Annual Review of Economics

History, Microdata, and Endogenous Growth

Ufuk Akcigit1,2,3 and Tom Nicholas4

1Department of Economics, University of Chicago, Chicago, Illinois 60637, USA; email: uakcigit@uchicago.edu
2National Bureau of Economic Research, Cambridge, Massachusetts 02138, USA
3Center for Economic and Policy Research, Washington, DC 20009, USA
4Harvard Business School, Harvard University, Cambridge, Massachusetts 02138, USA; email: tnicholas@hbs.edu

Keywords
economic development, growth, history, innovation

Abstract
The study of economic growth is concerned with long-run changes, and therefore, historical data should be especially influential in informing the development of new theories. In this review, we draw on the recent literature to highlight areas in which study of history has played a particularly prominent role in improving our understanding of growth dynamics. Research at the intersection of historical data, theory, and empirics has the potential to reframe how we think about economic growth in much the same way that historical perspectives helped to shape the first generation of endogenous growth theories.
1. INTRODUCTION

The study of history has been instrumental to the development of endogenous growth theory. Variation in growth rates across countries or regions may result from factors such as human capital advantages, differences in market incentives, or technological progress, all of which depend on long-term developments. In an early contribution to the endogenous growth literature, Romer (1993, p. 562) alluded to the importance of history, stating that "our knowledge of economic history, of what production looked like 100 years ago, and of current events convinces us beyond any doubt that discovery, invention, and innovation are of overwhelming importance in economic growth."

Endogenous growth theory implies a more complex set of interactions than does the neoclassical model, where convergence of output to a long-run steady state leaves less scope for the analysis of historical differences (Nunn 2009). Because the theory of endogenous growth focuses on the mechanisms driving the real economy, it provides a rich framework for examining the causes of underlying differences in development. In the neoclassical model, growth in the long run can only be determined by exogenous technological progress or population growth. In endogenous growth theory, by contrast, differences in growth can result from a variety of factors influenced by the actions of economic agents, such as human capital externalities (Lucas 1988), economy-wide learning by doing (Stokey 1991), spillovers from innovation (Romer 1990), or Schumpeterian creative destruction (Aghion & Howitt 1992, Grossman & Helpman 1991).

In this review, we examine the progression of the 25-year-old endogenous growth theory literature. We begin with a summary of recent advances and explain how this research agenda raises important questions for historical analysis. We then focus our attention on some of the big debates in the recent literature where theory has intersected with historical data. We start with the growth slowdown hypothesis before showing a range of historical facts that are pertinent to endogenous growth, including the relationship between innovation and economic growth, the impact on growth of population density and market access, and human capital development. We also highlight a growing body of evidence on the social origins of inventors and the relationship between inequality and growth. The increasing availability of new data, including historical US Censuses, has provided a new fact base on the history of innovation and growth. This increasing availability of new evidence should encourage researchers to revisit the classic ideas from the endogenous growth literature and to develop new theoretical insights.

2. ADVANCES IN ENDOGENOUS GROWTH THEORY, PAST AND PRESENT

The seminal papers in endogenous growth theory during the 1980s and 1990s built on the basic idea that endogenizing long-run growth in income per capita would more accurately capture the way in which economies actually behaved. First, newly collected historical data sets were created to test the convergence hypothesis implied by neoclassical growth. Because empirical work tended to reject the idea that productivity growth rates of poorer countries were higher than those of richer countries, historical data guided advances in theory in the direction of endogenous growth. Second, advances in theory led to the incorporation of monopolistic competition into modeling frameworks, allowing for increasing returns to scale through investment in R&D, partial exclusiveness through market power, and spillovers from innovation.

Several papers in this new literature were motivated by historical data or insights. The notion that technologies emerge as a result of costly R&D efforts by individuals and firms underpinned newly developed models of innovation-based growth. Romer (1986) introduced spillovers
and increasing returns into the production function using motivation from the history of economic thought going back to Adam Smith’s pin factory and Alfred Marshall’s notion of external economies. Furthermore, the theory drew on an observation from Kendrick’s (1976) influential analysis showing that the rate of growth in inputs could not fully explain the rate of growth of output in the United States between 1929 and 1969. According to Romer, a model of endogenous growth incorporating increasing returns to scale could help to explain this discrepancy.

When Lucas (1988, p. 17) introduced spillovers through human capital externalities to explain variation in growth, he “reviewed an example of the neoclassical model of growth [and] compared it to certain facts of U.S. economic history.” The growth process in his model was at least partly motivated by Jacobs’ (1969) highly regarded book, The Economy of Cities, which provides an array of case studies of how US cities developed over time through a variety of channels, including their ability to create an environment where agents could optimally learn about production in an organized way. Human capital externalities, Lucas argued, would be assumed away under a more aggregate country-level neoclassical model with factor price equalization and growth convergence. His assertion about human capital externalities became central because it offered an explanation for perpetual growth without running into diminishing returns.

Schumpeterian models of endogenous growth by Aghion & Howitt (1992) and Grossman & Helpman (1991) were naturally historically grounded given Joseph Schumpeter’s (1942) path-breaking insights into the growth process. Although Schumpeter wrestled with intellectual disciplinary distinctions, his frame of thought reflected the intersection of history and theory (McCraw 2010). The concept of creative destruction, which Schumpeter considered to be a historical norm, gave rise to the incorporation of quality-improving innovations into a generation of endogenous growth models. While the papers of Aghion & Howitt and Grossman & Helpman were largely ahistorical in content, the idea that they built on is historically significant because their theories operationalized Schumpeter’s arguments (Aghion et al. 2014). Furthermore, the apparatus of Schumpeterian growth theory lends itself directly to the historical analysis of major technology waves. For example, Helpman & Trajtenberg (1998) built a model where it takes time for new intermediate goods to develop before general purpose technologies can have an impact on economic growth through the displacement of older technologies. The canonical cases are all historical: steam in the eighteenth century, electricity in the nineteenth century, or information and communications technology (ICT) in the twentieth century.

Because these early theories of endogenous growth were developed in light of historical evidence, attention increasingly focused on how far the models could be refined given potential disconnects between theory and the data. In a well-known and impactful article, Jones (1995) argued that endogenous growth theories in the work of Romer (1990), Aghion & Howitt (1992), and Grossman & Helpman (1991) were inconsistent with historical data. For example, if the Romer-style production function reflected reality, then a greater population size should have been associated with a higher steady-state growth rate because a larger population would both raise the return to innovation through a market size effect and increase the supply of researchers pursuing potentially innovative ideas. Yet Jones showed that US total factor productivity growth remained relatively constant between 1950 and 1988 despite a large increase in the number of scientists and engineers engaged in R&D activities. Jones’ specification of a semiendogenous growth model to remove the impact of population size motivated the contentious scale effects debate in the endogenous growth literature. Debate focused on whether it made more sense to accept decreasing returns to knowledge production or to use more disaggregated frameworks at, say, the product-line level as a way of abstracting from an economy-wide scale effect. Building on the work of Young (1998), Howitt (1999) proposed product proliferation as a remedy to the scale effect problem. According to Howitt’s mechanism, as the number of product varieties in an economy grows,
the effectiveness of research effort on each variety decreases because the population gets spread more thinly over a larger number of varieties. This makes the expected growth rate in each variety independent of the overall population level while preserving the role of policy in economic growth.

As the theory of endogenous growth progressed, one of the main challenges in maintaining historical perspective stemmed from data availability. Klette & Kortum (2004) cited a range of empirical facts to motivate a novel Schumpeterian growth model of firm growth driven by R&D investment and technological innovation. They defined a firm as a collection of production units such that firms could grow in size by introducing better-quality versions of other firms’ products—a natural starting point given the importance of product innovation and business stealing to the process of creative destruction. Importantly, Klette & Kortum emphasized that firm-level studies of innovation are crucial to the development of new theories. However, of the 10 facts that they used as motivation for their model, few were based on historical studies. Even the papers cited as pertaining to these facts relied on R&D data sets with limited coverage of the firm-level distribution. Lack of data proved to be a major constraint on the development of new theories. Griliches (1998) noted how problems associated with measuring R&D and innovation narrowed the scope of growth-related questions that could be addressed, titling a section of his survey article “Data Woes.” He argued that researchers who “collect their own data on interesting aspects of the economy” should be encouraged because these contributions create the knowledge base for new theory development (Griliches 1998, p. 365).

In recent years, advances in endogenous growth theory have resulted from significant improvements in data coverage and availability. Firm-level panel data sets, the National Bureau of Economic Research patent data set and COMPUSTAT link, and US Census Bureau administrative data have created a new platform for theoretical and empirical studies. For example, Lentz & Mortensen (2008) applied and extended the Klette & Kortum (2004) model to a data set of Danish firms. They incorporate variation in the ability of firms to create new products, which leads to the reallocation of workers from less productive to more productive enterprises. Akcigit & Kerr (2018) also extended the Klette & Kortum framework using a rich array of US firm- and patent-level data to quantify the impact on growth of different types of innovation—own-product-line internal innovations versus external innovations that capture market share from other firms. Finally, Acemoglu et al. (2018) focused on firms with different types of innovative potential and the growth implications of R&D policies designed to reallocate resources to the most-productive firms in the economy. Together, these theories extend endogenous growth in important ways by introducing reallocation dynamics and heterogeneity across different types of firms.

While some other branches of the endogenous growth literature have been less data driven, the theoretical insights have helped to reiterate the significance of variables incorporated into some of the early models of endogenous growth. Human interactions have continued to be important, following the original insight of Lucas (1988). For example, Lucas & Moll (2014) and Perla & Tonetti (2014) built alternative theories where growth can result from less knowledgeable individuals or firms interacting with and imitating more knowledgeable individuals or firms to improve their own productivity. This research has also triggered a debate over how theory should be used to advance the research agenda on endogenous growth (Romer 2015). Mathematical results involve some degree of abstraction from the way an economy functions, and therefore, underlying beliefs about the growth process can determine the nature of the assumptions made. One could correct for this problem by testing the new theories empirically or by integrating theory more closely with data through calibration exercises. The relevance of a model would then be judged by its ability to match real-world data, which would negate the impact of any potential belief biases.
As Aghion & Durlaf (2009, p. 7) noted, “One indicator of a good theory is the extent to which the predictions delivered by the theory can be validated or falsified by the available evidence.”

The new trade and innovation literature, which can be categorized as part of the endogenous growth literature, since trade is often considered to be an important engine of economic development, has been informed by both theory and data. Following seminal work on heterogeneous firms and international trade (Helpman 2006, Melitz 2003), research on trade has shifted from country or industry studies to a focus on firms and products. In that spirit, Atkeson & Burstein (2010) used theory to show that the welfare consequences of trade liberalization can be surprisingly neutral. Whereas exporting firms invest in process innovation to capture profits in a larger market, this effect is counterbalanced by a reduction in new entry (and thereby product innovation) due to the fact that payoffs to entrants are lower in a more competitive marketplace. Akcigit et al. (2018a) combined theory and empirics to further show how trade openness can create powerful firm-level dynamics, specifically around productivity cutoff points (i.e., a firm’s own productivity relative to its foreign competitors). At these cutoffs, firms can face strong incentives to innovate either to defend their share of domestic markets or to increase their share of export markets. This model reflects empirical facts in global innovation and competition, including convergence to the US technology frontier during the 1970s and 1980s. It also provides a framework for examining how the introduction of R&D subsidies in the United States in the late twentieth century helped to promote innovation and maintain US domestic competitiveness.

These types of mechanisms are important to explore because they reveal how government policies can influence the rate of technological progress and therefore long-run growth. As a corollary, the endogenous growth literature has embraced the idea that economic growth has societal consequences, a long-standing historical idea given the sharp increase in the top 1% income share in the United States in the early twentieth century and the recent repeat of that increase since the 1980s (Piketty & Saez 2003). There are ostensibly two ways of thinking about the relationship between innovation and inequality from a theoretical standpoint. If innovation is associated with payoffs (e.g., rents from patents), then we should see a positive association between innovation and inequality. Alternatively, if innovation involves creative destruction as new entrants displace incumbents, then innovation could lead to lower income inequality. Whereas Jones & Kim (2018) suggested that top income inequality is negatively correlated with innovation because the creative destruction effect dominates, Aghion et al. (2019) found that the relationship goes the other way. We discuss further evidence on this issue in Section 4.4.

Finally, Akcigit et al. (2018b) used new data on inventors and firms active in R&D and a database of tax changes over the course of the twentieth century to explore how one of the most significant government policy levers—taxation—can impact innovation. Higher tax rates (personal and corporate) are shown to negatively influence the quantity and quality of innovation, as well as where inventors and firms chose to locate their R&D activities. Importantly, Akcigit et al.’s analysis revealed that business stealing across states in response to different tax regimes cannot fully account for the size of the estimated responses to taxation, implying that the impact of taxes is quantitatively large. This study illustrated how insights from the theory of endogenous growth can be used to frame an empirical analysis that directly speaks to the issue of innovation policy design.

3. THE GROWTH SLOWDOWN HYPOTHESIS

History has informed thinking about economic growth in relation to the oft-cited growth slowdown hypothesis. The idea that growth in the United States may be about to stall is a few decades old. Jones (1997) cautions that positive trajectories in research intensity, educational attainment, and world economy openness as drivers of economic growth since the 1950s will level off. Bringing
microfoundations into this debate, Jones (2009, 2010) identifies three empirical facts about US innovation that also imply a slowdown. He shows that, first, inventors are becoming older in age over time; second, knowledge is becoming specialized; and third, teams are increasingly the source of the most creative ideas. If teams increase creativity through the combination of specialized insights, and if it takes time to acquire this knowledge (say, through education or learning), then we would expect to see a right-shifting age at invention distribution over time. If inventors develop their best ideas when they are young but the demands of acquiring knowledge to innovate increase over time, then the right-shifting age at invention distribution (relative to the baseline age distribution of the population) would lead to a slowdown in economic growth.

Recent US productivity performance has certainly been weak. One potential explanation is a reduction in business dynamism. Since the start of the twenty-first century, the share of high-growth young firms in the US economy has fallen sharply relative to the share in the 1980s and 1990s (Decker et al. 2016). To the extent that high-growth young firms have typically generated a disproportionate share of employment growth and innovation, this fact has important implications for aggregate growth. In the most recent endogenous growth models, young firms are the main drivers of innovation (Acemoglu et al. 2018, Akcigit & Kerr 2018).

However, Hsieh & Klenow (2018) use a model of endogenous innovation to show that the emphasis on declining business dynamism may be misplaced. They argue that within-firm improvements in product quality by incumbents accounts for a larger share of economic growth than do new firms entering a market and replacing those incumbents. Yet the measurement of this effect is controversial. In Hsieh & Klenow’s model, a young firm would be expected to claim market share in the space of 5 years, but Haltiwanger (2018) points out that a longer time horizon is necessary to fully capture the contribution of young firms to creative destruction. For example, Amazon’s future was highly uncertain within a 5-year time frame of its founding in 1994, but it spurred a revolution in the retail sector over subsequent decades. The papers discussed above highlight how close attention to data and measurement is necessary to ensure that the parameters calibrated by endogenous growth models capture the true impact of young firms on business dynamics.

In this spirit, Akcigit & Ates (2018) rely on an intensive set of empirical observations to build a unifying theory of endogenous growth that can speak to the potential causes of the trends laid out by Decker et al. (2016), as well as several other, possibly related, empirical regularities. Their theory builds on the step-by-step tradition of Schumpeterian creative destruction, and they explicitly account for entrants, laggard firms, and frontier incumbents. The framework allows them to account for the endogenous dynamics of firm entry, markups, and job reallocation, among other factors, while also reasonably capturing the postentry dynamics of young firms. Replicating the transitional dynamics of the US economy in the past several decades, Akcigit & Ates use their model to identify which policies or structural changes might have played a role in declining business dynamism and related trends. An intriguing finding suggests that a decline in the knowledge diffusion from frontier firms to laggard ones was an important driver behind many of the observed trends. Thus, the theory suggests that further research should look into mechanisms that may impede such diffusion, such as the gradual weakening of antitrust laws.

Bloom et al. (2017) formalize the growth slowdown hypothesis through the lens of idea-based growth theory. Their theory is straightforward. Assume that economic growth is a function of the number of researchers engaged in the search for new ideas and the productivity of these researchers. The challenge is thus to explain, as does Jones (1995), how research productivity can be declining when standard endogenous growth models predict exponential growth from a constant number of research workers due to the scale effect. After ruling out a variety of potential explanations—including the argument that the invention of new varieties of goods as the economy grows can conceal productivity effects when they are measured only in aggregate—Bloom et al.
argue that the evidence is consistent with ideas becoming harder to find. They show, for example, that while Moore’s Law predicts that the number of transistors per square inch on an integrated circuit will double each year, it is much harder to generate the exponential growth behind Moore’s Law today relative to the time period in which Intel’s Gordon Moore first made his prediction.

History has played a particularly pivotal role in shaping the growth slowdown debate because of the influence of Gordon’s (2016) book, *The Rise and Fall of American Growth*. Between 1870 and 1970, the United States experienced the most transformational period in its history, marked by revolutions in technological development—including electricity and the internal combustion engine—that profoundly reshaped living standards and the nature of production. Gordon argues that recent developments in ICT are not as far reaching in terms of impact on productivity relative to the earlier generations of technological discovery. This is because a growing share of the economy is accounted for by services where the scope for technology leading to productivity improvements is more limited than in manufacturing. Furthermore, in Gordon’s view, a combination of an aging population, fiscal challenges, rising inequality, and lagging educational attainment represents a strong impediment to growth.

Yet the slowdown hypothesis has been critiqued. At one level, there are a multitude of confounding measurement issues to consider. Crafts (2016) notes that chain linking using overlapping indices to deflate nominal output during the highly cyclical 1930s and 1940s overstates the growth in total factor productivity during two decades that are crucial to Gordon’s analysis. Furthermore, recent official statistics may underestimate the true productivity surge associated with ICT because real output is difficult to quantify. If a better-quality product is introduced to displace an incumbent product, then it is standard practice to assume that the new product has the same quality-adjusted price as products in the incumbent product category. However, due to creative destruction, the new entrant will produce at a lower quality-adjusted price. The impact of this distortion in official statistics (leading to overestimation of inflation and underestimation of growth) can be large, accounting for at least 0.6% in missing growth per year (Aghion et al. 2017b). It is easy to see how upward-biased productivity growth in the past, combined with downward-biased productivity growth in the present, could undermine the growth slowdown hypothesis.

As a strong opponent of the idea that growth will slow down, Mokyr (2018) makes the case that innovation—as a key variable in endogenous growth theory—will produce favorable growth dynamics in the future, just as it did in the past. Moreover, current technical breakthroughs in the digital age, he argues, have reinvigorated invention with even more profound possibilities than those of innovations from history. His case rests on a detailed synthesis of the history of innovation mapped to the present. Consider, for example, how much growth was spurred by the invention of the barometer (invented in 1643) and the vacuum pump (invented sometime during the 1650s). The barometer can be used to measure air pressure, while the vacuum pump exploits changes in pressure within a chamber. The combination of these insights created a pathway to the introduction of the steam engine, one of the most important innovations during the Industrial Revolution. Similarly, Mokyr argues, modern tools like DNA sequencing, high-powered computers, and nanotechnology are likely to lead to the diffusion of new science and radical technological breakthroughs.

Taking the lessons of history further, we see that the Second Industrial Revolution between 1870 and 1914 rested at least partly on more effective ways to make steel and a better understanding of chemistry—the so-called tool of innovation. Relative to these earlier epochs, Mokyr (2018) argues, the current technological tailwind is strong. In that spirit, Brynjolfsson & McAfee (2016) maintain that developments in robotics and artificial intelligence have the power to radically transform productivity. Alternatively, Aghion et al. (2017c) model artificial intelligence as a continuation of efforts to automate production since the Industrial Revolution, arguing that the
growth effects are actually ambiguous. On the one hand, artificial intelligence may substitute for human research effort, thereby generating the increasing returns that are so central to endogenous growth. On the other hand, growth could stall if artificial intelligence permits a proliferation of business stealing, which ultimately constrains investment in innovation. How an economy’s labor market will adjust to this new wave of innovation in artificial intelligence is also a key area of interest in economic growth frameworks (Acemoglu & Restrepo 2017).

Finally, it is worth emphasizing that the growth slowdown hypothesis has been presented in the past only to be refuted by the subsequent record of development. Nelson & Wright (1992) argue that the United States became a manufacturing powerhouse in the late nineteenth and early twentieth centuries as a result of innovations like mass production techniques, building on favorable resource endowments and a large market size. However, they also note that global technological capabilities and convergence to the US frontier, along with relative underinvestment in education and R&D in the United States, would cause these advantages to progressively erode. Despite the pessimism displayed in Nelson & Wright’s paper, soon after its publication, the United States embarked on a decade of fundamental technological leadership in new areas of innovation, especially ICT. While Nelson & Wright did not attach their argument to a specific theory of economic growth, it is clear that it was at least partially motivated by the neoclassical growth model as a framework to study convergence. It is interesting to see how, several decades after the start of the endogenous growth literature, there is still an intense debate concerning the real-world applicability of growth theory. There is therefore immense scope for the development of new theories to enhance our understanding of past and future growth trajectories in more predictable ways.

4. LESSONS FROM HISTORY

Akcigit et al. (2017b) argue that more attention to history can facilitate a deeper understanding of the growth process, which should, in turn, lead to better theories. To the extent that growth is about long-run changes, it is important to integrate historical data. Akcigit et al. build on insights from some of the standard theories in the endogenous growth literature and use historical data to identify fruitful avenues for further research. While they are not always able to identify causality in their empirics, they present interesting correlations that are consistent with various theories. These should act as stepping stones to a more fundamental comprehension of micro- and macrolevel growth dynamics.

In line with the supposition that history can be used to illuminate long-run changes, Akcigit et al.’s (2017b) analysis is based on a comprehensive data set of over 6 million US patents granted between 1836 and 2004 and a data set of inventors matched to recently released complete-count US Censuses conducted between 1880 and 1940. Both sources allow the authors to gain robust insights into the characteristics of US inventive activity over long time horizons. They establish a series of micro- and macrolevel observations concerning the factors that were driving US technological development. Several of their results confirm existing theories, especially the importance of innovation, spillovers, population density, market access, and human capital. Other results suggest new directions for theory in relation to the impact of family structure, social mobility, and inequality. Several of Akcigit et al.’s key results are presented in this section to illuminate related research contributions.

4.1. Innovation and Spillovers

Consider Figure 1 (see Akcigit et al. 2017b, figure 17), which shows a strongly positive correlation between patents and economic growth in US states between 1900 and 2000. The impact
of innovation on economic growth has been central to the endogenous growth literature going back to Romer’s work. Yet it is perhaps surprising how difficult it has been to establish a robust empirical link between innovation and growth. To our knowledge, no paper has actually shown that innovation is related to US economic growth over the long run.

Akcigit et al.’s (2017b) analysis of these data indicates that the effect of innovation could be large. They show that a high-innovation state like New Jersey (at the 90th percentile) would be 26% richer than a low-innovation state like Mississippi (at the 10th percentile) over the span of the century. They use an instrumental variables strategy (using federal contracts for Second World War technological developments distributed by the Office of Scientific Research and Development) to suggest that the kind of relationship shown in Figure 1 could be causal. Further pinning down how innovation has contributed to economic growth remains a topic of considerable interest.

Of course, in a neoclassical steady-state world, transition dynamics should render these growth rate differences unobservable. States that have higher innovation should exhibit a tendency to slow down, whereas lagging states should catch up to the frontier because the marginal product of capital in these places would be relatively high. However, endogenous growth theory opens up the possibility that some places may grow faster than others, with long-term consequences. The key is to explain why these differences in innovation-driven growth rates might have occurred.

One potential explanation is that the benefits of innovation are highly localized, which, in turn, would affect the production of knowledge and the population. In a seminal paper, Jaffe et al. (1993) identify localized spillovers from patented inventions, spawning an entire research agenda devoted to examining the geography of innovation. In a recent contribution, Kantor & Whalley (2019) analyze the impact on subsequent county-level productivity of federally funded agricultural experiment stations (AESs), which were established in the late nineteenth century to promote
biological innovation in crops. Because AESs were opened in the locations of land grant colleges, which, in turn, were determined in a largely random manner under the 1862 Morrill Act, the research design permits a causal interpretation of the results.

Kantor & Whalley (2019) find that farms in counties located close to AESs witnessed significantly higher productivity relative to those further away. Differences in productivity as a consequence of proximity to an AES persisted for at least two to three decades. Although they find that the proximity effect started to attenuate thereafter, it persisted much longer, even to today, in cases where the AESs focused on difficult-to-replicate basic (as opposed to applied) research. The idea that initial conditions can have such long-lasting effects on growth trajectories is supported by an entire body of research work (for a review, see Nunn 2009). For example, Bleakley & Lin (2012) find that historical portage locations (places between navigable waterways where cargo would be transited), as hubs of commerce and manufacturing in the United States, continue to attract higher populations long after their natural advantages become irrelevant.

4.2. Population Density and Market Access

Figure 2 (see Akcigit et al. 2017b, figure 5A) shows that population density was much higher in the most inventive US states during the early twentieth century. This finding is naturally associated with endogenous growth theories, wherein people are responsible for the production of the most creative new ideas. From around 1920, the majority of people in the United States lived in urban areas, as defined by the US Census Bureau, such that innovative activity should be traceable to the rise of cities. These are the places where people and firms tend to interact and agglomerate. The agglomeration literature illustrates that physical proximity can promote creativity and the

Figure 2
Urbanization and innovation. For each US state, average number of patents per 10,000 individuals between 1940 and 1960 is plotted against the percentage of individuals living in urban areas in 1940. Urbanicity is equal to \(-1.452 + 0.063 \times \text{patents per capita}\). The slope coefficient is statistically significant at the 1% level. Figure adapted with permission from Akcigit et al. (2017b, figure 5A).
exchange of ideas among inventors (Carlino & Kerr 2015). To the extent that growth can be related to the number of researchers engaged in the search for new ideas and the productivity of these researchers, population density should be an important determinant of the rate of technological progress.

Cities have traditionally played a preeminent role in the rise of modern industry (Glaeser 2011). Ideas can spread from person to person much more easily in close physical environments, which accelerates the production and refinement of new innovations. While business-stealing effects can detract from these innovation-inducing exchanges, the consensus is that urbanization is an overwhelmingly significant driver of economic development (Rossi-Hansberg & Wright 2007). Indeed cities, or urban agglomerations more generally, may be an important factor for the production of the increasing returns that are key to endogenous growth. Akcigit et al. (2017b) find that US inventors in the early twentieth century tended to move to urban areas, especially those that were financially developed, because these places were more conducive to innovation.

The integration of cities or regions through transportation improvements should further magnify the long-run rate of growth by making increasing returns to scale even more feasible. A good historical example is the rapid diffusion of railroads in the United States during the nineteenth century, which facilitated an expansion in trade. Donaldson & Hornbeck (2016) use new data on the widening of the railroad network to estimate the value of agricultural land associated with higher levels of market access. Their estimates imply that, without railroads in 1890, the total value of agricultural land would have been lower by a sizeable 60%. Although their analysis stops short of estimating the additional impact of railroad-related market access on the manufacturing sector, the implication of their paper is broad from an endogenous growth perspective. Market integration is one of the defining characteristics of early US development because larger markets should have increased the expected return to research and development. Indeed, Rivera-Batiz & Romer (1991) provide theory insights showing that the integration of markets is much more significant in driving economic growth than the neoclassical growth model supposes.

Further examining the connection between the diffusion of railroads and endogenous growth, Perlman (2016) investigates the extent to which railroads may have impacted the pace of technological innovation. She uses patents and data on transportation systems, including railroads, for the period from 1790 to 1900. Her results suggest a close correspondence between patents and transportation, with a doubling in patents per capita being observable in a county in the two decades following the introduction of railroad access. She attempts to establish causality using straight-line distances between major preexisting trade locations to instrument for the opening of a railroad, the logic being that straight-line distances represent a predetermined way of predicting the places that railroads would connect. Interestingly, her study finds that most of the gains from transportation derived not from increased market access per se, but rather from the localized changes to counties that a new railroad induced. Transportation improvements, she argues, facilitated the movement of people and ideas across the United States producing new clusters of innovation and economic activity as the network expanded.

4.3. Human Capital

As documented in Section 2, following Lucas (1988), the role of human capital as a spur to economic development occupies a central place in the endogenous growth literature, although there is debate on the extent to which educational attainment promotes growth or is simply a result of it (Bils & Klenow 2000). Educated workers can facilitate learning and technology diffusion. According to Goldin & Katz (2007), human capital is a central determinant of development, a theory that
they powerfully illustrate by linking the rise of investment in education to the spectacular rate of growth in the United States during the twentieth century.

The impact of human capital accumulation on growth has become increasingly prominent in historical studies. According to Cantoni & Yuchtman (2014), the first universities established during the Commercial Revolution in the Middle Ages created a pathway to development by facilitating the development of markets. Their evidence comes from Germany, where the Papal Schism of the late fourteenth century altered conceptions of learning and piety, leading to the birth of universities and the transformation of educational institutions. Markets developed, for example, because of an augmented supply of legal talent to enforce property rights and foster commercial exchanges. Squicciarini & Voigtländer (2015) find that human capital was a central determinant of French industrialization during the eighteenth century. They isolate the importance of upper-tail knowledge in the right-sided skew of the distribution. Cities with denser subscriptions to the famous *Encyclopédie*—a grand compendium of ideas on politics, economics, science, and technology—grew faster and also generated significantly more high-quality innovations.

Akcigit et al. (2017b) offer microlevel evidence on the backgrounds of inventors, who can be thought of as knowledge producers in the upper tail of the human capital distribution. In the United States during the early twentieth century, inventors accounted for only approximately 0.02% of the working-age population; for comparison, doctors and lawyers accounted for 0.46%. Figure 3 (see Akcigit et al. 2017b, figure 6) shows the probability of becoming an inventor conditional upon education. Akcigit et al. find that, if an individual had at least a college degree, then they were four times as likely to become an inventor compared to an individual with a high school diploma. In their sample of individual inventors matched to US Census records, 40% had a college degree in 1940, compared with only 10% of the population who were not inventors. Using similar data, Sarada et al. (2017) find that inventors were different from the general population for a range of demographic measures.

The centrality of human capital can be further evidenced by work on the relationship between universities and innovation in the United States. For example, Furman & MacGarvie (2007) show that universities had a strong casual impact on the growth of nearby industrial pharmaceutical laboratories during the 1920s, 1930s, and 1940s. Focal universities could have a large impact on local knowledge production. Andrews (2017) finds that counties chosen to host a new college in

---

**Figure 3**

Education and probability of becoming an inventor. *(a)* Inventors per 10,000 individuals by education. *(b)* Percentage of inventors in each education category. Figure adapted with permission from Akcigit et al. (2017b, figure 6).
the United States from the nineteenth to the mid-twentieth century generated more innovation relative to observably similar counties that only just missed out in the site selection process. Most of this effect, he argues, comes from migration because the most productive inventors moved to college towns, perhaps because they provided better amenities. Aghion et al. (2009) assume that highly educated people tend to want to colocate with one another, which, in turn, can spur innovation, capital, and labor efficiencies and, thereby, economic growth.

4.4. Family Background, Social Mobility, and Inequality

Given the importance of human capital to innovation, it is crucial to understand the backgrounds of those who enter into careers as inventors. This type of research has a long tradition in economic history. Seminal work by Lamoreaux & Sokoloff (1999) and Khan & Sokoloff (2004) provides profiles of inventors in the late nineteenth and early twentieth centuries, and these researchers have also explored the institutions, like markets for technology or patent rights, that have tended to encourage innovation. New data sources, which have become increasingly available, profiling inventors for both historical and current time periods permit consideration of a wider array of variables. Armed with new data, researchers have made progress on addressing a range of important growth questions, including how innovation is related to societal outcomes that policy makers care about—specifically social mobility and inequality.

Akcigit et al. (2017b) pay particular attention to the family backgrounds of inventors. Using data on incomes provided by the 1940 US Census, Figure 4a illustrates the relationship between parental affluence and the propensity to become an inventor in the United States. Whereas the relationship is quite flat for most of the father’s income range, there is a noticeable increase in the probability of becoming an inventor at the higher end of the income range. Parental income may matter for a variety of reasons. It could lead to higher levels of investment in a child’s human capital, relax liquidity constraints, or reflect heritable ability traits (Shea 2000). Each of these factors would matter from the standpoint of how talent is allocated into inventive activity.

Of particular importance is the fact that the shape distribution shown in Figure 4 can be observed in the modern era and for non-US data. Bell et al. (2017) use US patent data and tax records from 1996 to 2014 to examine the demographic characteristics of inventors. They find that social class, race, and gender all matter. Their result pertaining to parental income is shown in Figure 4b. It is strikingly similar to Figure 4a, which plots parallel data from more than half a century earlier. The same shape distribution can be observed for inventors in Finland, as shown in Figure 4c, based on a figure by Aghion et al. (2017a). Aghion et al. use highly granular modern administrative data and patents to investigate the potential channels underlying this relationship. They find that parental income, education, and socioeconomic status, as well as a child’s IQ, are all closely intertwined as predictors of entry into invention. Interestingly, they show that the impact of the mother’s income is more pronounced than that of the father’s income, a result that agrees with work by sociologists and psychologists showing how factors like mothers’ education can have a profoundly important impact on children’s outcomes (Harding et al. 2015).

These research findings provide a rich microfoundation for thinking about the relationship between innovation and human capital accumulation. For example, Bell et al. (2017) report that the probability of a White child becoming an inventor is over three times larger than that of a Black child, and gender differences are also stark, with women accounting for less than 20% of inventors for the latest (1980) birth cohort that they observe. Hsieh et al. (2013) show how costly these racial and gender gaps can be in a model of occupation choice where everyone in society is efficiently allocated to a job based on their talent. By integrating that mechanism into a growth model and calibrating using summary statistics in the data, they show how aggregate growth improves when
Figure 4
Parental income and becoming an inventor. (a) The number of US inventors per 10,000 people plotted by their father's percentile of wage income in 1940. Panel adapted with permission from Akcigit et al. (2017b). (b) The number of US inventors per 1,000 people plotted by their parental household income percentile in modern times. Panel adapted with permission from Bell et al. (2017). (c) The number of Finnish inventors plotted by their biological father's income percentile. Panel adapted with permission from Aghion et al. (2017a).

these gaps are reduced. They estimate that a narrowing in racial and gender disparities between 1960 and 2010 can account for approximately one-quarter of GDP per capita growth.

Given the fact that disparities do persist, it is natural to think about the societal implications. Work by Thomas Piketty, Emmanuel Saez, and their coauthors (e.g., Piketty & Saez 2013) on top income shares has reshaped our understanding of social mobility and inequality and highlighted these factors as important outcome variables to consider. Akcigit et al. (2017b) probe historical data to investigate the link between inequality and innovation. Figure 5 shows that the relationship between the top income share and inventiveness in the United States during the early twentieth century is U shaped, with more patenting being associated with more income held by the top 1% in highly innovative states like New York, New Jersey, and Massachusetts and with the reverse holding true in less innovative states such as West Virginia and North Carolina. Akcigit et al.'s research also shows that innovative places were associated with social mobility, perhaps because they exhibited characteristics that were more open to innovation (Acemoglu et al. 2014, Florida 2002). These correlations represent an important first step in documenting the potential welfare
Figure 5
Relationship between the top 1% labor income share and inventiveness in US states. Figure adapted with permission from Akcigit et al. (2017b, figure 19A).

implications of innovation. They also provide a fact base on which future theory and empirical work can build.

5. FUTURE RESEARCH AT THE INTERSECTION OF THEORY AND HISTORY

Having outlined how historical evidence has informed the endogenous growth literature, we now turn to the scope of future research. The upshot is that exciting opportunities exist for empirical and theoretical advances in endogenous growth, given a watershed in new data availability.

The recent release by the Integrated Public Use Microdata Series of 100% complete-count Census data for the United States up to 1940 represents one of the most important changes to the stock of information available for historical analysis in the past several decades. Millions of observations are now available covering family background, race, gender, occupation, migration, earnings, and education. These variables provide enormous scope for investigating factors such as human capital accumulation. Large representative data sets are appealing because they allow for more precise and generalizable estimates. These data sets also widen the scope of questions that can be addressed (Abramitzky 2015). The practice of testing growth theories and collating new facts allows theories to be updated, which is necessary for the study of endogenous growth to progress.

Pairing up complete-count Census data with other sources provides an opportunity for new research. One motivation for Akcigit et al. (2017b) was to exploit complete-count Census data and match inventors to patent records to collect together the key facts about innovation and economic growth in the United States for historical time periods. Advances in machine learning techniques
and matching methods mean that it is possible to link individuals in Census records over time to create panel data sets (Abramitsky et al. 2018, Bailey et al. 2018). From that perspective, big data in history is well suited to studying aspects of economic growth, which by construction involves changes over longer time horizons. How innovation affects social mobility and inequality (and vice versa) is a pertinent example of a question that can be more effectively analyzed with long-run data.

Data advances due to digitization are also expanding the frontier of our knowledge. For example, Sequeira et al. (2017) digitize historical maps to create a data set covering the railroad network in the United States from 1850 to 1920. Railroads transported people as well as goods. Sequeira et al. therefore pair these data with digitized data on immigration patterns and Census data to investigate the link between migration and long-run economic growth at the county level. They find strong effects over time. Places where migrants located early on are associated with favorable outcomes today, including better human capital and more innovation. Akcigit et al. (2017a) find that technology areas in patent classes where foreign-born inventors were particularly concentrated between 1880 and 1940 exhibited much faster patent growth between 1940 and 2000, and the quality of these patents was also higher.

In terms of theory development, the endogenous growth literature has evolved more and more to incorporate realistic heterogeneity by way of firms and innovations. In the absence of computational power, earlier models of endogenous growth during the 1990s were concerned about analytical tractability and therefore mostly abstracted from meaningful firm- and innovation-level heterogeneity. Recent advances in both microdata and ease of computation have allowed researchers to think harder about firm-level heterogeneity, which has facilitated a closer integration of theory and empirics. For instance, among several other papers, Akcigit & Kerr (2018) introduce heterogeneity in innovation qualities, which they discipline using patent data, and Acemoglu et al. (2018) focus on heterogeneity in the innovativeness of firms, which they estimate using US Census of Manufacturing data. These types of exercises are crucial because they allow data and theory to be more closely matched, leading to more credible policy recommendations.

6. CONCLUSION

Above, we highlight how the 25-year-old endogenous growth literature has produced many important theoretical frameworks for understanding long-run economic growth. We describe recent attempts to use modern and historical data to test some of these theory insights, which have helped to build a more complete understanding of the growth process. Substantial progress has been made, especially through efforts to combine new microdata with theory and empirics. This body of research has the potential to influence a wide array of economic growth policies.

With respect to future research directions, our knowledge of macroeconomic growth can be significantly improved by continuing to combine micro and macro perspectives. A macroeconomy is made up of micro firms that hire inventors to produce new innovative ideas. Focusing on the microeconomics of firms, inventors, and ideas is particularly crucial because it captures the rich heterogeneity associated with innovation. As we show in our review of the literature, it is possible to explore these dynamics for both current and historical time periods.

Historical studies complement studies using modern data by determining the extent to which influences on the growth process remain constant or change over time and space, which is key to guiding current debates on innovation and industrial policy design. Further efforts to explore historical and modern data sets should provide researchers with valuable resources for studying innovation and growth dynamics from a long-run perspective. Given major advances in microdata availability in recent years, we are optimistic that these efforts will improve our ability to further unravel the causes of endogenous growth and its societal consequences.
DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

U.A. gratefully acknowledges financial support from the National Science Foundation grant CAREER 1654719 and the Alfred P. Sloan Foundation grant G-2018-11071. T.N. received funding from the Division of Research and Faculty Development at Harvard Business School.

LITERATURE CITED

Atkeson A, Burstein AT. 2010. Innovation, firm dynamics, and international trade. J. Political Econ. 118:433–84
## Contents

The Economics of Kenneth J. Arrow: A Selective Review  
*Eric S. Maskin* ................................................................................................................................. 1

Econometrics of Auctions and Nonlinear Pricing  
*Isabelle Perrigne and Quang Vuong* ................................................................................................. 27

The Economics of Parenting  
*Matthias Doepke, Giuseppe Sorrenti, and Fabrizio Zilibotti* ................................................................. 55

Markets for Information: An Introduction  
*Dirk Bergemann and Alessandro Bonatti* ............................................................................................. 85

Global Wealth Inequality  
*Gabriel Zucman* ............................................................................................................................... 109

Robustness in Mechanism Design and Contracting  
*Gabriel Carroll* ..................................................................................................................................... 139

Experiments on Cognition, Communication, Coordination, and Cooperation in Relationships  
*Vincent P. Crawford* .......................................................................................................................... 167

Bootstrap Methods in Econometrics  
*Joel L. Horowitz* ............................................................................................................................... 193

Experiments and Entrepreneurship in Developing Countries  
*Simon Quinn and Christopher Woodruff* ............................................................................................. 225

Bayesian Persuasion and Information Design  
*Emir Kamenica* .................................................................................................................................... 249

Transitional Dynamics in Aggregate Models of Innovative Investment  
*Andrew Atkeson, Ariel T. Burstein, and Manolis Chatzikonstantinou* ................................................... 273

Echo Chambers and Their Effects on Economic and Political Outcomes  
*Gilat Levy and Ronny Razin* ................................................................................................................ 303

Evolutionary Models of Preference Formation  
*Ingela Alger and Jürgen W. Weibull* ...................................................................................................... 329

Approximately Optimal Mechanism Design  
*Tim Roughgarden and Inbal Talgam-Cohen* ....................................................................................... 355
Auction Market Design: Recent Innovations
  Paul Milgrom ................................................................. 383

Fair Division in the Internet Age
  Hervé Moulin ................................................................. 407

Legislative and Multilateral Bargaining
  Hülya Eraslan and Kirill S. Evdokimov .................................. 443

Social Networks in Policy Making
  Marco Battaglini and Eleonora Patacchini .................................. 473

Econometric Analysis of Panel Data Models with Multifactor Error Structures
  Hande Karabiyik, Franz C. Palm, and Jean-Pierre Urbain .................. 495

Using Randomized Controlled Trials to Estimate Long-Run Impacts in Development Economics
  Adrien Bouguen, Yue Huang, Michael Kremer, and Edward Miguel ............ 523

Is Education Consumption or Investment? Implications for School Competition
  W. Bentley MacLeod and Miguel Urquiola .................................. 563

Productivity Measurement: Racing to Keep Up
  Daniel E. Sichel .............................................................. 591

History, Microdata, and Endogenous Growth
  Ufuk Akcigit and Tom Nicholas ........................................... 615

Production Networks: A Primer
  Vasco M. Carvalho and Alireza Tahbaz-Salehi .................................. 635

Economic Theories of Justice
  Marc Fleurbaey ............................................................... 665

Machine Learning Methods That Economists Should Know About
  Susan Athey and Guido W. Imbens ......................................... 685

Weak Instruments in Instrumental Variables Regression:
  Theory and Practice
  Isaiah Andrews, James H. Stock, and Liyang Sun .......................... 727

Taking State-Capacity Research to the Field:
  Insights from Collaborations with Tax Authorities
  Dina Pomeranz and José Vila-Belda ......................................... 755

Free Movement, Open Borders, and the Global Gains from Labor Mobility
  Christian Dustmann and Ian P. Preston ..................................... 783
Monetary Policy, Macropudential Policy, and Financial Stability
   David Martinez-Miera and Rafael Repullo ............................................. 809

Has Dynamic Programming Improved Decision Making?
   John Rust ................................................................. 833

The International Monetary and Financial System
   Pierre-Olivier Gourinchas, Hélène Rey, and Maxime Sauzet ................... 859

Symposium: Universal Basic Income

Universal Basic Income: Some Theoretical Aspects
   Maitreesh Gbatak and François Maniquet .............................................. 895

Universal Basic Income in the United States and Advanced Countries
   Hilary Hoynes and Jesse Rothstein ..................................................... 929

Universal Basic Income in the Developing World
   Abhijit Banerjee, Paul Niehaus, and Tavneet Suri ................................ 959

Indexes

Cumulative Index of Contributing Authors, Volumes 7–11 ......................... 985

Errata

An online log of corrections to Annual Review of Economics articles may be found at http://www.annualreviews.org/errata/economics