forms a product and 0 otherwise. The expected number of products we can make,  $\bar{p}_s(n)$ , is the average of  $p(n, s)$  over all subsets  $n \subseteq \mathcal{N}$ ; there are  $\binom{N}{n}$  such subsets. Therefore

$$\begin{aligned} \bar{p}(n, s) &= 1/\binom{N}{n} \sum_{n \subseteq \mathcal{N}} p(n, s) \\ &= 1/\binom{N}{n} \sum_{n \subseteq \mathcal{N}} \sum_{s \subseteq n} \text{prod}(s). \end{aligned} \quad (5)$$

Consider some particular combination of components  $s'$ . The double sum above will count  $s'$  once if  $s = n$ , but multiple times if  $s < n$ , because  $s'$  will belong to multiple sets  $n$ . How many? In any set  $n$  that contains  $s'$ , there are  $n - s$  free elements to choose, from  $N - s$  other components. Therefore eq. (5) will count every combination  $s$  a total of  $\binom{N-s}{n-s}$  times, and

$$\begin{aligned} \bar{p}_s(n) &= 1/\binom{N}{n} \binom{N-s}{n-s} \sum_{s \subseteq \mathcal{N}} \text{prod}(s) \\ &= \binom{n}{s} / \binom{N}{s} p(N, s). \end{aligned} \quad (6)$$

The same must be true when we replace  $n$  by  $n'$ , and therefore

$$\bar{p}_s(n)/\binom{n}{s} = \bar{p}_s(n')/\binom{n'}{s}.$$

### B. Proof of components invariant

The prevalence  $p_\alpha(n, s)$  is the number of products of size  $s$  we can make from  $n$  components and that contain component  $\alpha$ . Using the same notation as before, and with  $\beta$  some component,

$$p_\alpha(n, s) = \sum_{s \subseteq n} \text{prod}(s) \sum_{\beta \subseteq s} \delta_{\alpha\beta},$$

where the Kronecker delta function  $\delta_{\alpha\beta} = 1$  if  $\alpha = \beta$  and 0 otherwise. The expected prevalence of component  $\alpha$ ,  $\bar{p}_\alpha(n, s)$ , is the average of  $p_\alpha(n, s)$  over all subsets  $n \subseteq \mathcal{N}$ . Therefore

$$\begin{aligned} \bar{p}_\alpha(n, s) &= 1/\binom{N}{n} \sum_{n \subseteq \mathcal{N}} \sum_{s \subseteq n} \sum_{\beta \subseteq s} \text{prod}(s) \delta_{\alpha\beta} \\ &= 1/\binom{N}{n} \binom{N-s}{n-s} \sum_{s \subseteq \mathcal{N}} \sum_{\beta \subseteq s} \text{prod}(s) \delta_{\alpha\beta} \\ &= \binom{n}{s} / \binom{N}{s} p_\alpha(N, s). \end{aligned}$$

The same must be true when we replace  $n$  by  $n'$ , and therefore

$$\bar{p}_\alpha(n, s)/\binom{n}{s} = \bar{p}_\alpha(n', s)/\binom{n'}{s}.$$

### C. Expected innovation rate depends on size only

By definition,  $\bar{p}(n)$  is the mean of  $p(n)$ , and therefore  $p(n)$  is an unbiased estimator of  $\bar{p}(n)$ . How accurate it is depends on the details of the particular innovation domain; empirically, observations from our three domains (Figure 4, right) suggest it is reasonable.

Summing eq. (6) over  $s$ ,

$$\bar{p}(n) = \sum_{s=1}^n p(n, s) \binom{n}{s} / \binom{N}{s}.$$

Thus the number and size alone of the products in the universal recipe book alone determine the expected innovation rate.

### D. Rate of innovation for specific size distributions

Let  $f(s)$  be the probability distribution product size. Then

$$p(n, s) = p(n) f(s).$$

For a binomial size distribution,

$$f(s) = \binom{2\bar{s}}{s} \left(\frac{1}{2}\right)^{2\bar{s}},$$

and thus

$$p(n, s) = p(n) \binom{2\bar{s}}{s} \left(\frac{1}{2}\right)^{2\bar{s}}.$$

Substituting this into eq. (2) yields

$$\begin{aligned} p(n') &\simeq p(n) \sum_{s=0}^{2\bar{s}} \binom{2\bar{s}}{s} \left(\frac{1}{2}\right)^{2\bar{s}} \binom{n'}{s} / \binom{n}{s} \\ &\simeq p(n) \left(\frac{1}{2}\right)^{2\bar{s}} \sum_{s=0}^{2\bar{s}} \left(\frac{n'}{n}\right)^s \binom{2\bar{s}}{s} \\ &= p(n) (1/2)^{2\bar{s}} (1 + n'/n)^{2\bar{s}} \\ &= p(n) ((1 + n'/n)/2)^{2\bar{s}(n)}, \end{aligned}$$

For a Poisson size distribution,

$$f(s) = \bar{s}^s e^{-\bar{s}} / s!,$$

and thus

$$p(n, s) = p(n) \bar{s}^s e^{-\bar{s}} / s!,$$

Substituting this into eq. (2) yields

$$\begin{aligned} p(n') &\simeq p(n) \sum_{s=0}^{\infty} \bar{s}^s e^{-\bar{s}} / s! \binom{n'}{s} / \binom{n}{s} \\ &\simeq p(n) e^{-\bar{s}} \sum_{s=0}^{\infty} (\bar{s} \frac{n'}{n})^s / s! \\ &= p(n) e^{-\bar{s}} e^{\bar{s} n'/n} \\ &= p(n) e^{(n'/n-1)\bar{s}(n)}. \end{aligned}$$

Summarising our results,

$$p(n') \simeq p(n) \cdot \begin{cases} (n'/n)^{\bar{s}(n)}, & \text{constant size,} \\ ((1+n'/n)/2)^{2\bar{s}(n)}, & \text{binomial size,} \\ e^{(n'/n-1)\bar{s}(n)}, & \text{Poisson size.} \end{cases} \quad (7)$$

### E. Crossover of two components

Consider two components  $\alpha$  and  $\beta$  and assume  $p_\alpha(n) < p_\beta(n)$ . At what stage  $n_\times$  will their prevalences cross, such that  $p_\alpha(n') > p_\beta(n')$  for  $n' > n_\times$ ?

To answer this, we first calculate the expected prevalence  $p_\alpha(n')$  explicitly for specific distributions of the valence  $\bar{s}_\alpha$ , similar to how we did for  $p(n')$  above. We find

$$p_\alpha(n') \simeq p_\alpha(n) \cdot \begin{cases} (n'/n)^{\bar{s}_\alpha(n)}, & \text{constant valence,} \\ ((1+n'/n)/2)^{2\bar{s}_\alpha(n)}, & \text{binomial valence,} \\ e^{(n'/n-1)\bar{s}_\alpha(n)}, & \text{Poisson valence.} \end{cases} \quad (8)$$

To find the crossover, we evaluate eq. (8) for  $\alpha$  and  $\beta$  at  $n' = n_\times$  and set them equal. For example, for constant valence, we solve  $p_\alpha(n_\times/n)^{\bar{s}_\alpha} = p_\beta(n_\times/n)^{\bar{s}_\beta}$  for  $n_\times$ . Summarising the

results for the three distributions,

$$n_{\times} = n \cdot \begin{cases} e^{z(n)}, & \text{constant valence,} \\ 2e^{z(n)/2} - 1, & \text{binomial valence,} \\ 1 + z(n), & \text{Poisson valence,} \end{cases} \quad (9)$$

where the key parameter  $z(n)$  is the negative of the difference in the log of the prevalences over the difference in the valences:

$$z(n) = -\frac{\ln p_{\beta}(n) - \ln p_{\alpha}(n)}{\bar{s}_{\beta}(n) - \bar{s}_{\alpha}(n)}. \quad (10)$$

## F. Data

All records of innovation data are necessarily biased in two ways. First, they favor the trophy case of high-value components; the dustbin of ineffectual components (inventions rather than innovations, in Schumpeter’s terminology) have been largely forgotten. For gastronomy, there will have been countless unsuc-

cessful new ingredients or ingredient substitutes which provided little immediate or long-term benefit. They are not recorded in published cookbooks. Similarly for technology, unpopular or obsolete development tools are unlikely to be registered on the stackshare.io database or elsewhere.

Second, innovation records come to an abrupt end at the present moment in time, because we don’t know which new recipes, software or other novel products the future will bring. For this reason the results of our data analysis are most applicable away from the present-day boundary of 381 ingredients and 584 development tools. Innovation should be viewed as an infinite or large finite game rather than a finite one.

For our language data set, we used the `WordList[]` set of common English words from Mathematica 10. `WordList[]` has 40,127 words, from which we used the 39,915 that contain only the lowercase letters a–z; for instance, we removed all words with a hyphen.