The Innovation-Complexity Trade-off: How Bottlenecks Create Superstars and Constrain Growth

Seth G. Benzell and Erik Brynjolfsson

January 25, 2022

Abstract

We introduce a model of technological advances as allowing for greater productivity at the cost of increased complexity. Complex goods and services require a large number of strongly complementary inputs. Innovation increases the probability that an input will be flawed, leading to a skewed distribution of firm sizes, with more extreme superstars and a higher share of mediocre firms. Whether additional combinatoric innovation increases growth is determined by a trade-off between its contribution to productivity and effective complexity. We evaluate strategies to deal with complexity, including modularization, and show they are just as important for boosting long-term growth as innovation.

1 Background

A recent literature has pointed to, and attempted to explain, a 'modern productivity paradox'. The paradox is that there has been an explosion of exciting new technologies, including artificial intelligences which match or surpass human level performance in many domains, simultaneous with slowing productivity growth (Brynjolfsson et al., 2017). Researchers have variously attributed this incongruity to mismeasurement of the economic effects of the digital economy and to over-optimism about the importance of recent technologies [(Syverson, 2017), (Brynjolfsson et al., 2019), (Diewert and Fox, 1999)]. Perhaps the leading explanation is an adjustment friction that prevents some firms from adopting the most cutting-edge technologies, at least without making massive complementary intangible investments first (Brynjolfsson et al., 2020).

Another hypothesis is that new ideas are getting harder to find (Gold, 2021). Looking into data across several different industries, Bloom et al. (2020) argue there has been a decrease in the productivity of scientists in producing valuable new ideas. They show that U.S. agricultural R&D has nearly doubled since 1970 and semiconductor R&D has gone up by nearly 2000%, yet the TFP growth rate in agriculture has declined, and in semi-conductors has stagnated.

The idea that ideas are becoming harder to find is a subversive one for growth theory. Weitzman (1998), in his seminal analysis of the implications of idea generation for growth, shows that, under very general conditions, long-term economic growth is *not* limited by the quantity of new ideas. He starts from the premise that potential new ideas are combinations of old, already realized, ideas. If this is the case, then as new ideas are generated, the space of promising new ideas explodes combinatorically. The pace of economic improvement is limited by the willingness of households to save and invest, not by lack of ideas, which quickly become superabundant.

To reconcile these two perspectives, we develop a growth model where new ideas being recombinations of old ones guarantees their abundance, but introduces a complication. While Swiss watches and supercomputers are extremely powerful, creating them requires painstaiking precision. Even small errors can make the entire product useless, making improvements risky and even imitation hazardous. The difficulty in maintaining this precision has led semi-conductor manufacturers to 'clone' successful factories (on the off chance some seemingly incidental decision is essential), and watch manufacturers to modularize and standardize components.

Ours is not the first paper to emphasize the importance of complexity and complementarity in production. Hidalgo and Hausmann (2009) show that richer countries have more complex economies. And Kremer (1993) shows how tight complementarities between workers can drive growth and wage patterns.

2 Model

Every firm is based on an idea. As in Weitzman (1998), a new idea is a combination of previously mastered concepts or intermediate inputs. For a new idea to be discovered, developed, or implemented, each of its elements must be properly implemented. Because of this, each stage or element in the new process can be thought of as a bottleneck. Elements of an idea could correspond to a step in a production process, ingredients in a stew, or different specialists in a team.

Potential entrants are ex-ante uncertain about how successfully they can implement each component of a new idea. To capture the fact that precision is needed to implement complex ideas, the quality with which idea i is implemented, B_i, is as perfect as its most imperfectly implemented step $X_{i,m} \sim U[0,1]$. So,

$$B_i = min\{X_{i,1}, ..., X_{i,M(N_i)}\} \tag{1}$$

where $M(N_i)$ is the effective complexity, or number of fail points that need to be overcome, to implement an idea with N elements. In other words, the firm's productivity is a Leontieff function of it's M inputs.

However, there is an upside to increased complexity. While complex ideas are ones that require more elements to go correctly to be implemented properly, they can also reach heights not achievable through less complex means. We model this upside of complexity as an increased *maximum* quality of an idea with more elements. We denote by $P(N_i)$ the maximum possible final quality of an idea with N elements.

Therefore, the total realized quality of an idea is,

$$A_i = P(N_i)B_i(X_{i,1}, ..., X_{i,M(N_i)}) \tag{2}$$

and we assume that $P(N)$ and $M(N)$ are both weakly increasing in N. Firms face uncertainty before entering in how well they can overcome each idea fail-point $X_{i,m}$. They decide how much capital to rent *after* realizing their productivity draw. Each firm decides whether to enter in every period, paying a fixed cost F to do so.

To operate, a firm rents capital, which it uses in a decreasing returns to scale production technology. So each firm's ex-post profit, after realizing it's productivity draw B_i is

$$\pi_i = A_i \sqrt{K_i} - rK_i - F \tag{3}$$

where r is the interest rate and K_i is the amount of capital the firm decides to rent.

The economy as a whole is characterized by the following market clearing conditions: Total output is the combined net output of all I firms that enter with an idea

$$Y = \sum^{I} A_i \sqrt{K_i} - F = rK. \tag{4}$$

The capital each firm rents must sum to the to the total capital stock

$$K = \sum^{I} K_i. \tag{5}$$

In addition, there is a single inter-temporal market clearing condition. We have suppressed time subscripts until this point (and do so except in equations with an inter-temporal element), but they are indicated with t. The economy grows with capital appreciation. Capital does not depreciate and, by assumption, there is a constant saving rate.

$$K_{t+1} = sY_t + K_t. \tag{6}$$

All output is capital output, and (for a given level of complexity) the economic growth rate is constant and linear as in an AK model,

$$\frac{Y_{t+1}}{Y_t} = sr. \tag{7}$$

Firms enter, demanding capital and bidding up the interest rate, until the expected profit from entry is zero ($E[\pi_i] = 0$). We assume every firm enters with an idea of the same complexity N_i. Firms' desire to enter is increasing in their expected productivity, meaning that the interest rate, and therefore the growth rate, are functions of $P(N)$ and $M(N)$.

$$r = \left(1 - \sqrt{1 - 2F\frac{(M(N)+1)(M(N)+2)}{P(N)^2}}\right)^{-1} \tag{8}$$

so long as F is greater than 0, but small enough that the law of large numbers holds.[1]

Note that r is decreasing in N if $\frac{\partial M(N)}{N} \geq 1$ and $P(N)$ is constant. In other words, unless increased complexity sufficiently increases firms' maximum productivity, it will hold back growth.

3 Model Analysis

This model has implications both for the distribution of firm sizes and long run growth. The implications for both are tightly tied to expectations of $A(N)$ which in turn is determined by the rates at which maximum productivity and the count of fail points, $P(N)$ and $M(N)$, increase as complexity N increases.

3.1 Firm Size Distribution

How does the average, maximum and the distribution of firm sizes change as N grows?

The average firm must make zero profit in equilibrium. Therefore, average firm size, measured through firms' capital stock, is a function of the interest rate and the expected productivity of a startup

$$E[K_i] = \left(\frac{E[A_i]}{2r}\right)^2. \tag{12}$$

Whether an increase in expected productivity increases or decreases average firm size

[1]F must also not be so large that no firms enter. Equation 8 is derived by noting that

$$\pi_i = max(\pi_i|A_i)_{wrtK_i} = \frac{(A_i)^2}{2r} - F - \frac{(A_i)^2}{4r^2}, \tag{9}$$

then setting expected profit equal to zero

$$0 = E[\pi_i] = E[(A_i)^2](\frac{1}{2r} - \frac{1}{4r^2}) - F, \tag{10}$$

and finally by calculating expected productivity as a function of N

$$E[A_i^2] = \frac{2P(N)^2}{(M(N)+1)(M(N)+2)}. \tag{11}$$

depends on whether the economy is open or closed. In an open economy, with exogenous r, an increase in expected productivity causes firms to grow larger and move farther along their increasing marginal cost curve. In a closed economy, as we are focused on, r grows with equation 8 as expected productivity increases. The denominator grows at a faster rate than $E[A_i]$, meaning that an increase in expected productivity in a closed economy leads the average firm size to *decrease*. Intuitively, a more productive technology causes more firms to enter and bid up the interest rate, moving each firm back along its marginal cost curve and lowering average firm size.

Maximum firm size is only a function of the maximum potential productivity draw, $P(N)$, while median firm size is a function of the number of fail point $M(N)$ as well. So,

$$Max[K_i] = \frac{P(N)^2}{2r} \tag{13}$$

$$median[K_i] = \frac{median[A_i^2]}{4r^2} \tag{14}$$

$$median[A_i^2] = P^2 2^{\frac{-2}{M}} (2^{1/M} - 1)^2 \tag{15}$$

Whether an increase in complexity N increases or decreases expected productivity and the interest rate r, the **skewness**, i.e. the difference between average and median firm size, of the firm size distribution increases. This is the sense in which increased complexity creates a 'superstar firm' effect. Intuitively, this is because while an increase in P raises the maximum possible outcome, an increase in M makes perfection more difficult to achieve, and so median productivity will not grow as quickly (or may even shrink).

Figure 1 shows an example simulation of how the interest rate/growth rate, median, and average firm sizes evolve as complexity N increases, holding the stock of capital K constant. In this example, there are no productivity gains from increasing productivity (i.e. $P'(N) = 0$), but there is a gradual increase in the number of fail points $M(N) = N$. As can be seen, these assumptions generate an increase in average firm size as fewer firms enter because of their worsening expected productivity. However, median firm size grows more slowly, leading to an increasingly skewed distribution of firm sizes.

Has there been an increase in the skewness of US firm sizes? This is a main prediction

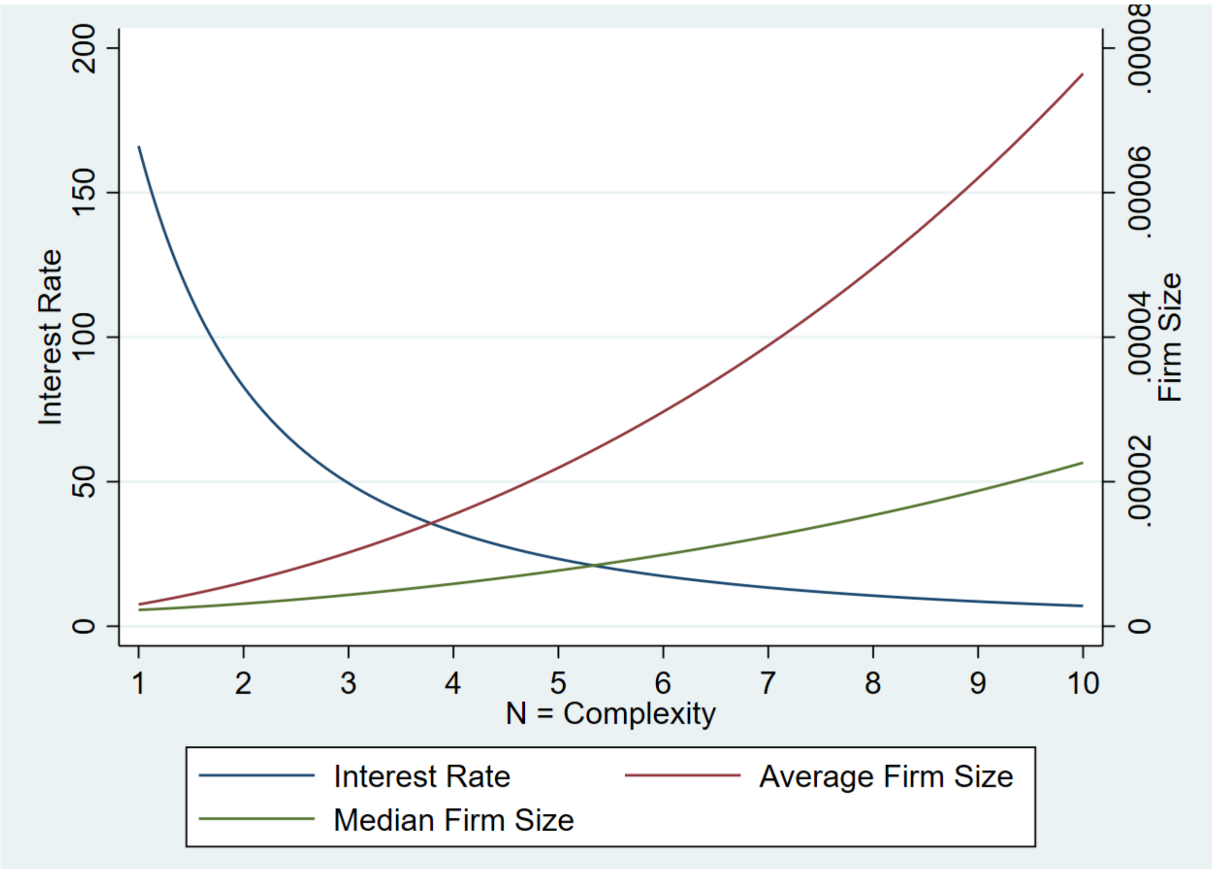

Figure 1: Simulated interest rate, average firm size, and median firm size for economies with different degrees of complexity N. $K = 1$; $F = .001$; $P(N) = 1$, $M(N) = N$

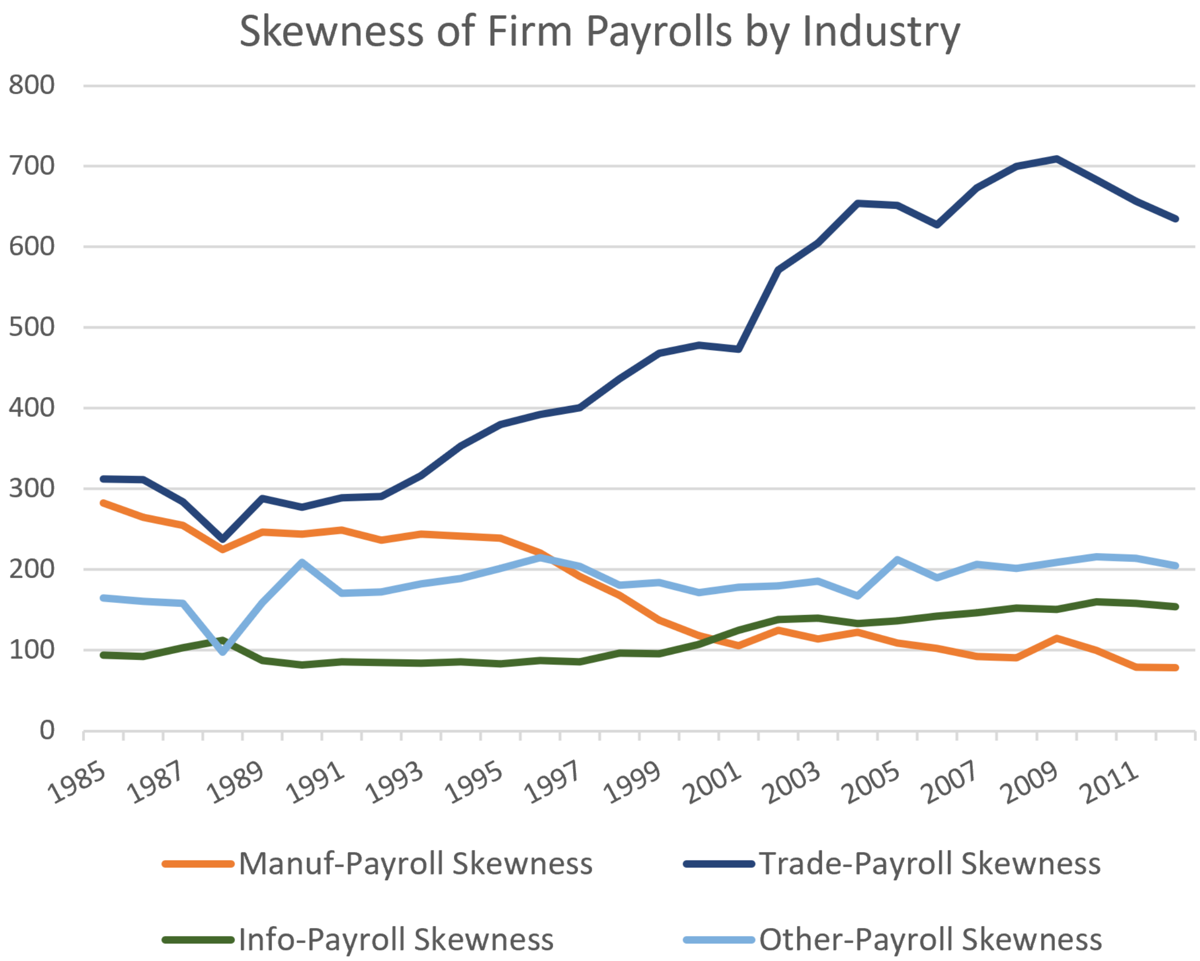

Figure 2: Skewness of firm payrolls by sector and year. Figure derived from restricted-use Census Longitudinal Business Database (LHD) microdata.

of our model, and it seems to hold water across industries. Increasing concentration of employment and sales in the most productive firms in most US industries since the early 1980's has been well documented in Autor et al. (2017). Figure 2 shows the skewness of payrolls over a similar interval for four sectors of the economy.[2] Each has experienced an increase in skewness except for manufacturing (an exception which is also noted in Autor et al. (2017)). The largest increase in skewness has come in the retail sector. This suggests that sector has seen the largest increase in effective complexity and number of fail-points.

[2]Thanks to Xiupeng Wang for providing the data underlying this figure.

4 Complexity and Long-Term Growth

At the macroeconomic level, our model is a variation on the canonical AK growth model. In this AK growth economy, growth is constant in percent terms for a constant r

$$g = Y_{t+1}/Y_t = sr \tag{16}$$

Therefore, understanding the direction of increasing complexity N on growth is equivalent to finding the sign of $\frac{\partial r}{\partial N}$. The interest rate r is monotonically decreasing in $\frac{(M(N)+1)(M(N)+2)}{P(N)^2}$. Therefore, for large M and P, we have the following (asymptotic, where M and P are large) condition for an increase in complexity N to increase the growth rate

$$\text{sgn}\left(\frac{\partial g}{\partial N}\right) = \text{sgn}\left(\frac{\frac{\partial P}{\partial N}}{P} - \frac{\frac{\partial M}{\partial N}}{M}\right) \tag{17}$$

What equation 17 tells us is that the role of a new technology in increasing growth rates is not just a function of growth in maximum productivity $P(N)$. Rather, the growth rate of the multiplicity of fail points in the new technology is also important. How easy it is to implement a new technology is just as important for growth as how productive the technology is when correctly implemented. But how fast does $M(N)$ grow? And what tools do we have to overcome growing complexity's drag on growth?

4.1 Overcoming Complexity: The Case of Polynomial Complexity Growth

If the number of fail points increases one-for-one with the number of elements, but productivity is stagnant (e.g. $M(N) = N$ and $P(N)$ is constant) then economic growth will slow as the economy gets more complex (i.e. $\left(\frac{\partial g}{\partial N}\right)$ will be negative).

To overcome this malus from increased complexity, new technologies need to increase maximum productivity at a faster rate than they increase the multiplicity of fail points. If both M and P grow linearly in N, these effects perfectly offset, and the growth rate does not change over time (although the distribution of firm sizes will continue to become increasingly skewed).

It is also easy to imagine the number of fail points M growing with the square of the level of complexity. Consider the case of a watchmaker who must precisely align every mechanism in their timepiece.[3] If a watch has N elements, and each element needs to be put into perfect pair-wise alignment with each other element (as well as be calibrated itself), then the effective number of fail points grows at the rate

$$M(N) = N + \frac{N(N-1)}{2} \tag{18}$$

which grows asymptotically with the term N^2.

Suppose that watchmaker is tempted to add a new, $N+1$th element to her watch. Her decision to do so is a function of how much the new element will raise the value of a perfect watch. If the rate she can raise the watch's quality is also polynomial, i.e. $P(N) = bN^a$, then (in the limit, as $N \to \infty$) she will be better off in expectation by adopting the new technology when $a > 2$. For $a < 2$ she will be worse off with the new technology (it's too hard to implement to be worth it), and for $a = 2$ she'll be indifferent. Note also that exponential growth in maximum productivity as a function of N (i.e. $M(N) = b^N$) is also always sufficient to overwhelm polynomial growth in effective complexity. This perhaps explains why the technology industry has been able to see such gains in average productivity growth. The velocity of Moore's Law overcame increasing complexity.

Another approach to improving products without an offsetting increased to effective complexity is what you might refer to as a "Coperinican Revolution." While usually products are augmented with additional doohickeys or code which add an additional functionality or to handle more exceptions and edge-cases – 'bells and whistles' in other words – sometimes an industry achieves an innovation that allows for radical simplification of design with no offsetting loss in quality. We call these events Copernican revolutions to recall the event where the helio (or sun) centered model of the solar system replaced that of the classical Earth-based system. It was long known that the movements of the planets through the sky were only very approximately modeled through circular orbits of these

[3] This example is inspired in part by Simon (1962).

planets around earth. To correct this imprecision, classical astronomers added 'epicycles' (or smaller circles that diverted the heavenly spheres during the larger circles of their orbits) to their models. This increased the precision of their predictions at the cost of increased complexity. When Copernicus re-oriented the heavens around the Earth, he simplified astronomical science, while at the same time unlocking new potential for greater precision. Eventually, the arbitrary-seeming paths of the planets and all manner of physical phenomena would be united by Newtonian physics, another example of a scientific revolution that simultaneously simplified and empowered.

In the world of technology, one major example of a Copernican revolution was the shift from propeller driven to jet airplanes. The Wright Brothers' original aircraft engine was a bespoke creation, hand milled without any formal design drawings. It weighed 180 pounds and created 12 horsepower. An example of a modern patented propeller engine weighed 660 pounds, created 600 horsepower, and has 196 labled parts in its patent drawings. The first feasible turbojet engine patented – Whittle's 1940 design – is remarkably simple in comparison. It contains closer to 50 labeled parts in its patenting document. However, it produced 2187 horsepower at a weight of 950 pounds. Whittle's innovation is an example of one that enabled follow up innovations by going back to the drawing board and moving to a new foundation with lowered complexity. Future jet engine innovators were able to create even better engines by layering additional innovations on Whittle's new base concept, whereas propeller engines have not advanced dramatically beyond their mid-century level, despite increasing complexity.[4]

4.2 The Role of Modularization

A final, but perhaps most important and reliable, method of reducing effective complexity and the number of fail-points in a systems is modularization. Through modularization, a system can maintain a high level of functionality, but reduce the number of steps that an implementer must perfect.

To show how this works in our model, let's return to our mythical watchmaker. In-

[4]Wright Brothers.org (2011), Frank (1946), Fuerlinger et al. (2006), Wikipedia contributors (2021b), Wikipedia contributors (2021a).

stead of taking the three approaches already discussed (abandoning progress, increasing the maximum productivity of more advanced watches, or going back to the drawing board inventing a digital watch) the watchmaker might try increasing the modularization of his watch-assembly.[5]

A system with N elements can be organized into $\frac{N}{\mu}$ modules of size μ. For simplicity, we focus on modules of equal sizes, so μ must be a divisor of N.

If our watchmaker does not use any modules, then each element N of the watch must be put into perfect configuration with each other, pairwise. As noted, this leads to the number of fail points growing at the rate $M(N) = N + \frac{N(N-1)}{2}$. In the limit, the largest term in $M(N)$ will be $\frac{N^2}{2}$. Therefore $\lim_{N\to\infty} \frac{\frac{\partial M}{\partial N}}{M(N)} = 1/N$

Now, suppose that the watchmaker divides his assembly task into μ sub-assemblies (which we model as modules). This will reduce the number of fail points per element to be assembled $M(N)$. Exactly how much it will reduce the number of fail points depends both on the number of modules selected and how those modules interact.

Suppose, most conservatively, that each module must be pair-wise calibrated with each other module, and that also within each module each element must be put into pair-wise. This means that the only benefit of modularization is that elements in different modules do not need to interact with each other. In this case, $M(N,\mu) = N + \frac{N(N/\mu-1)}{2} + \frac{\mu(\mu-1)}{2}$. Introducing modularization to a system that previously did not have modules (i.e. increasing μ from zero to some constant) will potentially introduce a one-time reduction in complexity. The optimal modularization of a system, in the limit when N is large and given these interaction rules, is to set $\mu = N^{2/3}$. This leads to $\min_\mu M(N,\mu) = N + \frac{N^{4/3}-N}{2} + \frac{N^{4/3}-N^{1/3}}{2}$. Optimal modularization reduces largest term in the rate of technology growth from $\frac{N^2}{2}$ to $N^{4/3}$.

Another possibility is that modules are completely compartmentalized, and therefore do not need to be calibrated to properly interact. Each module must still be made to work by calibrating all of its internal elements. In this case, $M(N,\mu) = N + \frac{N(N/\mu-1)}{2}$.

[5]There are other ways to frame of how introducing modularization may increase the chances that a watchmaker produces a perfect watch. For example, if the watchmaker needs to restart a full watch every time she makes a mistake (without modularization) or only a single module with modularization, then modularization greatly decreases the number of steps that must be re-done after an inevitable mistake. This alternative framing is the one discussed in Simon (1962).

Here it is easy to see that the best case scenario is to set $\mu = N$. Because each module is completely compartmentalized, the most efficient strategy is simply to give each element its own module. This will reduce the rate of effective complexity growth to $M(N) = N$. Equations 19 and 20 summarize how modularization can decrease effective complexity growth.

$$M(N,\mu) = \underbrace{N + \frac{N(N-1)}{2}}_{Naive} \geq \underbrace{N + \frac{N(N/\mu - 1)}{2} + \frac{\mu(\mu-1)}{2}}_{\text{Interaction Within \& Across Modules}} \geq \underbrace{N + \frac{N(N/\mu-1)}{2}}_{\text{Interaction Within Modules}} \tag{19}$$

$$\min_{\mu} \lim_{N\to\infty} M(N,\mu) = \underbrace{N + \frac{N(N-1)}{2}}_{Naive} \geq \underbrace{N + \frac{N^{4/3}-N}{2} + \frac{N^{4/3}-N^{1/3}}{2}}_{\text{Interaction Within \& Across Modules: } \mu=N^{2/3}} \geq \underbrace{N}_{\text{Interaction Within Modules: } \mu=N} \tag{20}$$

Each of these approaches to modularization yields polynomial complexity growth (i.e. $M(N) = bN^a$ is the largest term in the limit). This means, for each of these modularizations, the asymptotic growth rate of M will be $\frac{\frac{\partial M}{\partial N}}{M(N)} = a/N$. Therefore, if the rate of maximum product quality technology growth is also polynomial of degree between 1 and 2 (i.e. $P(M) = b_P N^{a_P}$) then modularization in this context can enable accelerating long term growth (i.e. by lowering a to below a_P). Through this mechanism modularization may be key to the very long-term fate of the economy and Humanity.

Modularization might also decrease the rate of complexity growth to less than polynomial. One way to do this is through commodification. If an element of a process becomes so routine to create and implement that it can be viewed as a commodity, this is an element that the new firm does not need to master. As an extreme case, $M(N) = 0$ and any rate of technological progress will be enough to guarantee accelerating long-run growth rates.

5 Conclusion

Atul Gawande, in his 2020 opus, "Checklist Manifesto" wrote "We have accumulated stupendous know-how... accomplished extraordinary things. Nonetheless, that know-how is often unmanageable. Avoidable failures are common and persistent, not to mention demoralizing and frustrating, across many fields—from medicine to finance, business to government... Knowledge has both saved us and burdened us."

In this paper we developed a micro-founded variation on the classic AK growth model, highlighting the role of complexity. We showed effective complexity can be modeled as the number of 'fail points' in a Leontieff production function that must be mastered by entering companies. The expected productivity of companies is increasing in the maximum potential quality of a product and decreasing in how difficult it is to create that product.

At the microeconomic level, this model predicts that increasing complexity will increase the skewness of the firm size distribution. We found evidence of this process at play in three of four US sectors over the last forty years.

At the macroeconomic level, this model explains how the rate of economic growth is determined not just by maximum productivity but also the difficulty in implementing new innovations. We examined four different methods for addressing the challenges brought on by increased complexity. Progress can be abandoned, maximum productivity enhanced, processes can be modularized (reducing effective complexity per number of elements) and 'Copernican Revolutions' can return us to a drawing board with lowered complexity.

We showed that the long-term growth rate is increasing in the number of elements to be combined so long as the growth rate of maximum productivity is faster than the growth rate of effective complexity. For a range of possible effective complexity and maximum product quality growth rates, modularization can make the difference between economic growth which is accelerating or decelerating long run.

References

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The Fall of the Labor Share and the Rise of Superstar Firms," Working Paper 23396, National Bureau of Economic Research May 2017.

Bloom, Nicholas, Charles I Jones, John Van Reenen, and Michael Webb, "Are ideas getting harder to find?," *American Economic Review*, 2020, *110* (4), 1104–44.

Brynjolfsson, Erik, Avinash Collis, and Felix Eggers, "Using massive online choice experiments to measure changes in well-being," *Proceedings of the National Academy of Sciences*, 2019, *116* (15), 7250–7255.

— , **Daniel Rock, and Chad Syverson**, "Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics," 2017.

— , — , **and** — , "The productivity J-curve: How intangibles complement general purpose technologies," *American Economic Journal: Applied Economics – Forthcoming*, 2020.

Diewert, W Erwin and Kevin J Fox, "Can measurement error explain the productivity paradox?," *The Canadian Journal of Economics/Revue canadienne d'Economique*, 1999, *32* (2), 251–280.

Frank, Whittle, "Aircraft propulsion system and power unit," July 16 1946. US Patent 2,404,334.

Fuerlinger, Josef, Heinz Lippitsch, and Johann Bayerl, "Piston type aircraft engine," September 28 2006. US Patent App. 11/339,847.

Gold, E Richard, "The fall of the innovation empire and its possible rise through open science," *Research Policy*, 2021, *50* (5), 104226.

Hidalgo, César A and Ricardo Hausmann, "The building blocks of economic complexity," *Proceedings of the national academy of sciences*, 2009, *106* (26), 10570–10575.

Kremer, Michael, "The O-ring theory of economic development," *The Quarterly Journal of Economics*, 1993, *108* (3), 551–575.

Simon, Herbert A, "The Architecture of Complexity," *Proceedings of the American Philosophical Society*, 1962, *106* (6), 467–482.

Syverson, Chad, "Challenges to mismeasurement explanations for the US productivity slowdown," *Journal of Economic Perspectives*, 2017, *31* (2), 165–86.

Weitzman, Martin L., "Recombinant growth," *Quarterly Journal of Economics*, 1998, *113* (2), 331–360.

Wikipedia contributors, "Gloster E.28/390," 2021. [Online; accessed July 2021].

— , "Power Jets W.2," 2021. [Online; accessed July 2021].

Wright Brothers.org, "1903 Wright engine," 2011.