Finding Needles in Haystacks:

Artificial Intelligence and Recombinant Growth

Ajay Agrawal[1], John McHale[2], and Alexander Oettl[3,4]

November 2017

Abstract

In many fields, innovation is predicated on discovering useful new combinations of existing knowledge in highly complex knowledge spaces. Such needle-in-a-haystack problems are pervasive in fields like genomics, drug discovery, materials science, and particle physics. We develop a combinatorial-based knowledge production function and embed it in the classic Jones growth model (1995) to explore how breakthroughs in artificial intelligence (AI) that dramatically improve prediction accuracy about which combinations are most valuable could enhance discovery rates and consequently economic growth. This production function is a generalization (and reinterpretation) of the Romer/Jones knowledge production function. Separate parameters control the extent of individual-researcher knowledge access, the effects of fishing out/complexity, and the ease of forming research teams.

[1] University of Toronto and NBER.
[2] National University of Ireland Galway and Whitaker Institute for Innovation and Societal Development.
[3] Georgia Institute of Technology.
[4] We thank Joshua Gans and Chad Jones for thoughtful input on this paper. We gratefully acknowledge financial support from the Social Sciences Research Council of Canada and the Whitaker Institute for Innovation and Societal Development.

The potential for continued economic growth comes from the vast search space that we can explore. The curse of dimensionality is, for economic purposes, a remarkable blessing. To appreciate the potential for discovery, one need only consider the possibility that an extremely small fraction of the large number of potential mixtures may be valuable. (Paul Romer, 1993, pp. 68-69)

Deep learning is making major advances in solving problems that have resisted the best attempts of the artificial intelligence community for years. It has turned out to be very good at discovering intricate structure in high-dimensional data and is therefore applicable to many domains of science, business, and government. (Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, 2015, p. 436)

1. Introduction

What are the prospects for technology-driven economic growth? Technological optimists point in part to the ever-expanding possibilities for combining existing knowledge into new knowledge (Paul Romer, 1990, 1993; Martin Weitzman, 1998; Brian Arthur, 2009; Erik Brynjolfsson and Andrew McAfee, 2014). The counter case put forward by technological pessimists is primarily empirical: Growth at the technological frontier has been slowing down rather than speeding up (Tyler Cowen, 2011; Robert Gordon, 2016). Gordon (2016, p. 575) highlights this slowdown for the US economy. Between 1920 and 1970, total factor productivity grew at an annual average compound rate of 1.89 percent, falling to 0.57 percent between 1970 and 1994, then rebounding to 1.03 percent during the information technology boom between 1994 and 2004, before falling again to just 0.40 percent between 2004 and 2014. Even the maintenance of this lowered growth rate has only been possible due to exponential growth in the number of research workers (Charles Jones, 1995). Nicholas Bloom, Charles Jones, John Van Reenen, and Michael Webb (2017) document that the total factor productivity in knowledge production itself has been falling both in the aggregate and in key specific knowledge domains such as transistors, healthcare, and agriculture.

Economists have given a number of explanations for the disappointing growth performance. Cowen (2011) and Gordon (2016) point to a "fishing out" or "low-hanging fruit" effect – good ideas are simply becoming harder to find. Benjamin Jones

(2009) points to the headwind created by an increased "burden of knowledge." As the technological frontier expands, it becomes harder and harder for individual researchers to know enough to find the combinations of knowledge that produce useful new ideas. This is reflected in PhDs being awarded at older ages and a rise in team production as ever-more specialized researchers must combine their knowledge to produce breakthroughs. Other evidence points to the physical, social, and institutional constraints that limit access to knowledge, including the need to be physically close to the sources of knowledge (Adam Jaffe, Manuel Trajtenberg, and Rebacca Henderson, 1993; Christian Catalini, 2017), the importance of social relationships in accessing knowledge (Joel Mokyr, 2002; Ajay Agrawal, Iain Cockburn, and John McHale, 2006; Agrawal, Devesh Kapur, and McHale, 2008), and the importance of institutions in facilitating – or limiting – access to knowledge (Jeff Furman and Scott Stern, 2011).

Despite the evidence of a growth slowdown, one reason to be hopeful about the future is the recent explosion in data availability under the rubric of "big data" and computer-based advances in capabilities to discover and process those data. We can view these technologies in part as "meta technologies" – technologies for the production of new knowledge. If part of the challenge is dealing with the combinatorial explosion in the potential ways that existing knowledge can be combined as the knowledge base grows, then meta technologies such as deep learning hold out the potential to partially overcome the challenges of fishing out, the rising burden of knowledge, and the social and institutional constraints on knowledge access.

Of course, meta technologies that aid in the discovery of new knowledge are nothing new. Mokyr (2002; 2017) gives numerous examples of how scientific instruments such as microscopes and x-ray crystallography significantly aided the discovery process. Nathan Rosenberg (1998) provides an account of how technology-embodied chemical engineering altered the path of discovery in the petro-chemical industry. Moreover, the use of artificial intelligence for discovery is itself not new and has underpinned fields such as cheminformatics, bioinformatics, and particle physics for decades. However, recent breakthroughs in AI such as deep learning have given a new impetus to these fields.[5] The convergence of GPU-accelerated computing power, exponential growth in data availability buttressed in part by open data sources, and the rapid advance in AI-based prediction technologies is leading to breakthroughs in solving many needle-in-a-haystack problems. If the curse of dimensionality is both the blessing and curse of discovery, advances in AI offer renewed hope of breaking the curse while helping to deliver on the blessing.

Understanding how these technologies could affect future growth dynamics is likely to require an explicitly combinatorial framework. Weitzman's (1998) pioneering development of a recombinant growth model has unfortunately not been well incorporated into the corpus of growth theory literature. Our contribution in this paper is thus twofold. First, we develop a relatively simple combinatorial-based knowledge production function that converges in the limit to the Romer/Jones function. The model allows for the consideration of how existing knowledge is combined to produce new knowledge and also how researchers combine to form teams. Second, while this function can be incorporated into existing growth models, the specific combinatorial foundations mean that the model provides insights into how new meta technologies such as artificial intelligence might matter for the path of future economic growth.

Our paper thus contributes to a recent but rapidly expanding literature on the effects of AI on economic growth. Much of the focus of this new literature is on how increased automation substitutes for labor in the production process. Building on the pioneering work of Joseph Zeira (1998), Daron Acemoglu and Pascual Restrepo (2017) develop a model in which AI substitutes for workers in existing tasks but also creates new tasks for workers to do. Philippe Aghion, Benjamin Jones, and Charles Jones (2017) show how automation can be consistent with relatively constant factor shares when the elasticity of substitution between goods is less than one. Central to their results is Baumol's "cost disease," which posits the ultimate constraint on growth to be from goods that are essential but hard to improve rather than goods whose production benefits from AI-driven technical change. In a similar vein, William Nordhaus (2015)

[5] See, for example, the recent survey of the use of deep learning in computational chemistry by Garrett Goh, Nathan Hodas, and Abhinav Vishnu (2017).

explores the conditions under which AI would lead to an "economic singularity" and examines the empirical evidence on the elasticity of substitution on both the demand and supply sides of the economy.

Our focus is different from these papers in that instead of emphasising the potential substitution of machines for workers in existing tasks, we emphasise the importance of AI in overcoming a specific problem that impedes human researchers – finding useful combinations in complex discovery spaces. Our paper is closest in spirit to Iain Cockburn, Rebecca Henderson, and Scott Stern (2017), which examines the implications of AI – and deep learning in particular – as a general purpose technology (GPT) for invention. We provide a suggested formalization of this key idea. Nielsen (2012) usefully illuminates the myriad ways in which "big data" and associated technologies are changing the mechanisms of discovery in science. Nielsen emphasizes the increasing importance of "collective intelligence" in formal and informal networked teams, the growth of "data-driven intelligence" that can solve problems that challenge human intelligence, and the importance of increased technology facilitating access to knowledge and data. We incorporate all of these elements into the model developed in this paper.

The rest of the paper is organized as follows. In the next section, we outline some examples of how advances in artificial intelligence are changing both knowledge access and the ability to combine knowledge in high dimensional data across a number of domains. In Section 3, we develop an explicitly combinatorial-based knowledge production function and embed it in the growth model of Jones (1995), which itself is a modification of Romer (1990). In Section 4, we extend the basic model to allow for knowledge production by teams. We discuss our results in Section 5 and conclude in Section 6 with some speculative thoughts on how an "economic singularity" might emerge.

2. How Artificial Intelligence is Impacting the Production of Knowledge: Some Motivating Examples

Breakthroughs in AI are already impacting the productivity of scientific research and technology development. It is useful to distinguish between such meta technologies that aid in the process of search (knowledge access) and discovery (combining existing knowledge to produce new knowledge). For search, we are interested in AIs that solve problems that meet two conditions: 1) potential knowledge relevant to the process of discovery is subject to an explosion of data that an individual researcher or team of researchers finds increasingly difficult to stay abreast of (the "burden of knowledge"); and 2) the AI predicts which pieces of knowledge will be most relevant to the researcher, typically through the input of search terms. For discovery, we also identify two conditions: 1) potentially combinable knowledge for the production of new knowledge is subject to combinatorial explosion; and 2) the AI predicts which combinations of existing knowledge will yield valuable new knowledge across a large number of domains. We now consider some specific examples of how AI-based search and discovery technologies may change the innovation process.

Search

Meta$^{\alpha}$ produces AI-based search technologies for identifying relevant scientific papers and tracking the evolution of scientific ideas. The company was acquired by the Chan-Zuckerberg Foundation, which intends to make it available free of charge to researchers. This AI-based search technology meets our two conditions for a meta technology for knowledge access: 1) the stock of scientific papers is subject to exponential growth at an estimated 8-9 percent per year (Lutz Bornmann and Rüdiger Mutz, 2015); and 2) the AI-based search technology helps scientists identify relevant papers, thereby reducing the "burden of knowledge" associated with the exponential growth of published output.

BenchSci is an AI-based search technology for the more specific task of identifying effective compounds used in drug discovery (notably antibodies that act as reagents in scientific experiments). It again meets our two conditions: 1) reports on compound efficacy are scattered through millions of scientific papers with little standardisation in how these reports are provided; and 2) an AI extracts compound-efficacy information, allowing scientists to more effectively identify appropriate compounds to use in experiments.

Discovery

Atomwise is a deep learning-based AI for the discovery of drug molecules (compounds) that have the potential to yield safe and effective new drugs. This AI meets our two conditions for a meta technology for discovery: 1) the number of potential compounds is subject to combinatorial explosion; and 2) the AI predicts how basic chemical features combine into more intricate features to identify potential compounds for more detailed investigation.

DeepGenomics is a deep learning-based AI that predicts what happens in a cell when DNA is altered by natural or therapeutic genetic variation. It again meets our two conditions: 1) genotype-phenotype variations are subject to combinatorial explosion; and 2) the AI "bridges the genotype-phenotype divide" by predicting the results of complex biological processes that relate variations in the genotype to observable characteristics of an organism, thus helping to identify potentially valuable therapeutic interventions for further testing.

3. A Combinatorial-Based Knowledge Production Function

Figure 1 provides an overview of our modelling approach and how it relates to the classic Romer/Jones knowledge production function. The solid lines capture the essential character of the Romer/Jones function. Researchers use existing knowledge – the standing-on-shoulders effect – to produce new knowledge. The new knowledge then becomes part of the knowledge base from which subsequent discoveries are made. The dashed lines capture our approach. The existing knowledge base determines the potential new combinations that are possible, the majority of which are likely to have no value. The discovery of valuable new knowledge is made by searching among the massive number of potential combinations. This discovery process is aided by meta technologies such as deep learning that allow researchers to identify valuable combinations in spaces where existing knowledge interacts in often highly complex ways. As with the Romer/Jones function, the new knowledge adds to the knowledge base – and thus the potential combinations of that knowledge base – which subsequent researchers have to work with. A feature of our new knowledge production function will be that the Romer/Jones function emerges as a limiting case both with and without team production of new knowledge. In this section, we first develop the new function without team production of new knowledge; in the next section, we extend the function to allow for team production.

The total stock of knowledge in the world is denoted as A, which we assume initially is measured discretely. An individual researcher has access to an amount of knowledge, A^{ϕ}(also assumed to be an integer), so that the share of the stock of knowledge available to an individual researcher is $A^{\phi-1}$.[6] We assume that $0 < \phi < 1$. This implies that the share of total knowledge accessible to an individual researcher is falling with the total stock of knowledge. This is a manifestation in the model of the "burden of knowledge" effect identified by Jones (2009) – it becomes more difficult to access all the available knowledge as the total stock of knowledge grows. The knowledge access parameter, ϕ, is assumed to capture not only what a researcher knows at a point in time but also their ability to find existing knowledge should they require it. The value of the parameter will thus be affected by the extent to which knowledge is available in codified form and can be found as needed by researchers. The combination of digital repositories of knowledge and search technologies that can

[6] Paul Romer emphasized the importance of distinguishing between ideas (a non-rival good) and human capital (a rival good). "Ideas are . . . the critical input in the production of more valuable human and non-human capital. But human capital is also the most important input in the production of new ideas. . . . Because human capital and ideas are so closely related as inputs and outputs, it is tempting to aggregate them into a single type of good. . . . It is important, nevertheless, to distinguish ideas and human capital because they have different fundamental attributes as economic goods, with different implications for economic theory" (Romer, 1993, p. 71). In our model, A^{ϕ} is a measure of a researcher's human capital. Clearly, human capital depends on the existing technological and other knowledge and the researcher's access to that knowledge. In turn, the production of new knowledge depends on the researcher's human capital.

predict what knowledge will be most relevant to the researcher given the search terms they input – think of the ubiquitous Google as well as more specialized search technologies such Meta$^{\alpha}$ and BenchSci – should increase the value of ϕ.

Innovations occur as a result of combining existing knowledge to produce new knowledge. Knowledge can be combined a ideas at a time, where $a = 0, 1 \ldots A^{\phi}$. For a given individual researcher, the total number of possible combinations of units of existing knowledge (including singletons and the null set)[7] given their knowledge access is:

(1) $$Z_i = \sum_{a=0}^{A^{\phi}} \binom{A^{\phi}}{a} = 2^{A^{\phi}}.$$

The total number of potential combinations, Z_i, grows exponentially with A^{ϕ}. Clearly, if A is itself growing exponentially, Z_i will be growing at a double exponential rate. This is the source of combinatorial explosion in the model. Since it is more convenient to work with continuously measured variables in the growth model, from this point on we treat A and Z_i as continuously measured variables. However, the key assumption is that the number of potential combinations grow exponentially with knowledge access.

The next step is to specify how potential combinations map to discoveries. We assume that a large share of potential combinations do not produce useful new knowledge. Moreover, of those combinations that are useful, many will have already been discovered and thus are already part of A. This latter feature reflects the fishing-out phenomenon. The per period translation of potential combinations into valuable new knowledge is given by the (asymptotically) constant elasticity discovery function:

[7] Excluding the singletons and the null set, total number of potential combinations would be $2^{A^{\phi}} - A^{\phi} - 1$. As singletons and the null set are not true "combinations," we take equation (1) to be an approximation of the true number of potential combinations. The relative significance of this approximation will decline as the knowledge base grows, and we ignore it in what follows.

$$(2) \qquad \dot{A}_i = \beta\left(\frac{Z_i^{\theta}-1}{\theta}\right) = \beta\left(\frac{\left(2^{A^{\phi}}\right)^{\theta}-1}{\theta}\right) \qquad for \quad 0<\theta\leq 1$$

$$= \beta \ln Z_i = \beta \ln\left(2^{A^{\phi}}\right) = \beta \ln(2) A^{\phi} \qquad for \quad \theta = 0,$$

where β is a positively valued knowledge discovery parameter and use is made of L'Hôpital's rule for the limiting case of $\theta = 0$.[8]

For $\theta > 0$, the elasticity of new discoveries with respect to the number of possible combinations, Z_i, is:

$$(3) \qquad \frac{\partial \dot{A}}{\partial Z_i}\frac{Z_i}{\dot{A}} = \frac{\beta Z_i^{\theta-1}}{\beta\left(\frac{Z_i^{\theta}-1}{\theta}\right)} = \left(\frac{Z_i^{\theta}}{Z_i^{\theta}-1}\right)\theta,$$

which converges to θ as the number of potential combinations goes to infinity. For $\theta = 0$, the elasticity of new discoveries is:

$$(4) \qquad \frac{\partial \dot{A}}{\partial Z_i}\frac{Z_i}{\dot{A}} = \frac{\beta}{Z_i}\frac{Z_i}{\beta lnZ_i} = \frac{1}{lnZ_i},$$

[8] L'Hôpital's rule is often useful where a limit of a quotient is indeterminate. The limit of the term in brackets on the right-hand-side of equation (2) as θ goes to zero is 0 divided by 0 and is thus indeterminate. However, by L'Hôpital's rule, the limit of this quotient is equal to the limit of the quotient produced by dividing the limit of the derivative of the numerator with respect to θ by the limit of the derivative of the denominator with respect to θ. This limit is equal to $\ln(2)A^{\phi}$.

which converges to zero as the number of potential combinations goes to infinity.

A number of factors seem likely to affect the value of the fishing-out/complexity parameter, θ. First are basic constraints relating to natural phenomena that limit what is physically possible in terms of combining existing knowledge to produce scientifically or technologically useful new knowledge. Pessimistic views on the possibilities for future growth tend to emphasize such constraints. Second is the ease of discovering new useful combinations that are physically possible. The potentially massive size and complexity of the space of potential combinations means that finding useful combinations can be a needle-in-the-haystack problem. Optimistic views of the possibilities for future growth tend to emphasize how the combination of AI (embedded in algorithms such as those developed by Atomwise and DeepGenomics) and increases in computing power can aid prediction in the discovery process, especially where it is difficult to identify patterns of cause and effect in high dimensional data. Third, recognizing that future opportunities for discoveries are path dependent (see, for example, Weitzman, 1998), the value of θ will depend on the actual path that is followed. To the extent that AI can help identify productive paths, it will limit the chances of economies going down technological dead-ends.

There are L_A researchers in the economy each working independently, where L_A is assumed to be measured continuously. (In Section 4, we consider the case of team production in an extension of the model.) We assume that some researchers will duplicate each other's discoveries – the standing-on-toes effect. To capture this effect, new discoveries are assumed to take place "as if" the actual number of researchers is equal to L_A^{λ}, where $0 \leq \lambda \leq 1$. Thus the aggregate knowledge production function for $\theta > 0$ is given:

(5) $$\dot{A} = \beta L_A^{\lambda}\left(\frac{\left(2^{A^{\phi}}\right)^{\theta} - 1}{\theta}\right).$$

At a point in time (with given values of A and L_A), how does an increase in θ affect the rate of discovery of new knowledge, $\dot{A}$? The partial derivative of $\dot{A}$ with respect to θ is:

(6) $$\frac{\partial \dot{A}}{\partial \theta} = \frac{\beta L_A^{\lambda}\left(\theta \ln(2)\, A^{\phi} - 1\right)2^{A^{\phi}\theta}}{\theta^2} + \frac{\beta L_A^{\lambda}}{\theta^2}.$$

A sufficient condition for this partial derivative to be positive is that that term in square brackets is greater than zero, which requires:

(7) $$A > \left(\frac{1}{\theta \ln(2)}\right)^{\frac{1}{\phi}}.$$

We assume this condition holds. Figure 2 shows an example of how $\dot{A}$ (and also the percentage growth of A given that A is assumed to be equal to 100) varies with θ for different assumed values of ϕ. Higher values of θ are associated with a faster growth rate. The figure also shows how θ and ϕ interact positively: Greater knowledge access (as reflected in a higher value of ϕ) increases the gain associated with a given increase in the value of θ.

We assume, however, that θ itself evolves with A. A larger A means a bigger and more complex discovery search space. We further assume that this complexity will eventually overwhelm any discovery technology given the power of the combinatorial explosion as A grows. This is captured by assuming that θ is a declining function of A;

that is, $\theta = \theta(A)$, where $\theta'(A) < 0$. In the limit as A goes to infinity, we assume that $\theta(A)$ goes to zero, or:

(8) $$\lim_{A\to\infty} \theta(A) = 0.$$

This means that the discovery function converges asymptotically (given sustained growth in A) to:

(9) $$\dot{A} = \beta \ln(2) L_A^{\lambda} A^{\phi}.$$

This mirrors the functional form of the Romer/Jones function and allows for decreasing returns to scale in the number of researchers, depending on the size of λ. While the form of the function is familiar by design, its combinatorial-based foundations have the advantage of providing richer motivations for the key parameters in the knowledge discovery function.

We use the fact that the functional form of equation (9) is the same as that used in Jones (1995) to solve for the steady state of the model. More precisely, given that the limiting behaviour of our knowledge production function mirrors the function used by Jones and all other aspects of the economy are assumed to be identical, the steady-state along a balanced growth path with constant exponential growth will be the same as in that model.

As we have nothing to add to the other elements of the model, we here simply sketch the growth model developed by Jones (1995), referring the reader to the original for details. The economy is composed of a final goods sector and a research sector. The final goods sector uses labor, L_Y, and intermediate inputs to produce its output. Each new idea (or "blueprint") supports the design of an intermediate input, with each input being supplied by a profit-maximizing monopolist. Given the blueprint, capital, K, is transformed unit for unit in producing the input. The total labor force, L, is fully allocated between the final goods and research sectors, so that $L_Y + L_A = L$. We assume the labor force to be equal to the population and growing at rate $n(> 0)$.

Building on Romer (1990), Jones (1995) shows that the production function for final goods can be written as:

(10) $$Y = (AL_Y)^{\alpha} K^{1-\alpha},$$

where Y is final goods output. The intertemporal utility function of a representative consumer in the economy is given by:

(11) $$U = \int_0^{\infty} u(c) e^{-\rho t} dt,$$

where c is per capita consumption and ρ is the consumer's discount rate. The instantaneous utility function is assumed to exhibit constant relative risk aversion, with a coefficient of risk aversion equal to σ and a (constant) intertemporal elasticity of substitution equal to $1/\sigma$.

Jones (1995) shows that the steady-state growth rate of this economy along a balanced growth path with constant exponential growth is given by:

(12) $$g_A = g_y = g_c = g_k = \frac{\lambda n}{1 - \phi},$$

where $g_A = \dot{A}/A$ is the growth rate of the knowledge stock, g_y is the growth rate of per capita output y ($where\ y = Y/L$), g_c is the growth rate of per capita output c ($where\ c = C/L$), and g_k is the growth rate of the capital labor ratio ($where\ k = K/L$).

Finally, the steady-state share of labor allocated to the research sector is given by:

$$(13) \qquad s = \frac{1}{1 + \frac{1}{\lambda\left[\frac{\rho(1-\phi)}{\lambda n} + \frac{1}{\sigma} - \phi\right]}}.$$

We can now consider how changes in the parameters of knowledge production given by equation (5) will affect the dynamics of growth in the economy. We start with improvement in the availability of AI-based search technologies that improve a researcher's access to knowledge. In the context of the model, the availability of AI-based search technologies – e.g., Google, Meta$^{\alpha}$, BenchSci, etc. – should increase the value of ϕ and reduce the "burden of knowledge" effect. From equation (12), an increase in this parameter will increase the steady steady-state growth rate and also the growth rate and the level of per capital output along the transition path to the steady state.

We next consider AI-based technologies that increase the value of the discovery parameter, β. As β does not appear in the steady state in equation (12), the steady-state growth rate is unaffected. However, such an increase will raise the growth rate (and level) along the path to that steady state.

The most interesting potential changes to the possibilities for growth come about if we allow a change to the fishing-out/complexity parameter, θ. We assume that the economy is initially in a steady state and then experiences an increase in θ as the result of the discovery of a new AI technology. Recall that we assume that θ will eventually converge back to zero as the complexity that comes with combinatorial explosion eventually overwhelms the new AI. Thus, the steady state of the economy is unaffected. However, the transition dynamics are again quite different, with larger increases in knowledge for an given starting of the knowledge stock along the path back to the steady state.

Using Jones (1995) as the limiting case of the model is appealing because we avoid unbounded increases in the growth rate, which would lead to the breakdown of any reasonable growth model and indeed a breakdown in the normal operations of any actual economy. It is interesting to note, however, what happens to growth in the economy if instead of assuming that θ converges asymptotically to zero, it stays at some positive value (even if very small). Dividing both sides of equation (5) by A gives an expression for the growth rate of the stock of knowledge:

$$(14) \qquad \frac{\dot{A}}{A} = \frac{\beta \ln(2)\, L_A^{\lambda}}{A}\left(\frac{\left(2^{A^{\phi}}\right)^{\theta} - 1}{\theta}\right).$$

The partial derivative of this growth rate with respect to A is:

$$(15) \qquad \frac{\partial\left(\frac{\dot{A}}{A}\right)}{\partial A} = \frac{L_A^{\lambda}\beta}{\theta A^2}\left[1 + \left(2^{A^{\phi}}\right)^{\theta}(\phi\theta \ln(2) A^{\phi} - 1)\right].$$

The key to the sign of this derivative is the sign of the term inside the last round brackets. This term will be positive for a large enough A. As A is growing over time (for any positive number of researchers and existing knowledge stock), the growth rate must eventually begin to rise once A exceeds some threshold value. Thus, with a fixed positive value of θ (or with θ converging asymptotically to a positive value), the growth rate will eventually begin to grow without bound.

4. A Combinatorial-Based Knowledge Production Function with Team Production: An Extended Model

Our basic model assumes that researchers working alone combine the knowledge to which they have access, A^{ϕ}, to discover new knowledge. In reality, new discoveries are increasingly being made by research teams (Benjamin Jones, 2009; Nielsen, 2012; Agrawal, Avi Goldfarb, Florenta Teodoridis, 2016). Assuming initially no redundancy in the knowledge that individual members bring to the team – i.e., collective team knowledge is the sum of the knowledge of the individual team members – combining individual researchers into teams can greatly expand the knowledge base from which new combinations of existing knowledge can be made. This also opens up the possibility of a positive interaction between factors that facilitate the operation of larger teams and factors that raise the size of the fishing out/complexity parameter, θ. New meta technologies such as deep learning can be more effective in a world where they are operating on a larger knowledge base due to the ability of researchers to more effectively pool their knowledge by forming larger teams.

We thus extend in this section the basic model to allow for new knowledge to be discovered by research teams. For a team with m members and no overlap in the knowledge of its members, the total knowledge access for the team is simply mA^{ϕ}. (We later relax the assumption of no knowledge overlap within a team.) Innovations occur as a result of the team combining existing knowledge to produce new knowledge. Knowledge can be combined by the team a ideas at a time, where $a = 0, 1 \ldots mA^{\phi}$. For a given team j with m members, the total number of possible combinations of units of existing knowledge (including singletons and the null set) given their combined knowledge access is:

$$(16) \quad Z_j = \sum_{a=0}^{mA^{\phi}} \binom{mA^{\phi}}{a} = 2^{mA^{\phi}}.$$

Assuming again for convenience that A^{Φ} and Z can be treated as continuous, the per period translation of potential combinations into valuable new knowledge by a team is again given by the (asymptotic) constant elasticity discovery function:

$$(17) \quad \dot{A}_j = \beta\left(\frac{Z_j^{\theta}-1}{\theta}\right) = \beta\left(\frac{\left(2^{mA^{\Phi}}\right)^{\theta}-1}{\theta}\right) \qquad for \quad 0 < \theta \leq 1$$

$$= \beta \ln Z_j = \beta \ln\left(2^{mA^{\Phi}}\right) = \beta \ln(2) m A^{\Phi} \qquad for \quad \theta = 0,$$

where use is again made of L'Hôpital's rule for the limiting case of $\theta = 0$.

The number of researchers in the economy at a point in time is again L_A (which we now assume is measured discretely). Research teams can potentially be formed from any possible combination of the L_A researchers. For each of these potential teams, an entrepreneur can coordinate the team. However, for a potential team with m members to form, the entrepreneur must have relationships with all m members. The need for a relationship thus places a constraint on feasible teams. The probability of a relationship existing between the entrepreneur and any given researcher is η, and thus the probability of relationships existing between all members of a team of size m is η^m. Using the formula for a binomial expansion, the expected total number of feasible teams is:

$$(18) \quad S = \sum_{m=0}^{L_A} \binom{L_A}{m} \eta^m = (1+\eta)^{L_A}.$$

The average feasible team size is then given by:

$$(19)\quad \bar{m} = \frac{\sum_{m=0}^{L_A} \binom{L_A}{m} \eta^m m}{\sum_{m=0}^{L_A} \binom{L_A}{m} \eta^m}.$$

Factorizing the numerator and substituting in the denominator using equation (18), we obtain a simple expression for the average feasible team size:

$$(20)\quad \bar{m} = \frac{\sum_{m=0}^{L_A} \binom{L_A}{m} \eta^m m}{\sum_{m=0}^{L_A} \binom{L_A}{m} \eta^m} = \frac{(1+\eta)^{L_A-1}\eta L_A}{(1+\eta)^{L_A}} = \left(\frac{\eta}{1+\eta}\right) L_A.$$

Figure 3 shows an example of the full distribution of teams sizes (with $L_A = 50$) for two different values of η. An increase in η (i.e. an improvement in the capability to form teams) will push the distribution to the right and increase the average team size.

We can now write down the form that the knowledge production function would take if all possible research teams could form (ignoring for the moment any stepping-on-toes effects):

$$(21)\quad \dot{A} = \left(\sum_{m=0}^{L_A} \binom{L_A}{m} \eta^m \beta \frac{\left(2^{mA^\phi}\right)^\theta - 1}{\theta} \right) \qquad for \quad 0 < \theta \leq 1.$$

We next allow for the fact that only a fraction of the feasible teams will actually form. Recognising obvious time constraints on the ability of a given researcher to be part of multiple research teams, we impose the constraint that each researcher can only be part of one team. However, we assume the size of any team that successfully forms is drawn from the same distribution over sizes as the potential teams. Therefore, the expected average team size is also given by equation (18). With this restriction, we can solve for the total number of teams, *N*, from the equation $L_A = N\left(\frac{\eta}{1+\eta}\right)L_A$, which implies $N = \frac{1+\eta}{\eta}$.

Given the assumption that the distribution of actual team sizes is drawn from the same distribution as the feasible team sizes, the aggregate knowledge production function (assuming $\theta > 0$) is then given by:

$$(22)\qquad \dot{A} = \frac{\frac{1+\eta}{\eta}}{(1+\eta)^{L_A}}\left(\sum_{m=0}^{L_A}\binom{L_A}{m}\eta^m\beta\frac{\left(2^{mA^{\phi}}\right)^{\theta}-1}{\theta}\right)$$

$$= \frac{1}{(1+\eta)^{L_A-1}\eta}\left(\sum_{m=0}^{L_A}\binom{L_A}{m}\eta^m\beta\frac{\left(2^{mA^{\phi}}\right)^{\theta}-1}{\theta}\right),$$

where the first term is the actual number of teams as a fraction of the potentially feasible number of teams. For $\theta = 0$, the aggregate knowledge production function takes the form:

$$(23)\qquad \dot{A} = \frac{1}{(1+\eta)^{L_A-1}\eta}\left(\sum_{m=0}^{L_A}\binom{L_A}{m}\eta^m m\beta\ln(2)A^{\phi}\right)$$

$$= \frac{1}{(1+\eta)^{L_A-1}\eta}\left((1+\eta)^{L_A-1}\eta L_A\beta\ln(2)A^{\phi}\right),$$

$$= \beta L_A\ln(2)A^{\phi}.$$

To see intuitively how an increase in η could affect aggregate knowledge discovery when $\theta > 0$, note that from equation (20) an increase in η will increase the average team size of the teams that form. From equation (16), we see that for a given knowledge access by an individual researcher, the number of potential combinations increases exponentially with the size of the team, m (see Figure 4). This implies that combining two teams of size m' to create a team of size $2m'$ will more than double the new knowledge output of the team. Hence, there is a positive interaction between θ and η. On the other hand, when $\theta = 0$, combining the two teams will exactly double the new knowledge output given the linearity of the relationship between team size and knowledge output. In this case, the aggregate knowledge is invariant to the distribution of team sizes.

To see this formally, note that from equation (23) we know that when $\theta = 0$, the partial derivative of $\dot{A}$ with respect to η must be zero since η does not appear in the final form of the knowledge production function. This results from the balancing of two effects as η increases. The first (negative) effect is that the number of teams as a share of the potentially possible teams falls. The second (positive) effect is that the amount of new knowledge production if all possible teams do form rises. We can now ask what happens if we raise θ to a strictly positive value. The first of these effects is unchanged. But that second effect will be stronger provided that the knowledge production of a team for any given team size rises with θ. A sufficient condition for this to be true is that:

(24) $$A > \left(\frac{1}{\theta \ln(2) m}\right)^{\frac{1}{\phi}} \qquad \text{for all } m > 0.$$

We assume that the starting size of the knowledge stock is large enough so that this condition holds. Moreover, the partial derivative of $\dot{A}$ with respect to η will be larger the larger is the value of θ. We show these effects for a particular example in Figure 5.

The possibilities of knowledge overlap at the level of the team and duplication of knowledge outputs between teams creates additional complications. To allow for stepping-on-toes effects, it is useful to first rewrite equation (20) as:

$$(25) \qquad \dot{A} = \left(\frac{1+\eta}{\eta}\right)\left(\frac{\eta}{1+\eta}\right)L_A \, \frac{1}{(1+\eta)^{L_A-1}\eta L_A}\left(\sum_{m=0}^{L_A}\binom{L_A}{m}\eta^m \beta \, \frac{\left(2^{mA^{\phi}}\right)^{\theta} - 1}{\theta}\right).$$

We introduce two stepping-on-toes effects. First, we allow for knowledge overlap within teams to introduce the potential for redundancy of knowledge. A convenient way to introduce this effect is to assume that the overlap reduces the *effective* average team size in the economy from the viewpoint of generating new knowledge. More specifically, we assume the effective team size is given by:

$$(26) \qquad \bar{m}^e = \bar{m}^{\gamma} = \left(\left(\frac{\eta}{1+\eta}\right)L_A\right)^{\gamma},$$

where $0 \leq \gamma \leq 1$. The extreme case of $\gamma = 0$ (full overlap) has each team acting as if it had effectively a single member; the opposite extreme of $\gamma = 1$ (no overlap) has no knowledge redundancy at the level of the team. Second, we allow for the possibility that new ideas are duplicated across teams. The effective number of non-idea-duplicating teams is given by:

$$(27) \qquad N^e = N^{1-\psi} = \left(\frac{1+\eta}{\eta}\right)^{1-\psi},$$

where $0 \leq \psi \leq 1$. The extreme case of $\psi = 0$ (no duplication) implies that the effective number of teams is equal to the actual number of teams; the extreme case of $\psi = 1$ (full duplication) implies that a single team produces the same number of new ideas as the full set of teams.

We can now add the stepping-on-toes effects – knowledge redundancy within teams and discovery duplication between teams – to yield the general form of the knowledge production function for $\theta > 0$:

$$(28) \qquad \dot{A} = \left(\frac{1+\eta}{\eta}\right)^{1-\psi} \left(\left(\frac{\eta}{1+\eta}\right) L_A\right)^{\gamma} \frac{1}{(1+\eta)^{L_A-1}\eta L_A} \left(\sum_{m=0}^{L_A} \binom{L_A}{m} \eta^m \beta \frac{\left(2^{mA^{\phi}}\right)^{\theta} - 1}{\theta}\right).$$

If we take the limit of equation (24) as θ goes to zero, we reproduce the limiting case of the knowledge production function. Ignoring integer constraints on L_A, this knowledge production function again has the form of the Romer/Jones function:

$$(29) \qquad \dot{A} = \left(\frac{1+\eta}{\eta}\right)^{1-\psi} \left(\left(\frac{\eta}{1+\eta}\right) L_A\right)^{\gamma} \frac{1}{(1+\eta)^{L_A-1}\eta L_A} \left(\sum_{m=0}^{L_A} \binom{L_A}{m} \eta^m \beta \ln(2)\, m A^{\phi}\right)$$

$$= \left(\frac{1+\eta}{\eta}\right)^{1-\psi} \left(\left(\frac{\eta}{1+\eta}\right) L_A\right)^{\gamma} \frac{(1+\eta)^{L_A-1}\eta L_A}{(1+\eta)^{L_A-1}\eta L_A} \beta \ln(2)\, A^{\phi}$$

$$= \left(\frac{1+\eta}{\eta}\right)^{1-\psi} \left(\left(\frac{\eta}{1+\eta}\right)\right)^{\gamma} \beta \ln(2)\, {L_A}^{\gamma} A^{\phi}.$$

We note finally the presence of the relationship parameter η in the knowledge production equation. This can be taken to reflect in part the importance of (social) relationships in the forming of research teams. Advances in computer-based technologies such as email and file sharing (as well as policies and institutions) could also affect this parameter (see, for example, Agrawal and Goldfarb (2008) on the effects of the introduction of precursors to today's internet on collaboration between researchers). Although not the main focus of this paper, being able to incorporate the effects of changes in collaboration technologies increases the richness of the framework for considering the determinants of the efficiency of knowledge production.

5. Discussion

5.1 Something new under the sun? Deep learning as a new tool for discovery

Two key observations motivate the model developed above. First, using the analogy of finding a needle in a haystack, significant obstacles to discovery in numerous domains of science and technology result from highly non-linear relationships of causes and effect in high dimensional data. Second, advances in algorithms such as deep learning (combined with increased availability of data and computing power) offer the potential to find relevant knowledge and predict combinations that will yield valuable new discoveries.

Even a cursory review of the scientific and engineering literatures indicates that needle-in-the-haystack problems are pervasive in many frontier fields of innovation, especially in areas where matter is manipulated at the molecular or sub-molecular level. In the field of genomics, for example, complex genotype-phenotype interactions make it difficult to identify therapies that yield valuable improvements in human health or agricultural productivity. In the field of drug discovery, complex interactions between drug compounds and biological systems present an obstacle to identifying promising new drug therapies. And in the field of material sciences, including nanotechnology, complex interactions between the underlying physical and chemical mechanisms increases the challenge of predicting the performance of potential new materials with potential applications ranging from new materials to prevent traumatic brain injury to lightweight materials for use in transportation to reduce dependence on carbon-based fuels (National Science and Technology Council, 2011).

The apparent speed with which deep learning is being applied in these and other fields suggests it represents a breakthrough general purpose meta technology for predicting valuable new combinations in highly complex spaces. Although an in-depth discussion of the technical advances underlying deep learning is beyond the scope of this paper, two aspects are worth highlighting. First, previous generations of machine learning were constrained by the need to extract features (or explanatory variables) by hand before statistical analysis. A major advance in machine learning involves the use of "representation learning" to automatically extract the relevant features.[9] Second, the development and optimization of multilayer neural networks allows for substantial improvement in the ability to predict outcomes in high-dimensional spaces with complex non-linear interactions (LeCun, Bengio, and Hinton, 2015). A recent review of the use of deep learning in computational biology, for instance, notes that the "rapid increase in biological data dimensions and acquisition rates is challenging conventional analysis strategies," and that "[m]odern machine learning methods, such as deep learning, promise to leverage very large data sets for finding hidden structure within them, and for making accurate predictions" (Christof Angermueller, Tanel Pärnamaa, Leopold Parts, and Oliver Stegle, 2016, p.1). Another review of the use of deep learning in computational chemistry highlights how deep learning has a "ubiquity and broad applicability to a wide range of challenges in the field, including quantitative activity relationship, virtual screening, protein structure prediction, quantum chemistry, materials design and property prediction" (Goh, Hoda, and Vishu, 2017).

Although the most publicized successes of deep learning have been in areas such as image recognition, voice recognition, and natural language processing, parallels to the way in which the new methods work on unstructured data are increasingly being identified in many fields with similar data challenges to produce research breakthroughs.[10] While these new general purpose research tools will not displace

[9] As described by LeCun, Bengio, and Hinton (2015, p. 436), "[c]onventional machine-learning techniques were limited in their ability to process natural data in their raw form. For decades, constructing a pattern-recognition or machine-learning system required careful engineering and considerable domain expertise to design a feature extractor that transformed the raw data (such as the pixel values of an image) into a suitable internal representation or feature vector from which the learning subsystem, often a classifier, could detect or classify patterns in the input. . . . Representation learning is a set of methods that allows a machine to be fed with raw data and to automatically discover the representations needed for detection or classification."
[10] A recent review of deep learning applications in biomedicine usefully draws out these parallels: "With some imagination, parallels can be drawn between biological data and the types

traditional mathematical models of cause and effect and careful experimental design, machine learning methods such as deep learning offer a promising new tool for discovery – including hypothesis generation – where the complexity of the underlying phenomena present obstacles to more traditional methods.[11]

5.2 Meta ideas, meta technologies, and general purpose technologies

We conceptualize AIs as general purpose meta technologies – that is, general purpose technologies (GPTs) for the discovery of new knowledge. Figure 6 summarises the relationship between Paul Romer's broader idea of meta ideas, meta technologies, and GPTs. Romer defines a meta idea as an idea that supports the production and transmission of other ideas (see, for example, Romer, 2008). He points to such ideas as the patent, the agricultural extension station, and the peer-review system for research grants as examples of meta ideas. We think of meta technologies as a subset of Romer's meta ideas (the area enclosed by the dashed lines in Figure 6), where the idea for how to discover new ideas is embedded in a technological form such as an algorithm or measurement instrument.

Elhanan Helpman (1998, p. 3) argues that a "drastic innovation qualifies as a GPT if it has the potential for pervasive use in a wide range of sectors in ways that drastically change their mode of operation." He further notes two important features necessary to qualify as a GPT: "generality of purpose and innovational complementarities" (see also Bresnahan and Trajtenberg, 1995). Not all meta

---

of data deep learning has shown the most success with – namely image and voice data. A gene expression profile, for instance, is essentially a 'snapshot,' or image, of what is going on in a given cell or tissue in the same way that patterns of pixilation are representative of the objects in a picture" (Polina Mamoshina, Armando Vieira, Evgeny Putin, and Alex Zhavoronkov, 2016, p. 1445).

[11] A recent survey of the emerging use of machine learning in economics (including policy design) provides a pithy characterization of the power of the new methods: "The appeal of machine learning is that it manages to uncover generalizable patterns. In fact, the success of machine learning at intelligence tasks is largely due to its ability to discover complex structure that was not specified in advance. It manages to fit complex and very flexible functional forms to the data without simply overfitting; it finds functions that work well out of sample" (Sendhil Mullainathan and Jann Spiess, 2017, p. 88).

technologies are general purpose in this sense. The set of general purpose meta technologies is given by the intersection of the two circles in Figure 6. Cockburn, Henderson, and Stern (2017) give the example of functional MRI as an example of a discovery tool that lacks the generality of purpose required for a GPT. In contrast, the range of application of deep learning as a discovery tool would appear to qualify it as a GPT. It is worth noting that some authors discuss GPTs as technologies that more closely align with our idea of a meta technology. Rosenberg (1998), for example, provides a fascinating examination of chemical engineering as an example of GPT. Writing of this branch of engineering, he argues that a "discipline that provides the concepts and methodologies to generate new or improved technologies over a wide range of downstream economic activity may be thought of as an even purer, or higher order, GPT" (Rosenberg, 1998, p. 170).

Our concentration on general purpose meta technologies (GPMTs) parallels Cockburn, Henderson, and Stern's (2017) idea of a general purpose invention of a method of invention. This idea combines the idea of a GPT with Zvi Griliches' (1957) idea of the "invention of a method of invention," or IMI. Such an invention has the "potential for a more influential impact than a single invention, but is also likely to be associated with a wide variation in the ability to adapt the new tool to particular settings, resulting in a more heterogeneous pattern of diffusion over time" (Cockburn, Henderson, and Stern, 2017, p. 4). They see some emerging AIs such as deep learning as candidates for such general purpose IMIs and contrast these with AIs underpinning robotics that, while being GPTs, do not have the characteristic features of an IMI.

5.3 Beyond AI: potential uses of the new knowledge production function

Although the primary motivation for this paper is to explore how breakthroughs in AI could affect the path of economic growth, the knowledge production function we develop is potentially of broader applicability. By deriving the Romer/Jones knowledge production function as the limiting case of a more general function, our analysis may also contribute to providing candidate micro-foundations for that function.[12] The key conceptual change is to model discovery as operating on the space of potential combinations (rather than directly on the knowledge base itself). As in Weitzman (1998), our production function focuses attention explicitly on how new knowledge is discovered by combining existing knowledge, which is left implicit in the Romer/Jones formulation. While this shift in emphasis is motivated by the particular way in which deep learning can aid discovery – allowing researchers to uncover otherwise hard-to-find valuable combinations in highly complex spaces – the view of discovery as the innovative combination of what is already known has broader applicability. The more general function also has the advantage of providing a richer parameter space for mapping how meta technologies or policies could affect knowledge discovery. The ϕ parameter captures how access to knowledge at the individual researcher level determines the potential for new combinations to be made given the inherited knowledge base. The θ parameter captures how the available potential combinations (given the access to knowledge) map to new discoveries. Finally, the η parameter captures the ease of forming research teams and ultimately the average team size. To the extent that the capacity to bring the knowledge of individual researchers together through research teams directly affects the possible combinations, the ease of team formation can have an important effect on how the existing knowledge base is utilized for new knowledge discovery.

[12] In developing and applying the Romer/Jones knowledge production function, growth theorists have understood its potential combinatorial underpinnings and the limits of the Cobb-Douglas form. Charles Jones (2005) observes in his review chapter on "Growth and Ideas" for the *Handbook of Economic Growth*: "While we have made much progress in understanding economic growth in a world where ideas are important, there remain many open, interesting research questions. The first is 'What is the shape of the idea production function?' How do ideas get produced? The combinatorial calculations of Romer (1993) and Weitzman (1998) are fascinating and suggestive. The current research practice of modelling the idea production function as a stable Cobb-Douglas combination of research and the existing stock of ideas is elegant, but at this point we have little reason to believe it is correct. One insight that illustrates the incompleteness of our knowledge is that there is no reason why research productivity should be a smooth monotonic function of the stock of ideas. One can easily imagine that some ideas lead to domino-like unravelling of phenomena that were previously mysterious . . . Indeed, perhaps decoding of the human genome or the continued boom in information technology will lead to a large upward shift in the production function for ideas. On the other hand, one can equally imagine situations where research productivity unexpectedly stagnates, if not forever then at least for a long time" (Jones, 2005, p. 1107).

We hope this more general function will be of use in other contexts. In a recent commentary celebrating the 25th anniversary of the publication of Romer (1990), Joshua Gans (2015) observes that the Romer growth model has not been as influential on the design of growth policy as might have been expected despite its enormous influence on the subsequent growth theory literature. The reason he identifies is that it abstracts away "some of the richness of the microeconomy that give rise to new ideas and also their dissemination" (Gans, 2015). By expanding the parameter space, our function allows for the inclusion of more of this richness, including the role that meta technologies such as deep learning can play in knowledge access and knowledge discovery but potentially other policy and institutional factors that affect knowledge access, discovery rates, and team formation as well.

6. Concluding Thoughts: A Coming Singularity?

We developed this paper upon a number of prior ideas. First, the production of new knowledge is central to sustaining economic growth (Romer, 1990, 1993). Second, the production of new ideas is fundamentally a combinatorial process (Weitzman, 1998). Third, given this combinatorial process, technologies that predict what combinations of existing knowledge will yield useful new knowledge hold out the promise of improving growth prospects. Fourth, breakthroughs in AI represent a potential step change in the ability of algorithms to predict what knowledge is potentially useful to researchers and also to predict what combinations of existing knowledge will yield useful new discoveries (LeCun et al., 2015).

In a provocative recent paper, William Nordhaus (2015) explored the possibilities for a coming "economic singularity," which he defines as "[t]he idea . . . that rapid growth in computation and artificial intelligence will cross some boundary or singularity after which economic growth will accelerate sharply as an ever-accelerating pace of improvements cascade through the economy." Central to Nordhaus' analysis is that rapid technological advance is occurring in a relatively small part of the economy (see also Aghion, Jones, and Jones, 2017). To generate more broadly based rapid growth, the products of the new economy need to substitute for products on either the demand- or supply-side of the economy. His review of the evidence – including, critically, the relevant elasticities of substitution – leads him to conclude that a singularity through this route is highly unlikely.

However, our paper's analysis suggests an alternative route to an economic singularity – a broad-based alteration in the economy's knowledge production function. Given the centrality of new knowledge to sustained growth at the technological frontier, it seems likely that if an economic singularity were to arise, it would be because of some significant change to the knowledge production function affecting a number of domains outside of information technology itself. In a world where new knowledge is the result of combining existing knowledge, AI technologies that help ease needle-in-the haystack discovery challenges could affect growth prospects, at least along the transition path to the steady state. It doesn't take an impossible leap of imagination to see how new meta technologies such as AI could alter – perhaps modestly, perhaps dramatically – the knowledge production function in a way that changes the prospects for economic growth.

References

Acemoglu, Daron and Pascual Restrepo (2017),The Race Between Machine and Man: Implications of Technology for Growth, Factor Shares and Employment," NBER Working Paper 22252.

Aghion, Phillipe, Benjamin Jones, and Charles Jones (2017), "Artificial Intelligence and Economic Growth," paper presented at the NBER Conference on Research Issues in Artificial Intelligence, Toronto, September, 2017.

Agrawal, Ajay and Avi Goldfarb (2008), "Restructuring Research: Communication Costs and the Democratization of University Innovation," *American Economic Review*, 98(4): 1578-1590.

Agrawal, A., D. Kapur, D and J. McHale (2008), "How to Spatial and Social Proximity Affect Knowledge Flows: Evidence from Patent Data," *Journal of Urban Economics*, 64: 258-269.

Agrawal, A., I. Cockburn, J. McHale, (2006), "Gone But Not Forgotten: Knowledge Flows, Labor Mobility, and Enduring Social Relationships," *Journal of Economic Geography*, 6(5): 571-591.

Agrawal, A., Avi Goldfarb, and Florenta Teodordis (2016), "Understanding the Changing Structure of Scientific Inquiry," *American Economic Journal: Applied Economics*, 8(1): 100-128.

Angermuller, Christof, Tanel Pärnamaa, Leopold Parts, and Oliver Stegle (2016), "Deep Learning for Computational Biology," *Molecular Systems Biology,* 12 (878): 1-16.

Arthur, Brian W. (2009), *The Nature of Technology: What it is and How it Evolves*, Penguin Books, London.

Bloom, Nicholas, and Charles Jones, John Van Reenen, and Michael Webb (2017), "Are Ideas Getting Harder to Find?" Stanford University.

Bornmann, Lutz and Rüdiger Mutz (2015), "Growth Rates of Modern Science: A Bibliometric Analysis Based on the Number of Publications and Cited References," *Journal of the Association for Information Science and Technology*, 66(11): 2215-2222.

Bresnahan, Timoty, and Manuel Trajtenberg (1995), "General Purpose Technologies: Engines of Economic Growth," *Journal of Econometrics*, 65: 83-108.

Catalini, Christian (2017), "Microgeography and the Direction of Inventive Activity," *Management Science*, 10.1287/mnsc.2017.2798.

Cockburn, Iain, Rebecca Henderson, and Scott Stern (2017), "The Impact of Artificial Intelligence on Innovation," paper presented at the NBER Conference on Research Issues in Artificial Intelligence, Toronto, September, 2017.

Cowen, Tyler (2011), *The Great Stagnation: How America Ate All the Low-Hanging Fruit of Modern History, Got Sick, and Will (Eventually) Feel Better*, Dutton, Penguin Group, New York.

Furman, Jeffrey, and Scott Stern (2011), "Climbing atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research," *American Economic Review,* 101: 1933-1963.

Gans, Joshua (2015), "The Romer Model Turn 25" Digitopoly Blog, Available at: https://digitopoly.org/2015/10/03/the-romer-model-turns-25/ (accessed August 21, 2017).

Goh, Garrett, Nathan Hodas, and Abhinav Vishnu (2017), "Deep Learning for Computational Chemistry," *Journal of Computational Chemistry*, 38(16): 1291-1307.

Gordon, Robert (2016), *The Rise and Fall of American Growth: The U.S. Standard of Living Since the Civil War*, Princeton University Press, Princeton, NJ.

Grilliches, Zvi (1957), "Hybrid Corn: An Exploration in the Economics of Technical Change," *Econometrica*, 25(4): 501-522.

Helpman, Elhanan (1998), "Introduction," in Elhanan Helpman editor, *General Purpose Technologies and Economic Growth*, The MIT Press, Cambridge, MA.

Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson (1993), "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 108(3): 577-598.

Jones, Benjamin (2009), "The Burden of Knowledge and the 'Death of the Renaissance Man': Is Innovation Getting Harder," *Review of Economics and Statistics*, 76(1): 283-317.*Economy*, 103(4): 759-784.

Jones, Charles (2005), "Growth and Ideas," *Handbook of Economic Growth, Volume 1B*, Phillipe Aghion and Steven Durlauf, editors, Elsevier.

LeCun, Yann, Yoshua Benigo, and Geoffrey Hinton (2015), "Deep Learning" *Nature*, 521: 436-444.

Mamoshina, Polina, Armando Vieira, Evgeny Putin, and Alex Zhavoronkov (2016), "Applications of Deep Learning in Biomedicine," *Molecular Pharmaceutics*, 13: 1445-1454.

Mokyr, Joel (2002), *The Gifts of Athena: Historical Origins of the Knowledge Economy*, Princeton University Press, Princeton, NJ.

Mokyr, Joel (2017), """The Past and Future of Innovation: Some Lessons from Economic History," paper presented at the NBER Conference on Research Issues in Artificial Intelligence, Toronto, September, 2017.

Mullainathan, Sendhil, and Jann Spiess (2017), "Machine Learning: An Applied Econometric Approach," *Journal of Economic Perspectives*, 31(2): 87-106.

National Science and Technology Council (2011), "Materials Genome Initiative for Global Competitiveness," Washington D.C.

Nielsen, Michael (2012), *Reinventing Discovery: The New Era of Networked Science*, Princeton University Press, Princeton, NJ.

Nordhaus, William (2015), "Are We Approaching an Economic Singularity? Information Technology and the Future of Economic Growth," NBER Working Paper 21547.

Romer, Paul (1990), "Endogenous Technical Change," *Journal of Political Economy*, 94: S71-S102.

Romer, Paul (1993), "Two Strategies for Economic Development: Using and Producing Ideas," *Proceedings of the World Bank Annual Conference on Development Economics, 1992.* The World Bank, Washington D.C.

Romer, Paul (2008), "Economic Growth" published in *The Concise Encyclopaedia of Economics*, Library of Economic Liberty, available at: http://www.econlib.org/library/Enc/EconomicGrowth.html.

Rosenberg, Nathan (1998), "Chemical Engineering as a General Purpose Technology," in Elhanan Helpman editor, *General Purpose Technologies and Economic Growth*, The MIT Press, Cambridge, MA.

Weitzman, Martin (1998), "Recombinant Growth," *Quarterly Journal of Economics*, 113: 331-360.

Zeira, Joseph (1998), "Workers, Machines, and Economic Growth," *Quarterly Journal of Economics*, 113(4): 1091-1117.

Figure 1. Romer/Jones and Combinatorial-Based Knowledge Production Functions

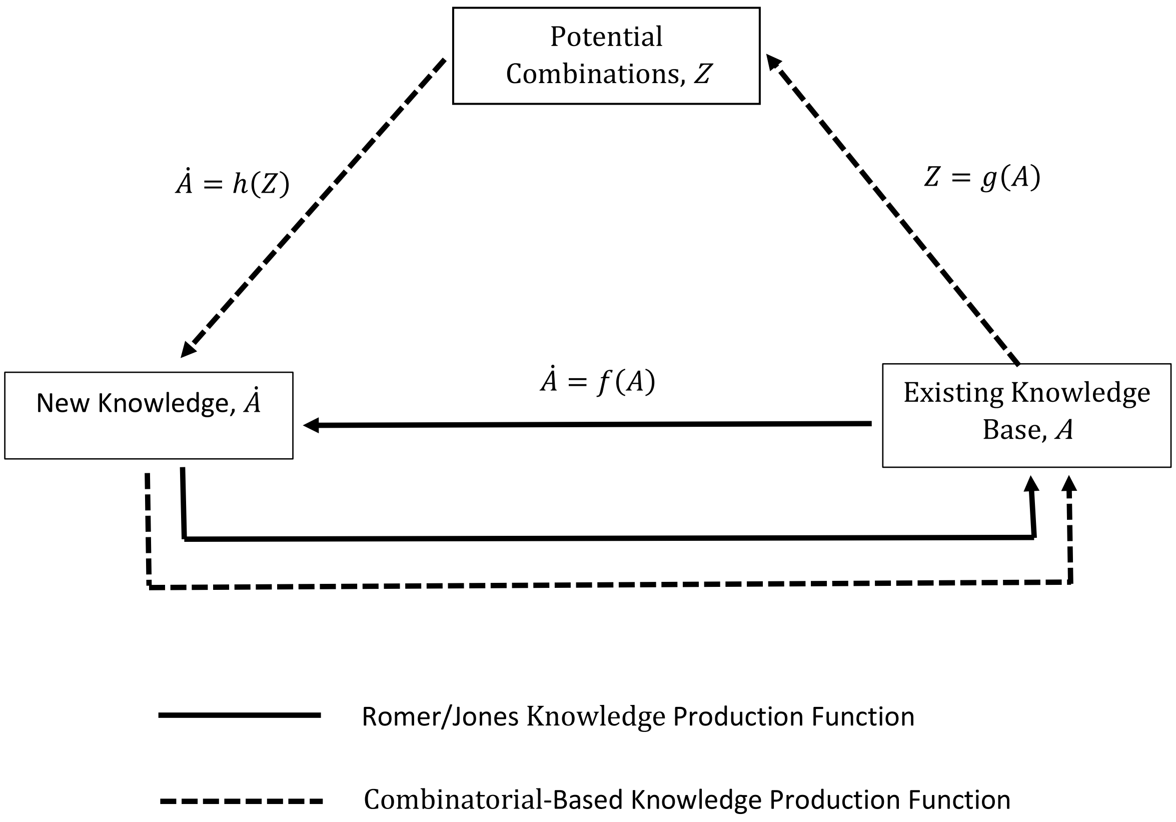

Figure 2. Relationships Between New Knowledge Production, θ, and ϕ

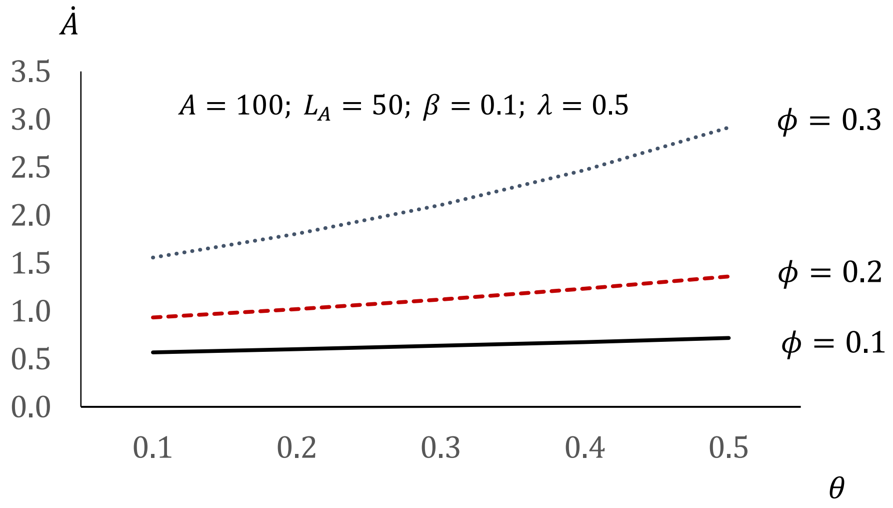

Figure 3. Example of How the Distribution of Team Size Varies with η

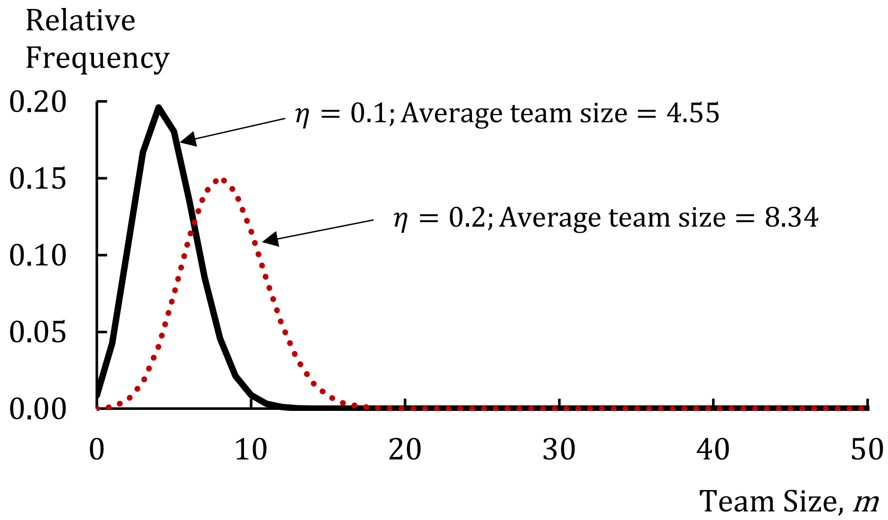

Figure 4. Team Knowledge Production and Team Size

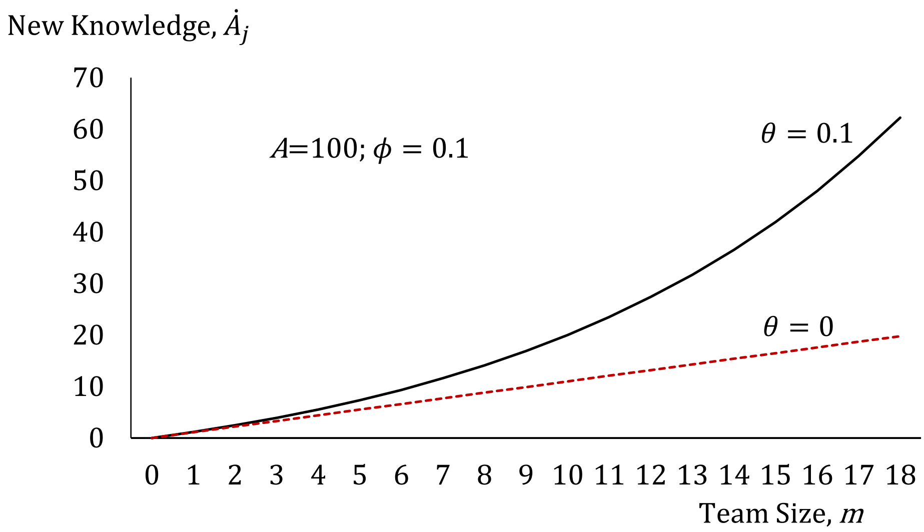

Figure 5. Relationships Between New Knowledge Production, η, and θ.

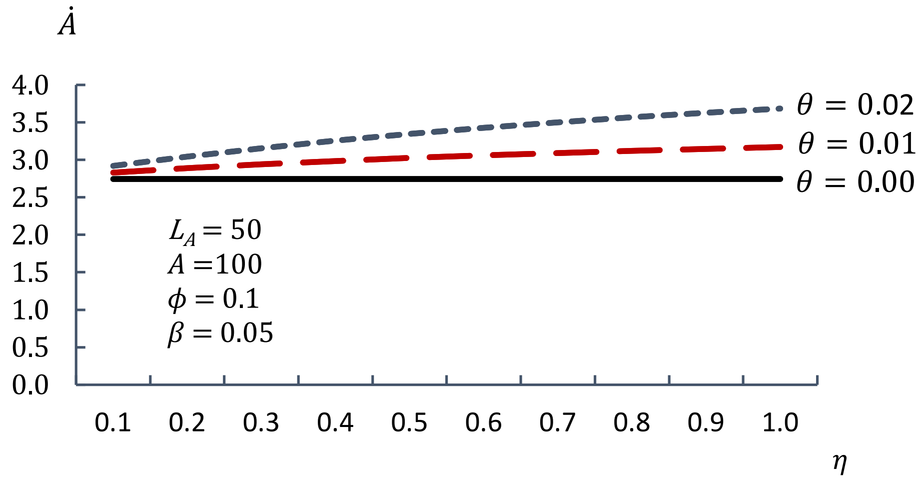

Figure 6. Relationships between Meta Ideas, Meta Technologies, and General Purpose Technologies