Paul Romer: Ideas, Nonrivalry, and Endogenous Growth*

Charles I. Jones†
Stanford University, Stanford CA 94305-5015, USA
chad.jones@stanford.edu

Abstract
In 2018, Paul Romer and William Nordhaus shared the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel. Romer was recognized “for integrating technological innovations into long-run macroeconomic analysis”. This article reviews his prize-winning contributions. Romer, together with others, rejuvenated the field of economic growth. He developed the theory of endogenous technological change, in which the search for new ideas by profit-maximizing entrepreneurs and researchers is at the heart of economic growth. Underlying this theory, he pinpointed that the nonrivalry of ideas is ultimately responsible for the rise in living standards over time.

Keywords: Economic growth; endogenous growth theory; ideas; nonrivalry; technical change
JEL classification: O3; O4

I. Introduction
When Paul Romer began working on economic growth in the early 1980s, a conventional view among economists (e.g., in the models taught in graduate school) was that productivity growth could not be influenced by anything in the rest of the economy. As in Solow (1956), economic growth was exogenous. Other models had been developed in the 1960s, as discussed further below, but these failed to capture widespread attention. Romer developed endogenous growth theory, emphasizing that technological change is the result of efforts by researchers and entrepreneurs who respond to economic incentives. Anything that affects their efforts, such as tax policy, basic research funding, and education, for example, can potentially influence the long-run prospects of the economy.

Romer’s fundamental contribution is his clear understanding of the economics of ideas and how the discovery of new ideas lies at the heart of economic growth. His 1990 paper is a watershed (Romer, 1990a). It stands
as the most important paper in the growth literature since Solow’s Nobel-recognized work. In this article, I review Romer’s prize-winning work, putting it into the context of the surrounding literature and providing a retrospective on how this research has led to the modern understanding of economic growth. The remainder of this introduction seeks to distill these insights into several pages, while the rest of the article delves more deeply into the details.

The history behind Romer’s path-breaking 1990 paper is fascinating and is engagingly presented by Warsh (2006). Romer had been working on growth for around a decade. The words in his 1983 dissertation and in Romer (1986) grapple with the topic and suggest that knowledge and ideas are important to growth. Of course, at some level, everyone knew that this must be true (and there is an earlier literature containing these words). However, what Romer did not yet have – and what no research had yet fully appreciated – was the precise nature of how this statement comes to be true. By 1990, though, Romer had it, and it is truly beautiful. One piece of evidence showing that he at last understood growth deeply is that the first two sections of the 1990 paper are written very clearly, with brilliant examples and precisely the right mathematics serving as the light switch that illuminates a previously dark room.

Here is the key insight: ideas – designs or blueprints for doing something or making something – are different from nearly every other good in that they are nonrival. Standard goods in classical economics are rival: as more people drive on a highway or require the skills of a particular surgeon or use water for irrigation, there are fewer of these goods to go around. This rivalry underlies the scarcity that is at the heart of most of economics and gives rise to the fundamental theorems of welfare economics.

Ideas, in contrast, are nonrival: as more and more people use the Pythagorean theorem or the Java programming language or even the design of the latest iPhone, there is not less and less of the idea to go around. Ideas are not depleted by use, and it is technologically feasible for any number of people to use an idea simultaneously once it has been invented. Consider oral rehydration therapy, one of Romer’s favorite examples. Until recently, millions of children died of diarrhea in developing countries. Part of the problem is that parents, seeing a child with diarrhea, would withdraw fluids. Dehydration would set in, and the child would die. Oral rehydration therapy is an idea: dissolving a few inexpensive minerals, salts, and a little sugar in water in just the right proportions produces a solution that rehydrates children and saves their lives. Once this idea was discovered, it could be used to save any number of children every year – the idea (the chemical formula) does not become increasingly scarce as more people use it.
How does the nonrivalry of ideas explain economic growth? The key is that nonrivalry gives rise to increasing returns to scale. The “standard replication argument” is a fundamental justification for constant returns to scale in production. If we wish to double the production of computers from a factory, one way to do this is to build an equivalent factory across the street and populate it with equivalent workers, materials, and so on. That is, we replicate the factory exactly. This means that production with rival goods is, at least as a useful benchmark, a constant returns process.

What Romer stressed is that the nonrivalry of ideas is an integral part of this replication argument: firms do not need to reinvent the idea for a computer each time a new computer factory is built. Instead, the same idea (i.e., the detailed set of instructions for how to make a computer) can be used in the new factory, or indeed in any number of factories, because it is nonrival. Because there is constant returns to scale in the rival inputs (the factory, workers, and materials), there is therefore increasing returns to the rival inputs and ideas taken together: if you double the rival inputs and the quality or quantity of the ideas, then you will more than double total production.

Once you have increasing returns, growth follows naturally. Output per person then depends on the total stock of knowledge; the stock does not need to be divided up among all the people in the economy. Contrast this with capital in a Solow model. If you add one computer, you make one worker more productive. If you add a new idea – think of the computer code for the first spreadsheet, or a word processor, or even the Internet itself – you can make any number of workers more productive. With nonrivalry, growth in income per person is tied to growth in the total stock of ideas (i.e., an aggregate) not to growth in ideas per person.

It is very easy to get growth in an aggregate in any model, even in a Solow model, because of population growth. More autoworkers produce more cars. In a Solow model, this cannot sustain per capita growth because we need growth in cars per autoworker. However, according to Romer, this is not the case: more researchers produce more ideas, which ultimately makes everyone better off because of nonrivalry. Throughout history – 25 years, 100 years, or even 1,000 years – the world is characterized by substantial growth, both in the total stock of ideas and in the number of people making them. Because ideas are nonrival, this is all that is required for sustained growth in living standards.

Finally, the increasing returns associated with nonrivalry means that a perfectly competitive equilibrium with no externalities will not exist and cannot decentralize the optimal allocation of resources. Instead, some departure is necessary. Romer emphasized that both imperfect competition and externalities to the discovery of new ideas are likely to be important. Imperfect competition provides the profits that incentivize entrepreneurs to
innovate. Later inventors and researchers benefit from the insights of those who came before.

Research on economic growth has been monumentally influenced by Romer’s contributions. While it can be difficult to adequately compensate for the knowledge spillovers, the 2018 Nobel Prize in economics is a well-deserved reward.

The remainder of this article is laid out as follows. Section II discusses Romer’s early contributions to growth, especially Romer (1986). Section III turns to the key insights of the 1990 paper. Section IV provides a brief survey of the many research directions that this paper opened up. Finally, Section V steps back and considers precisely how Romer’s insights lead to an understanding of sustained exponential growth. Section VI concludes.¹

II. Romer (1986) and the Explosion of Growth Research

The first key contribution that Romer made to the study of economic growth was to remind the profession of the ultimate importance of this topic. At a time when much of macroeconomics was devoted to studying inflation and unemployment, Romer emphasized the centrality of questions such as: “what determines the long-run rate of economic growth in living standards?” This reminder came in the form of his 1983 dissertation (Romer, 1983) and the key growth publication it led to (Romer, 1986). The substantive contribution of that paper was to build a model in which the long-run growth rate was determined endogenously, and to highlight that, because of externalities, the equilibrium growth rate might be lower than is optimal. In this way, Romer was a key founder of what came to be known as endogenous growth theory.

While there is substantial merit to the details of this work, one of its most important contributions lay in reintroducing economic growth as an active field of study to the economics profession. As one piece of supporting evidence, I searched for the topics “economic growth” and “endogenous growth theory” in Microsoft Academic.² Between 1973 and 1983, 49 papers on economic growth that received more than 1,000 citations were published. Between 1986 and 1996, the number rose more than five-fold, to 266.

Of course, there are many reasons for the explosion of research on economic growth during this period. In addition to Romer’s early work, Bob Lucas’ Marshall Lecture (Lucas, 1988), delivered by one of

¹The scientific background report prepared by the Prize Committee (2018) contains an excellent overview of the contributions of both Romer and Nordhaus, and how they are linked.
²Microsoft Academic (https://preview.academic.microsoft.com) allows you to sort by the number of citations, while Google Scholar does not.

Romer’s dissertation advisers at Chicago, was supremely influential, with its elegant modeling, beautiful prose, and its reminder that we care not only about growth rates but also about how enormous differences in living standards can persist across countries. Another important factor was the newly available data on cross-country economic performance. Summers and Heston (1984, 1988) used the International Comparisons Project to construct internationally comparable data on more than 100 countries back to 1950. Maddison (1982) made his career by looking back even further in time, to the year 1500, to shed light on how growth changed before and after the Industrial Revolution. These data featured in several prominent empirical papers, such as Baumol (1986), Barro (1991), Mankiw et al. (1992), and Barro and Sala-i-Martin (1992), all of which complemented, inspired, and pushed forward the economic theories that were being developed.

The confluence of new theories and new data made this an extremely exciting subject and, for a time, it seemed like half the profession was working on economic growth.3 One of the key theoretical developments was the re-discovery of the “AK” growth model. The use of the word “re-discovery” is intentional and highlights an important feature of the growth literature: many of the insights of new growth theory were anticipated, often to a great extent, by insights developed in the growth literature of the 1950s and 1960s. Perhaps the best illustration of this point is the AK model. Early versions of this model were developed by Von Neumann (1945) and Frankel (1962), and Solow (1956) explicitly analyzes this special case. In the 1980s, Lucas (1988), Jones and Manuelli (1990), and Rebelo (1991) fit this mold, as does Romer (1986), albeit in a slightly broader interpretation.

In its simplest form, the AK model can be expressed as

\[ Y_t = AK_t \] (1)

and

\[ K_t = sY_t - \delta K_t, \] (2)

where A is an exogenous and constant productivity parameter and s is an exogenous, constant investment rate. In this set-up, K is interpreted as physical capital, but in Romer (1986) K was interpreted as knowledge, and in Lucas (1988) it was replaced by human capital.

Putting these equations together,\[ g_Y \equiv \frac{\dot{Y}_t}{Y_t} = sA - \delta. \] (3)

The growth rate of the economy is endogenously determined by fundamental parameters of the economic environment. In this example, a permanent increase in the investment rate \( s^4 \) will permanently raise the growth rate of the economy.

The AK insight (i.e., a linear differential equation of the form \( \dot{K}_t = sAK_t \) could generate an endogenous exponential growth rate) was the theoretical spark that lit a thousand lamps. It led to the development of a host of different growth models and, together with the widespread availability of data that allowed income comparisons across countries and over time, fed the explosion of the new growth literature.

III. Romer (1990)

The watershed contribution highlighted by the Prize Committee is Romer (1990a), and this paper makes three key contributions:

1. it identifies the nonrivalry of ideas as crucial to economic growth;
2. it highlights the role of profit-maximizing entrepreneurs and imperfect competition;
3. it places the key AK linearity in the idea production function.

The Nonrivalry of Ideas

It is the first insight that is truly fundamental and leads to a deep, intuitive understanding of economic growth. Many others before Romer had appreciated that ideas are special; some of these will be discussed further below. Indeed, Romer himself, in his 1986 paper, writes extensively about knowledge. However, in the first eight pages of his 1990 paper, Romer put the nonrivalry of ideas on center stage in a way that no other paper had done before. Everything of ultimate importance follows from that insight.

Whereas Solow divided the world into capital and labor, Romer makes a more basic distinction: between ideas, on the one hand, and everything else (call them “objects”) on the other.\(^5\) Objects are the traditional goods

\(^4\)Note that \( s \) could be an investment rate in physical capital, human capital, or knowledge, depending on the interpretation of \( K \), and it is typically endogenized and, in turn, depends on policy.

\(^5\)Romer (1993) introduced this “ideas versus objects” language.
that appear in economics, including capital, labor, human capital, land, highways, lawyers, a barrel of oil, a bushel of soybeans, or even a certain volume of clean air. An idea is a design, a blueprint, or a set of instructions for starting with existing objects, and transforming or using them in some way that generates either more output or more utility. Examples include calculus, the recipe for a new antibiotic, Beethoven’s Fifth Symphony, the design of the latest quantum computer, and the technology that converts my key strokes into electric signals that show up as letters on my computer screen.

What is the distinction between ideas and objects? Objects are rival: one person’s use of an object precludes the simultaneous use of the object by others. In contrast, ideas are nonrival: an idea can be used simultaneously by any number of people. If I use a machine, or a gallon of gasoline, or an hour of a doctor’s time, others cannot use the same thing; this is the sense in which objects are scarce. Indeed, economics is precisely the study of the allocation of such scarce resources. But any number of people can simultaneously use calculus and any number of assembly lines can use the design for the latest computer chip. New ideas are scarce in that we are always looking for better ideas, but existing ideas are not scarce in the same way that a building or a doctor is – and this has far-reaching consequences for economics.

This simple distinction is at the heart of Romer’s theory of economic growth. Solow’s hypothesis was that the endogenous accumulation of objects (capital) could explain growth. This hypothesis certainly had empirical credibility: a large part of the difference between the world today and the world in 1800 is the vast amount of capital we now have at our disposal. However, what Solow showed is that the accumulation of objects in a neoclassical setting runs into diminishing returns and cannot, by itself, sustain exponential growth. However, ideas are different. Nonrivalry means that the accumulation of ideas is not ruled out as a theory of growth by Solow’s famous result.

To see why, Romer highlighted that the nonrivalry of ideas means that production is characterized by increasing returns to scale. This reasoning featured “precisely the right mathematics serving as the light switch”, as mentioned above. Let $A$ denote an index of the level of technology (e.g., the stock of knowledge or the number of ideas) and let the vector $X$ denote all the other – rival – inputs into production. For example, suppose we are producing a new state-of-the-art antibiotic. $A$ denotes the knowledge in the economy for making pharmaceuticals, and one piece of that knowledge is the newly invented instructions for manufacturing the new antibiotic. Perhaps it requires some extremely rare and special ingredients, and can only be assembled by the most expert chemists. The ingredients, the chemists, the laboratory equipment, etc., are all part of the vector $X$. © The editors of The Scandinavian Journal of Economics 2019.
Indeed, the presence of human capital or costly training, or even inputs in finite supply, do not change Romer’s basic insight; these are just part of the list of rival inputs. Finally, let $Y$ denote the number of doses of the antibiotic that are produced. Then the basic production function for the antibiotic is

$$Y = F(A, X).$$  \hfill (4)

Given a certain quantity of objects $X$ and the set of knowledge $A$, the function $F(\cdot)$ tells us how many doses of the antibiotic are produced.

Now consider the properties of $F(\cdot)$. In particular, suppose we wish to double the production of antibiotics. The *standard replication argument* tells us that one way to do this is by doubling the objects ($X$): if we set up a second lab that is identical in every way – with identical chemists, identical collections of the rare ingredients, and identical equipment – then we should be able to produce an identical number of doses, so that the second lab will double production. In other words, for some $\lambda > 1$,

$$F(A, \lambda X) = \lambda Y.$$  \hfill (5)

That is, there is constant returns to scale to the objects in production. But equation (5) makes clear that we have implicitly used the nonrivalry of ideas in the replication argument: the new lab can use the same set of instructions for building the new antibiotic. In particular, we do not have to re-invent the idea. Once the instructions have been invented, they can be used in one lab, two labs, or any number of labs simultaneously.

This means that, as long as more knowledge is useful, if we double the objects and double the knowledge as well, we will more than double the output:

$$F(\lambda A, \lambda X) > F(A, \lambda X) = \lambda Y.$$  \hfill (6)

That is, production is characterized by *increasing returns to scale*. Because of the nonrivalry of ideas, there is constant returns to objects and therefore increasing returns to objects and ideas taken together. Each idea only needs to be invented once and then it is technologically feasible for the idea to be used by any number of people or firms simultaneously and repeatedly.

Of course, just because it is feasible to do so does not mean that this will always happen. Intellectual property rights, such as patents, might restrict the use of ideas, and ideas can be forgotten so that they might need to be reinvented, such as concrete in ancient Rome. These are important considerations, but they do not change the essential point.

We now reach a crossroads in our narrative, and there are different ways in which we can proceed. First, of course, one would like to know precisely how the increasing returns that results from nonrivalry leads to the possibility of sustained exponential growth. A loose guide to the answer is

that Solow had constant returns to objects and therefore diminishing returns to capital, and this is what dooms growth in a neoclassical model. Now that we have increasing returns to objects and ideas together, it is not clear that diminishing returns to – well, to what? – dooms us in the same way. More care is required to make this intuition precise, and we will come back to this absolutely critical point a little later. Before that, however, we follow Romer (1990a) closely and discuss the role of imperfect competition and profits in incentivizing innovation.

The Role of Imperfect Competition

With increasing returns to scale, inputs cannot each be paid their marginal product in some perfectly competitive economy, as there is not enough output to go around. As an application of Euler’s theorem, the constant returns to rival objects implies that paying each object its marginal product will exactly exhaust output, so then nothing is left to pay for ideas. As a result, if ideas are to receive some compensation, at least one other factor must be paid less than its marginal product. Conceptually, this means that a perfectly competitive equilibrium with no externalities will not exist. The growth literature developed two basic ways to handle this problem: externalities and imperfect competition. Both imply that at least one factor will be paid less than its (social) marginal product, and therefore the equilibrium will not generally be efficient.

At a basic level, this point is profound. In particular, it implies that a pure perfectly competitive allocation cannot decentralize the optimal allocation of resources, and there is a role for institutions other than basic property rights and the potential for distortionary taxes to improve the allocation of resources. An unregulated economy is no longer the “best of all possible worlds”.

Romer (1990a) imports the models of imperfect competition developed by Dixit and Stiglitz (1977) and Ethier (1982) into growth theory, and the result is elegant, both conceptually and from the standpoint of the economics of the real world. A key to making these models applicable was the recognition that ideas, while nonrival, are not pure public goods. Recall that the definition of a pure public good is something that is both nonrival and nonexcludable. While nonrivalry is a property of the economic environment, excludability is a function of institutions and the decisions that societies make. Institutions such as the patent system – or even just trade secrets – can allow ideas to be partially excludable, at least for a certain period of time.

Using this insight, the decentralized allocation in the Romer model features entrepreneurial researchers who hunt for new ideas because of the financial rewards that can be earned by innovating. Successful innovators
are awarded a patent that gives them the exclusive right to produce with their invention. This allows them to charge a mark-up over marginal cost, subject to imperfect competition, and to earn the profits that ultimately serve as the carrot that motivates the search for new ideas.

Relative to the other growth models of the time, this framework had a ring of truth that seemed to be missing from the literature: growth occurs because of innovation from private entrepreneurs, who are incentivized by the prospect of earning profits from their ideas. The contrast with other models is worth appreciating. Solow (1956) obtained growth only by assuming exogenous technological change in which productivity improvements occur without anyone taking any actions. AK models based on physical capital, such as Rebelo (1991), assume that capital actually does not face any diminishing returns, at least in some sector, and there is no mention of new technologies being developed. Similarly, Lucas (1988) has individuals accumulating human capital, but again there are no inventions or new technologies. Finally, the learning-by-doing models of Arrow (1962a) and Frankel (1962) admit that technologies can improve but this is entirely an accidental by-product of the production process itself rather than a result of intentional efforts to innovate.

Constructing a model in which long-run growth is driven by the active search for new ideas by entrepreneurs who are motivated by profits was the second key contribution of Romer (1990a).

**Linearity in the Idea Production Function**

The third contribution of Romer (1990a) was to put the AK structure in the idea production function. Essentially, this worked like the simple model given earlier in Section II, where, as in Romer (1986), the “K” is interpreted as knowledge rather than as physical capital. In fact, to avoid the confusion between knowledge and capital, Romer (1990a) follows the notation we have adopted above and denotes the stock of knowledge by \( A \), following tradition and reserving \( K \) for physical capital.

With \( A_t \) as the stock of ideas at date \( t \), the flow of new ideas is denoted \( \dot{A}_t \equiv \frac{dA_t}{dt} \). Romer assumes the production of new ideas is given by

\[
\dot{A}_t = \theta H_{A_t} A_t,
\]

where \( H_{A_t} \) is the amount of human capital devoted to research and \( \theta > 0 \) is a parameter governing the productivity of research. Notice that there are increasing returns to \( H_A \) and \( A \) in this idea production function, and Romer uses externalities to handle the increasing returns in this case: individual researchers are small and take the time path of \( A_t \) as exogenously given. The new ideas they produce lead to positive knowledge spillovers that raise the productivity of future researchers.
The economy is endowed with a constant $\bar{H}$ units of human capital that can be used to produce either consumption goods or ideas. Letting $s_t \equiv H_{A_t}/\bar{H}$ be the fraction of the stock of human capital that is devoted to research, we can rewrite equation (7) as

$$\frac{\dot{A}_t}{A_t} = \theta s_t \bar{H}. \quad (8)$$

Profits incentivize people to search for new ideas, and wages in production encourage human capital to make consumption goods. The market equilibrium then determines a value for $s_t$, which turns out to be constant in the long run. Then equation (8) pins down the long-run growth rate of the economy. This growth rate is a function of the productivity parameter $\theta$, the amount of human capital in the economy $\bar{H}$, and the equilibrium fraction of human capital allocated to research $s$. Taxes, research subsidies, expiring patents, and other features of the economy can influence the long-run growth rate by affecting $s$ through the market for entrepreneurs.

In this way, the three key ingredients – the nonrivalry of ideas, the profit motive of imperfect competition, and putting the key linear differential equation in the idea production function – combine to generate the fundamental insights of Romer (1990a).

Looking Backwards

As I mentioned earlier in this article, it is remarkable how many of the discoveries of new growth theory were actually re-discoveries of old growth theory. For reasons we will leave to the historians, many of the insights from the 1960s were forgotten by the heyday of the 1980s and 1990s. It is worth reviewing the contributions from the 1960s in light of Romer’s 1990 paper in order to see both the similarities and the differences. In this task, we are helped appreciably by the fact that Romer himself is careful to discuss the earlier literature.  

Arrow (1962b) views invention as the production of information, a commodity he describes as having “peculiar attributes” in that information should be available free of charge (as that is its marginal cost) although this would provide no incentive for investment in research. Uzawa (1965) considers optimal technical change in a setting where the growth rate of technology is determined by the fraction of the labor force working in education. Shell (1966) clearly recognizes the failure of models of perfect competition to deliver optimal resource allocation in the presence of ideas.

\footnote{See also Romer (1987, 1990b, 2019).}
and explicitly assumes that knowledge is a pure public good. Phelps (1966) studies the “golden rule” for research investment that maximizes steady-state consumption.

The two papers closest to Romer (1990a) are Nordhaus (1969a) and Shell (1973). Nordhaus (1969a) notes the public-good characteristics of knowledge and explicitly introduces “a multitude of temporary little monopolies on information” based on fixed-life patents. His model generates a steady state with exponential growth at a rate that is proportional to the population growth rate. From the standpoint of Romer (1990a), this was viewed as something of a limitation, though I will come back to this point in Section IV.7

Each of the papers mentioned up until now notes that knowledge is special, and they often speak of a public good. They also assume a production function that exhibits increasing returns when $A$ is included as a factor of production, such as $Y = AK^\alpha L^{1-\alpha}$. However, with one exception, the papers do not clearly connect the nonrivalry of knowledge to the increasing returns they have implicitly assumed. In fact, this particular Cobb–Douglas specification was so common in the “exogenous $A$” case that the production function was often assumed to take this form automatically, rather than justified on first principles. It is an interesting accident that the production function of exogenous growth models, if naively used in an attempt to endogenize $A$, leads to the right destination even if the nuances of nonrivalry are not fully appreciated. Later researchers, such as Griliches (1979), who thought carefully about where to put ideas in a production function, in some ways chose the natural assumption of treating knowledge as another form of rival capital, putting the stock inside the constant returns as in $Y = A^\beta K^\alpha L^{1-\alpha-\beta}$.

The clear exception to this statement is Shell (1973). Shell states that technical knowledge is a public good and uses the replication argument to justify constant returns to capital and labor, and therefore increasing returns to capital, labor, and technical knowledge. He further appreciates that this means that a perfectly competitive model cannot support innovation because payments to conventional factors would exhaust output. Despite this promising start, the models in Shell (1973) are less satisfying. They do not generate stable exponential growth and are not fully worked out. Shell is unable to provide a sharp characterization of economic growth, beyond noting that the growth rate of productivity is projected to increase over time under some conditions.

7Phelps (1966) also derived the result, discussed in detail below, that long-run growth in per capita income is driven by population growth. Judd (1985) obtains this result in an expanding-variety set-up. The learning-by-doing models of Arrow (1962a) and Sheshinski (1967) contain a similar result.

The “new growth” insights of old growth theory were impressive, culminating with Nordhaus (1969a) and Shell (1973). Nevertheless, for reasons that remain unclear, these papers represented the end of an era of active research on endogenous technical change rather than a beginning.¹

IV. The Growth Literature after Romer (1990)

The explosion of the new growth literature was already well under way when Romer (1990a) was being presented in conferences in 1988. In this section, we highlight several different directions in which the literature evolved.

Probably the most important is Schumpeterian growth theory, associated most prominently with Aghion and Howitt (1992). Romer developed his theory in the context of the “love of variety” models of Dixit and Stiglitz (1977) and Ethier (1982), in which an idea is the discovery of how to produce a new good, such as the internal combustion engine, the laser, or a new antibiotic. In this setting, new inventions take a tiny share of the market from each existing producer while at the same time increasing the overall size of the market. The Schumpeterian model in Aghion and Howitt (1992), in contrast, is based on “quality ladders” in which a new innovation is a perfect substitute for some existing good, except it is of higher quality or can be produced more cheaply. Other important early contributions to Schumpeterian growth theory include Segerstrom et al. (1990) and Grossman and Helpman (1991a,b).

In Schumpeterian growth models, innovation is associated with creative destruction, in which the profit stream of a previous innovator is destroyed (or stolen) by the creation of a new innovator. Firms can be created and destroyed in the process of growth. In recent years, this dynamic has been explored extensively, allowing the growth literature to connect to a wealth of empirical facts from firm-level data. For examples, see Klette and Kortum (2004), Aghion et al. (2014, 2017), Akcigit and Kerr (2018), and Atkeson and Burstein (2019). Schumpeterian growth models also rely heavily on the nonrivalry of ideas and the increasing returns that it implies, although this is typically not emphasized.

Another important direction in which the growth literature expanded was backwards in time. Romer (1986) originally noted that if we look back over centuries rather than just decades, it appears that growth rates have

¹For example, as of January 2019, the Nordhaus paper had 297 citations and the Shell paper had 143 citations, both from Google Scholar. Nordhaus also has a book-length treatment published in the same year (Nordhaus, 1969b) but it is focused on other questions related to inventive activity rather than on the source of sustained exponential growth.

been increasing over time. It is interesting to ask to what extent models that have been designed to explain growth in the past century can help us to understand growth since the Industrial Revolution, and even before. Kremer (1993) first documented that the prediction of Romer (1990a) that growth rates should be increasing in the overall size of the economy (captured by $\bar{H}$ in the model and proxied by population in the data) actually holds up very well when one looks at economic growth over thousands of years. Subsequent research by Galor and Weil (2000), Jones (2001), Hansen and Prescott (2002), and Lucas (2002) explored the causes of growth over the very long run; see the review of unified growth theory in Galor (2005) for more details.

Much of Acemoglu’s work on the direction of technical change can be interpreted as adding a second dimension to Romer (1990a): what if researchers can invent two different kinds of ideas? In Acemoglu (1998), ideas can augment either skilled labor or unskilled labor. Acemoglu (2003) revisits the famous Uzawa theorem and asks if there is any reason why technical change would tend to be labor-augmenting instead of capital-augmenting. Acemoglu et al. (2012) consider environmental damage and study a world in which firms can invent “clean” or “dirty” technologies. Finally, Hemous and Olsen (2016) and Acemoglu and Restrepo (2018) study the effects of automation on economic growth, unemployment, and inequality.

The connections between economic growth and international trade have been explored in a wide range of papers. Krugman (1979) and Grossman and Helpman (1989, 1991a) were some of the early and very influential papers. Romer himself explored this topic in Rivera-Batiz and Romer (1991) and Romer (1994). More recent papers include Atkeson and Burstein (2010), Ramondo et al. (2016), and Arkolakis et al. (2018).

In the end, this all-too-brief review has omitted many other topics and lines of research that have been influenced by Romer’s work. When an author has been cited more than 80,000 times, I suppose this is inevitable.

V. Romer and Sustained Economic Growth

The substantive economic question at the heart of this branch of the growth literature is this: how do we understand sustained economic growth? How is it, for example, that US GDP per person has grown at a relatively steady

---

9Lee (1988) has an elegant model that combines a Romer-like production function for ideas with a Malthusian dynamic for population to predict that the growth rate of technology will itself grow exponentially over time.

rate of 2 percent per year for something like the past 150 years? In this section, I’d like to step back and provide an overview of the modern answer to this question – or at least the answer that I find most persuasive – and to show how this answer ties directly to Romer’s work.

Revisiting the Answer in Romer (1990)

The basic intuition behind sustained economic growth in Romer (1990a) goes as follows. Ideas are nonrival, so output per person depends on the overall stock of knowledge. Therefore, we need to understand why knowledge can grow at a constant exponential rate. This question in turn is answered by a key assumption: that the growth rate of knowledge is proportional to the amount of research effort (i.e., $H_A$ above). A constant amount of research effort therefore generates exponential growth in knowledge.

Stated this way, there are two important problems with Romer’s answer. First, it turns out that if you are willing to make the key assumption that constant research effort generates exponential growth in knowledge, the fact that ideas are nonrival then plays no essential role. Second, there is substantial empirical evidence suggesting that this assumption is not a good one.

To see the first problem, consider the following departure from Romer’s model, taken recently by Akcigit et al. (2016). Suppose, counterfactually, that knowledge, $A$, is completely rival, just like capital and labor, so that it enters inside the constant returns to scale piece of the production function:

$$Y_t = A_t^\theta K_t^\alpha L_t^{1-\alpha-\beta},$$

(9)

$$\frac{\dot{A}_t}{A_t} = \theta L_A = \theta \bar{s}_A \bar{L},$$

(10)

$$\dot{K}_t = \bar{s}_K Y_t - \delta K_t,$$

(11)

and

$$L_{Y_t} + L_A = \bar{L}.$$

(12)

In this set-up, $A$ is just like another form of rival capital, except it has a different production function that involves a linear differential equation. We assume that a constant fraction $\bar{s}_A$ of a fixed population is engaged

---

10I think of the growth literature as consisting of two main branches. (1) How do we understand frontier growth? (2) Why are some countries so much richer than others? While there is obviously some overlap between these two questions, it is helpful to think of them as conceptually distinct.
in producing new ideas, so that knowledge grows exponentially at rate
\( g_A \equiv \theta \bar{s}_A \bar{L} \). It is then straightforward to solve the rest of the model in order to see that \( y \equiv Y/L \) (and consumption per person) will exhibit steady-state growth at the rate\(^{11}\)

\[
g_y = \frac{\beta \theta \bar{s}_A \bar{L}}{1 - \alpha}.
\] (13)

In other words, if we are willing to assume that the idea production function involves a linear differential equation, then the nonrivalry of ideas becomes unnecessary and there is no need for increasing returns in the goods production function. In terms of the taxonomy of Romer’s contributions that I laid out in Section III, the presence of the third contribution (linearity in the idea production function) renders the other two contributions (nonrivalry of ideas and imperfect competition) inessential. But the inverse also holds: if we do away with the third contribution, the nonrivalry of ideas becomes much more important, a point that is explored in detail below.\(^{12}\)

The second problem is that there is ample evidence suggesting that Romer’s original specification of the idea production function is, from an empirical standpoint, flawed. As in equation (10), the original Romer formulation states that the growth rate of productivity is proportional to the amount of resources devoted to research. The problem with this formulation is easy to see: productivity growth rates are relatively stable over time, while the resources devoted to innovation show large exponential trends. Jones (1995) made this point using aggregate evidence. Ngai and Samaniego (2011) provide evidence at the industry level in the United States. In the presence of the growing resources devoted to research that we observe empirically in the past 50 or 100 years, the Romer set-up suggests that we should see large increases in growth rates. Put differently, if one adds population growth to Romer (1990a), then the model no longer generates a steady-state growth rate. Instead, growth rates explode over time.

\(^{11}\)To see this, let \( k \equiv K/L \) and \( a \equiv A/L \). Then \( y = k^\alpha a^\beta (1 - \bar{s}_K)^{1-\alpha-\beta} \) so that \( g_y = \alpha g_k + \beta g_a \). However, in the steady state, \( g_k = g_y \) and \( g_a = g_A \), as there is no population growth. Combining these equations gives the result.

\(^{12}\)The way I’ve stated this ignores another subtlety: the linear differential equation itself gives rise to dynamic increasing returns. For example, suppose \( \dot{A}_t / A_t = \bar{s}_L \) and \( Y_t = A_t^\beta [(1 - \bar{s})\bar{L}]^{1-\beta} \). These equations can be combined to yield \( Y_t = A_0^\beta \exp(\beta \bar{s}_L)[(1 - \bar{s})\bar{L}]^{1-\beta} \), which obviously exhibits increasing returns to scale. So it would not be correct to say that all increasing returns have been eliminated, and a perfectly competitive model with no externalities cannot decentralize the optimal allocation here. On a related point, one could naturally ask of such a model: “why does the idea production function look this way?” The nonrivalry of ideas would be one possible answer.
More recently, Bloom et al. (2019) have examined a host of evidence at different levels of aggregation confirming that the Romer specification of the idea production function is misguided. They look at a wide range of evidence, from Moore’s law for semiconductors, to agricultural innovations, to medical innovations that reduce cancer and heart disease mortality, to firm-level data. Wherever they look, they find that “research productivity” – defined as the ratio of the growth rate $\dot{A}_t/A_t$ to research effort – is declining rapidly. What was assumed to be constant by Romer (1990a) is, in fact, falling sharply in the data. The typical estimate in Bloom et al. (2019) suggests that research productivity declines at a rate of about 5 percent per year, meaning that the level of research productivity falls by half every 12 years.

The weaknesses highlighted in this section are serious and require remedy. However, in the next section, I show that dropping the empirically untenable assumption that productivity growth is proportional to the level of research addresses both problems. The augmented model highlights the essential contribution of Romer (1990a) and provides a sharp and elegant intuition for how sustained exponential growth is possible.

The Modern Romerian Answer

The modern version of Romer’s answer to the question of where exponential growth comes from can be explained quite simply. In particular, we take the basic Romer (1990a) set-up and change it only slightly, relaxing the assumption that constant research effort can generate constant exponential growth:

\begin{align*}
Y_t &= A_t^\sigma L Y_t \\
\dot{A}_t &= \theta L_{At} A_t^{-\beta} \\
L_{Yt} + L_{At} &= L_t = L_0 e^{nt}.
\end{align*}

Notice

\begin{align*}
L_{At} &= \bar{s} L_t.
\end{align*}

Here, equation (14) says that output is produced by a production function that has constant returns to objects (here just labor), and increasing returns to objects and ideas together. The parameter $\sigma$ measures the degree of increasing returns to scale in the goods production function. Notice

---

13The stylized model in this section is a simplified version of the models in Jones (1995), Kortum (1997), and Segerstrom (1998). As noted earlier, related models can be found in Phelps (1966), Nordhaus (1969a), Judd (1985), and Gomulka (1990).
that we are leaving out capital just to make the model as simple as possible. Including capital does not change anything in a substantitive way. The original Romer model also distinguished between human capital and unskilled labor, but we drop this distinction again for simplicity.

Equation (15) is the new production function for ideas. The parameter $\beta > 0$ captures the extent to which ideas – or, more precisely, “proportional improvements in productivity” – are becoming harder to find. Romer (1990a) had a similar production function, but with $\beta = 0$. The evidence discussed in the previous section suggested that $\beta > 0$ is more consistent with the data. Bloom et al. (2019) find values of $\beta \approx 1/4$ for Moore’s law and semiconductor production, $\beta \approx 1$ for firm-level data, and $\beta \approx 3$ for the aggregate economy.

Equation (16) is the resource constraint for labor – labor can be used to make either goods or new ideas. For reasons that will become clear below, we allow for population growth at an exogenous rate $n$.

Finally, equation (17) specifies that a constant fraction $\bar{s}$ of the population works to make ideas, so that $1-\bar{s}$ works to make goods. In richer models (including those cited in footnote 13), this allocation is endogenized and studied more carefully. Again, to get to the main point as quickly as possible, it is appropriate to specify a simple rule for the allocation.

Once the model is set up this way, the beauty of Romer (1990a) emerges naturally. First, notice that when we talk about economic growth, we are speaking of growth in consumption per person. In the model, this is $y \equiv Y/L$ and, from equation (14), this is given by

$$y_t = A_t^{\sigma} (1 - \bar{s}).$$  

The key insight of nonrivalry already shines in this equation: consumption per person is proportional to the overall stock of knowledge, $A$, raised to some power. This is because ideas are nonrival: each idea can raise each person’s consumption. Contrast this with a Solow model, where $y_t = k_t^{\alpha}$; that is, consumption per person is proportional to capital per person, not to the aggregate stock of capital. To increase the productivity of each person in the economy, you need to give each person a computer; however, you can increase the productivity of any number of people by inventing a single new idea. In other words, the essential role of the nonrivalry of ideas has already been embedded in the goods production function, exactly as in Romer (1990a).

Taking logs and derivatives of equation (18), we see that the growth rate of consumption per person is proportional to the growth rate of the overall stock of ideas: $g_y = \sigma g_A$. Therefore, the next step is to figure out the long-run growth rate of ideas.

To do this, consider the idea production function in equation (15), rewritten as
In order for the left-hand side of this equation, $\frac{\dot{A}_t}{A_t}$, to be constant, we need the right-hand side to be constant. This can only occur if the numerator and the denominator on the right-hand side grow at the same rate: $g_{L_A} = \beta g_A$. Solving this for $g_A$, we obtain
\begin{equation}
   g_A = \frac{g_{L_A}}{\beta} = \frac{n}{\beta}.
\end{equation}

The growth rate of the stock of ideas in the long run equals the growth rate of research effort deflated by the extent to which ideas are becoming harder to find, $\beta$. In the long run, research effort can only grow at the rate of growth of the overall population.

Finally, combining this last expression with $g_y = \sigma g_A$, we have
\begin{equation}
   g_y = \frac{\sigma n}{\beta} \equiv \gamma n.
\end{equation}

This is the key result that explains long-run growth. In the long run, the growth rate of consumption per person is the product of two terms, $\gamma$ and $n$. Here, $\gamma$ measures the overall degree of increasing returns in the economy and is itself equal to $\sigma$ (the static degree of increasing returns in goods production) deflated by $\beta$, the extent to which ideas are becoming harder to find. The long-run growth rate is then the product of the overall degree of increasing returns to scale and the rate at which the scale of the economy is itself growing, $n$.\(^{14}\)

Where does this result come from? It is tied intimately to the nonrivalry of ideas. Recall that the original lesson of Romer (1990a) was that nonrivalry implies increasing returns. However, increasing returns means that bigger is more productive. An economy with more people will, other things equal, have a larger number of researchers and therefore a larger number of ideas in the long run. Because these ideas are nonrival, a larger number of ideas raises everyone’s income, and an economy with more people will be richer.

An implication of this statement is that if two otherwise identical economies have different population growth rates, the economy with the higher rate of population growth will grow faster: the growth rate of

\(^{14}\)In my earlier work on this topic, I have typically written the idea production function in terms of the level change rather than in terms of the percentage change: $\dot{A}_t = \theta L_A,\dot{A}_t^\phi$. The two expressions are equivalent, with $\beta \equiv 1 - \phi$. Here, the Romer case corresponds to $\beta = 0$, which is then $\phi = 1$ in my earlier work. Similarly, the overall degree of increasing returns to scale is $\gamma \equiv (\sigma/\beta) = \sigma/(1 - \phi)$. The “$\phi$” approach makes clear the connection to increasing returns in that $1/(1 - \phi)$ measures the dynamic increasing returns associated with the idea production function. Both approaches have their merits.
researchers will be higher, so the growth rate of ideas will be higher, and this delivers a higher growth rate of income per person.

In this way, the relaxation of the empirically problematic assumption that $\beta = 0$ allows the crucial role of nonrivalry to come to the forefront. Why has the United States grown at 2 percent per year for the past 150 years? According to this theory, an important part of the explanation is that research effort has itself been growing exponentially. By any measure, we have been throwing ever larger quantities of resources into the hunt for new ideas. Just as more workers produce more cars, more researchers generate more ideas. Exponential growth in research effort leads to a growing stock of knowledge, and the fact that this knowledge is nonrival translates this into growing income per person.

Notice that it is precisely the nonrivalry of ideas that is crucial. If you try to replace knowledge with capital in this story, it no longer works. In the presence of population growth, it is easy to get aggregate capital to grow, just as we have gotten the aggregate stock of knowledge to grow. However, because capital is rival, growth in the aggregate capital stock at the rate of population growth does not translate into growth in capital per person or income per person. In contrast, because knowledge is nonrival, growth in the aggregate stock of knowledge at the rate of population growth will cause income per person to grow.

The result that long-run growth in income per person is proportional to the rate of population growth might well seem surprising, and perhaps even controversial. This is a point that is still being debated in the growth literature. It leads to many additional questions. Should Germany be richer than Luxembourg because it has more people? Should Ghana grow faster than China because it has a faster rate of population growth? Both these questions highlight the fact that ideas flow across borders, so an individual country’s economic performance is not intimately tied to its own population (see Coe and Helpman, 1995; Eaton and Kortum, 1996; Klenow and Rodriguez-Clare, 2005). What happens to growth if population growth comes to an end? These and other questions are raised by this theory of economic growth. Jones (2005) and Bloom et al. (2019) are two useful references for the reader who would like to learn more about this general topic.

The initial paragraphs of this section might appear superficially critical of Romer (1990a), but I hope by now that I have conveyed that the truth is actually the opposite. Removing the AK structure from the idea production function allows the importance of nonrivalry to be fully appreciated. Such AK structures were in vogue during the late 1980s, so it was natural for Romer to include one. But this inclusion turned out to be a distraction from the central contribution that Romer himself emphasized: the nonrivalry of ideas is what makes sustained exponential growth possible.
On this point, I cannot help but make a comment: the gods of growth theory are human after all! Solow’s brilliant and enduring Nobel-winning work fails to explain growth as an endogenous outcome. Romer’s contribution, which will be celebrated for centuries, has a blemish. We stand on the shoulders of giants, but it seems there are always a few more cherries to be found among the higher branches. This is somehow fitting, because if it were not the case, then growth itself might come to an end.

This leads to one final substantive remark. Even in light of Romer’s contributions, sustained exponential growth is surprising in another way: it hinges on an idea production function in which the difficulty of finding new ideas does not change dramatically over time. In the context of the simple model above, $\beta$ needs to be constant. There is no obvious reason why this should be the case. Diamond mines can have rich veins interspersed with poor ones, and, in the end, a mine contains only a finite supply of diamonds. Similarly, it could be easy to find ideas for a while and then harder, or vice versa. Or it could become ever easier over time until we suddenly run out of new ideas. The shape of the idea production function remains an intriguing subject of study.

VI. Conclusion

Beginning in the 1980s, the field of economic growth experienced a renaissance. Insights from the 1950s and 1960s were rediscovered, but the field, pushed forward by Romer, Lucas, Barro, Aghion and Howitt, Grossman and Helpman, and many others, achieved new heights and opened door after door for future research. Romer’s ultimate contribution stems from an appreciation that ideas, because they are nonrival, are in their very essence different from other goods. Romer was the first to truly understand the implications of nonrivalry – and the increasing returns that it implies – for economic growth.

We have a theory of endogenous technological change emphasizing the role of entrepreneurs and researchers, motivated in part by the profits associated with the discovery of new ideas. We have a world where the fundamental theorems of welfare economics no longer hold; there is an important role for government beyond simply providing secure property rights and stepping aside. And we have an understanding of sustained exponential growth, like that experienced in many countries for the past 150 years.

Romer’s contribution is essentially a beautiful theorem about the way the world works, of a kind that is extremely rare. It will be remembered and built upon for as long as economics exists.

References


