

Daron Acemoglu: 2005 John Bates Clark Medalist

Robert Shimer

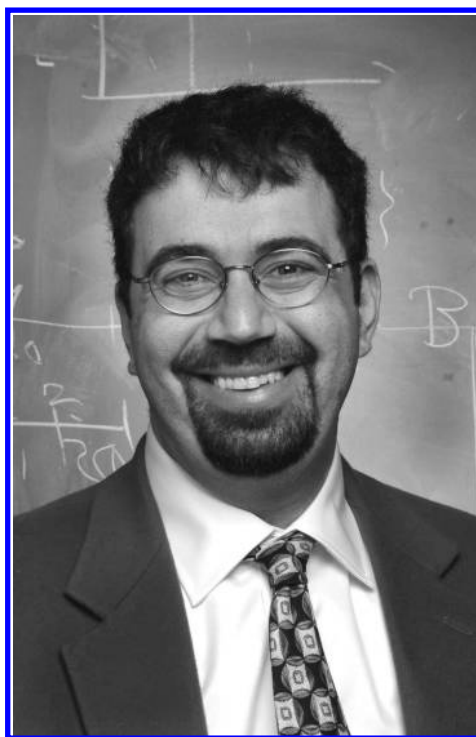
Daron Acemoglu, winner of the 2005 John Bates Clark Medal, uses theoretical and empirical analysis to tackle critical issues in a variety of fields in economics, including labor economics, macroeconomics, and political economy. His unparalleled combination of originality, thoroughness, and prolificacy has propelled him to the frontier of each field that he has explored.

Daron was born in Turkey and studied economics in the United Kingdom. He earned his Ph.D. from the London School of Economics (LSE) in 1992, where his thesis examiners recognized his accomplishments. In a letter supporting Daron for a thesis prize, one of his examiners, James Malcomson, wrote: “The thesis consisted of seven substantive chapters, each of which formed a paper in its own right. Each of these chapters was itself of very high quality. Indeed, I would consider even the weakest three of them to have been more than sufficient for the award of a PhD.” The letter by his other examiner, Christopher Pissarides, was similar: “I can say with some degree of certainty that it was the best thesis that I had ever examined at L.S.E., or elsewhere for that matter. It was original, full of important ideas and massive, without superfluous material.”

After a year lecturing at the LSE, Daron joined the MIT economics faculty in 1993, where he has been ever since. I was in the second year of my Ph.D. at MIT when Daron arrived. It would be amusing to write that I knew Daron would win the Clark Medal from the moment I met him, but I didn’t. Daron was so young that until he opened his mouth—that is, for about 60 seconds—I thought he was a newly-matriculated Ph.D. student.

In the winter of his first year at MIT, I attended his course “Economic Fluctuations, Labor Markets, and Complementarities.” It is not a coincidence that

■ *Robert Shimer is Professor of Economics, University of Chicago, Chicago, Illinois, and Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts.*



Daron Acemoglu

this title is a decent summary of much of my subsequent research. I worked closely under Daron's supervision during my dissertation research and had the opportunity to write four papers with him shortly after I completed my degree. Although our research interests have subsequently drifted apart, I have followed his papers with great interest. Some of his beliefs may have changed over time—my notes from that first course quote Daron as asserting that “the transversality condition is the most important thing in macroeconomics”—but his curiosity, enthusiasm, energy, breadth, and love for economics have not.

These may be part of the reason why Daron is such an extraordinary researcher. Daron can focus his intelligence tirelessly on the problem at hand. He does this better than almost anyone, not just because he has more energy or is smarter than others, but also because he loves economic research, both the product and, as importantly, the process. Daron's large set of publications, mostly in top refereed journals, provide some empirical evidence of his energy. Table 1 lists a subset of those papers, the ones I refer to throughout this essay using the associated number. Some of Daron's other outputs are harder to measure, including his generosity with his time to his current and former students and to his colleagues, both at MIT and in the broader profession.

Daron has another characteristic which is unfortunately too rare: although he started as an applied economic theorist, the Clark medal committee notes that “his

work is always motivated by real-world questions that arise when facts are difficult to reconcile with existing theory.” This combination of theory and data appears throughout almost all of his papers. Daron is motivated by a core set of questions and uses the best tools available to answer them. What determines the accumulation of human capital both during formal schooling and on the job? How do the implications of labor market frictions depend on the information available to job searchers? How do economic incentives affect the type of technological change that we observe? Why are there such enormous differences in output per worker and total factor productivity across countries? He tackles these questions repeatedly, sometimes using models, sometimes using novel data and empirical approaches, and often combining the two.

In the remainder of this essay, I explore how Daron has answered these and other questions. I touch on each of the papers mentioned in Daron’s Clark Medal citation ([2], [14], [15], [20], [22], [25], [27], [30], and [32]), and all the other papers listed in Table 1. I omit some of Daron’s other work, including almost all of his current working papers, in part because he can write faster than I can digest his research. I view this essay as an attempt at a road map through Daron’s fascinating list of publications and refer the reader to the original source for details.

I am not going to argue that Daron’s papers are always totally convincing; I am still too influenced by his critical assessment of the literature on “Economic Fluctuations, Labor Markets, and Complementarities” to be able to do that. But Daron is always thorough and ambitious. His research has created numerous novel directions for others to follow and has demonstrated the importance of mastering both theoretical and empirical approaches to economic problems.

Human Capital Accumulation

In a series of papers published in the late 1990s, Daron shows how labor market frictions can modify the conclusions of standard competitive theories of human capital accumulation. Each of these papers develops a simple model to illustrate its main economic mechanism. The later papers also use novel data to establish the model’s empirical relevance.

In [1], Daron develops a model that shows how labor market frictions can lead to externalities in human capital accumulation, a phenomenon that Robert E. Lucas (1988) argued is important for understanding the difference in income across countries. Daron assumes that the investments that both workers and firms make before they meet are complementary inputs into production. In a competitive labor market, each factor of production receives its marginal product, giving workers and firms the correct incentive to invest. With labor market frictions, however, workers must invest in human capital without knowing the type of job they will eventually get, and, symmetrically, firms must invest in physical capital or technology without knowing which workers they will hire. If some workers invest more—for example, because human capital accumulation becomes cheaper for

Table 1

Selected Papers by Daron Acemoglu

-
1. 1996. "A Microfoundation for Social Increasing Returns in Human Capital Accumulation." *Quarterly Journal of Economics*, 111(3): 779–804.
 2. 1997. "Training and Innovation in an Imperfect Labour Market." *Review of Economic Studies*, 64(3): 445–464.
 3. 1998. "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *Quarterly Journal of Economics*, 113(4): 1055–1089.
 4. 1999. "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence." *American Economic Review*, 89(5): 1259–78.
 5. 2002. "Directed Technical Change." *Review of Economic Studies*, 69(4): 781–809.
 6. 2002. "Technical Change, Inequality, and the Labor Market." *Journal of Economic Literature*, 40(1): 7–72.
 7. 2003. "Labor- and Capital-Augmenting Technical Change." *Journal of the European Economic Association*, 1(1): 1–37.
 8. 2003. "Patterns of Skill Premia." *Review of Economic Studies*, 70(2): 199–230.
 9. 2003. "Why Not a Political Coase Theorem? Social Conflict, Commitment, and Politics." *Journal of Comparative Economics*, 31(4): 620–652.
 10. 2005. "Politics and Economics in Weak and Strong States." *Journal of Monetary Economics*, 52(7): 1199–1226.
 11. With Joshua Angrist. 2000. "How Large Are Human Capital Externalities? Evidence from Compulsory Schooling Laws." *NBER Macroeconomics Annual*, 15(1): 9–59.
 12. With Veronica Guerrieri. 2005. "Capital Deepening and Non-Balanced Economic Growth." MIT Department of Economic Working Paper 06-24, June.
 13. With Simon Johnson. 2005. "Unbundling Institutions." *Journal of Political Economy*, 113(5): 949–95.
 14. With Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, 91(5): 1369–1401.
 15. With Simon Johnson, and James A. Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics*, 117(4): 1231–94.
 16. With Simon Johnson, and James A. Robinson, 2005. "Institutions as a Fundamental Cause of Long-Run Growth." In *Handbook of Economic Growth*, vol. 1, ed. Philippe Aghion and Steven Durlauf, 385–472. Elsevier.
 17. With Simon Johnson, and James A. Robinson, 2005. "The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth." *American Economic Review*, 95(3): 546–79.
 18. With Simon Johnson, James A. Robinson, and Yunyong Thaicharoen. 2003. "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth." *Journal of Monetary Economics*, 50(1): 49–123.
 19. With Joshua Linn. 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *Quarterly Journal of Economics*, 119(3): 1049–1090.
 20. With Jörn-Steffen Pischke. 1998. "Why Do Firms Train? Theory and Evidence." *Quarterly Journal of Economics*, 113(1): 79–119.
 21. With Jörn-Steffen Pischke. 1999. "Beyond Becker: Training in Imperfect Labour Markets." *Economic Journal*, 109(453): 112–142.
 22. With Jörn-Steffen Pischke. 1999. "The Structure of Wages and Investment in General Training." *Journal of Political Economy*, 107(3): 539–572.
 23. With James A. Robinson. 2000. "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective." *Quarterly Journal of Economics*, 115(4): 1167–1199.
 24. With James A. Robinson. 2001. "A Theory of Political Transitions." *American Economic Review*, 91(4): 938–963.
 25. With James A. Robinson. 2002. "Inefficient Redistribution." *American Political Science Review*, 95(3): 649–661.

Table 1—continued

-
-
26. With James A. Robinson. 2005. *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.
 27. With Robert Shimer. 1999. "Efficient Unemployment Insurance." *Journal of Political Economy*, 107(5): 893–928.
 28. With Robert Shimer. 1999. "Holdups and Efficiency with Search Frictions." *International Economic Review*, 40(4): 827–49.
 29. With Robert Shimer. 2000. "Productivity Gains from Unemployment Insurance." *European Economic Review*, 44(7): 1195–1224.
 30. With Robert Shimer. 2000. "Wage and Technology Dispersion." *Review of Economic Studies*, 67(4): 585–607.
 31. With Thierry Verdier. 2000. "The Choice between Market Failures and Corruption." *American Economic Review*, 90(1): 194–211.
 32. With Fabrizio Zilibotti. 1997. "Was Prometheus Unbound by Chance? Risk, Diversification, and Growth." *Journal of Political Economy*, 105(4): 709–51.
 33. With Fabrizio Zilibotti. 2001. "Productivity Differences." *Quarterly Journal of Economics*, 116(2): 563–606.
-

them—all firms will find it profitable to invest more, since firms do not know whether they will hire one of these workers. But this process in turn increases every worker's marginal product of human capital, encouraging investment by all workers.

These externalities have two main implications. First, any equilibrium is inefficient. The exact nature of the inefficiency depends on details of the model, particularly with regard to workers' bargaining power, but typically either the level of human capital accumulation is too low, or the level of physical capital accumulation is too low, or both. Second, when there are nonconvexities in firms' choice of technology, there may be multiple Pareto-ranked equilibria. In one possible equilibria, both workers and firms invest little, the marginal product of human and physical capital is low, and so the incentive to invest is low; while, in another possible equilibria, investments and marginal products are higher.

In [4], Daron uses a similar human capital externality to explain trends in U.S. wage inequality. There are two types of workers, skilled and unskilled. Firms can open jobs that are suited to one type of worker or middling jobs that can employ either type of worker. The advantage to opening a middling job is that it is easier to fill, since the firm can hire either type of worker. The disadvantage is that the technology is not well-suited to the type of worker the firm ultimately hires. Daron shows that, depending on workers' skill distribution, there can be a pooling equilibrium where all firms open middling jobs or a separating equilibrium where some firms open each type of job. An increase in the share of skilled workers, like the United States has experienced during the last several decades, can move the economy from the pooling to the separating result. Since this outcome increases the dispersion in technology utilized by each type of worker, it leads to more wage inequality.

This finding is consistent with the trends in wage inequality since the 1970s, a period when the number of college graduates and the college skill premium increased. Moreover, the mechanism is consistent with a variety of microeconomic

evidence. For example, following Sicherman (1991), Daron examines a question from the Panel Study of Income Dynamics (PSID), which asks, “How much formal education is required to get a job like yours?” He finds a decrease in both the amount and variance of overeducation between 1976 and 1985, consistent with a move from a pooling to a separating outcome. Daron also shows that the distribution of jobs shifted between 1983 and 1993 towards industries and occupations at the extremes of the wage distribution, again consistent with the underlying model. Although each piece of evidence might be “open to alternative interpretations,” Daron argues that “together this evidence . . . suggests that the changing composition of jobs may be an important component of the changes in the structure of wages.”

In a paper coauthored with Joshua Angrist [11], Daron offers another test for human capital externalities. The authors argue that cross-state variation in child labor and compulsory schooling laws led to exogenous differences in human capital accumulation across U.S. states between 1920 and 1960. If human capital externalities are important for earnings, the resulting differences in schooling should predict differences in wages. Unfortunately, the authors cannot reject the hypothesis that the external effect is zero and their point estimate is small: a one-year increase in everyone else’s schooling raises a worker’s wage by about 1 percent.

In another series of papers, Daron and Steve Pischke examine why firms help pay for workers to accumulate general human capital while employed. In a competitive economy, a firm does not profit when its employees accumulate general human capital since competition will simply drive up the wage. As a result, firms will not be willing to pay for such training (Becker, 1962). Daron and Steve note that this prediction is at odds with the observation that many firms do provide both generous wages and significant training to young workers, notably in German apprenticeship programs.

In [2], Daron suggests that search frictions can help to resolve this inconsistency. Because it is costly for a worker to move to another firm, an incumbent employer can reap at least some of the gain from training its existing employees. In [20], Daron and Steve argue that asymmetric information about a worker’s ability may also play a role (see also Greenwald, 1986). They assume that when a firm hires a worker, neither party knows the worker’s ability. Instead, the worker’s ability is revealed to the worker and employer, but not to other firms, during a training phase. Firms then keep the most able trained workers and lay off the rest. Adverse selection makes it costly for a trained worker to move to another firm; other employers will assume that an unemployed worker has low ability and so that worker will receive a low wage. This limitation on interfirm mobility holds down the wage of trained workers and implies that training is profitable even when it is transferable to other employers. In [22], Daron and Steve discuss a number of other mechanisms that introduce a wedge between the value of working for the current employer and the worker’s marginal product, including firm-specific human capital, efficiency wage concerns, and union wage setting.

To establish that asymmetric information may explain why firms pay for general training, Daron and Steve use German data in [20] to show the importance of adverse selection. They observe that in the case of workers forced to leave their apprenticeship to do their military service, adverse selection should not be important. Thus “military quitters” should earn more than other quitters upon reemployment. They may even earn more than stayers. (On the one hand, stayers are selected for their high ability. On the other hand, there is more competition for the service of military quitters, since no adverse selection problem exists for this group.) These predictions are borne out in data from the German Qualification and Career Survey. In the baseline specification, military quitters earn 4.5 percent more than other quitters upon reemployment and 3.3 percent more than stayers, consistent with the existence of at least some adverse selection.

None of these papers is the last word on these subjects. I could quibble with whether certain of these mechanisms are likely to be economically significant. Someone else might quibble with details in the papers’ empirical implementation. Daron would undoubtedly have a thoughtful and lengthy response to each of us. But at a more basic level, these papers showed how labor market imperfections, like search frictions and asymmetric information, can affect something that might have seemed only tangentially related: human capital accumulation both prior to and on the job.

Directed Search

Shortly after I completed my Ph.D., I had the pleasure of writing a series of papers with Daron on directed search. In a more standard random search model, unemployed workers and firms with job vacancies are viewed as passive agents, periodically learning about each other’s existence. A directed search model recognizes that workers and firms face trade-offs between the difficulty of finding a trading partner and the resulting terms of trade. A worker searching for a lower wage may find a job faster. Symmetrically, a firm offering a high wage may fill its vacancies faster than if it offers a low wage. The distinction is important not just because directed search sometimes seems like a more realistic assumption. We found that the assumption can substantially alter the predictions of equilibrium search models.

A few memories stand out from my experience writing with Daron. He was bursting with ideas and could write drafts faster than I could edit them. He had a keen sense of how to model an economic mechanism without sacrificing analytic tractability. At the same time, his goal was to expound mechanisms that he believed, either because of intuition or more frequently because of his knowledge of the relevant evidence, were empirically relevant even if they were not necessarily the simplest. Finally, he was a patient teacher who built my confidence and helped me make the transition from student to colleague.

One of our papers [27] explores the idea that unemployment insurance affects

the type of jobs that workers seek and that this might improve productive efficiency in a case where markets are incomplete. To show this, we write down an equilibrium model of competitive search (Moen, 1997), where insurance markets are incomplete. In the simplest version of the model, homogeneous firms simultaneously decide whether each firm should create a single job vacancy, and those that do announce wage offers. Risk-averse workers observe all the offers and apply for one of the jobs. Because of matching frictions, markets do not clear. However, a firm that offers a higher wage is more likely to attract a worker and fill its vacancy, and likewise a worker who applies for a higher-wage job faces more competition and so is more likely to stay unemployed. We find that more risk-averse workers apply for lower-wage jobs to reduce their unemployment risk, which encourages too much entry by rent-seeking firms. Unemployment insurance can undo this inefficiency by encouraging workers to seek higher-wage jobs.

We then extend the model to introduce a physical capital investment decision that firms must make before they learn whether they have hired successfully. When workers are risk averse and do not have access to unemployment insurance, excess entry discourages firms from making large investments, because the physical capital is less likely to be utilized. Unemployment insurance can undo this inefficiency as well. Thus well-designed unemployment insurance both mitigates excess job creation and encourages firms to create better jobs. In [29], we assess the quantitative importance of this mechanism in a calibrated model. We argue that it can potentially justify the high level of unemployment insurance and associated high level of unemployment in some continental European countries.

In [30], we examine what happens when it is costly for workers to acquire information about available jobs. The basic setup is similar to the one in [27]: firms decide whether to create a job vacancy, what wage to offer, and (in some versions) how much to invest. The difference is that it is no longer free for workers to observe these wages. Instead, as in Burdett and Judd (1983), workers must pay a cost for each wage that they observe. We find there is always an equilibrium in which the labor market breaks down: all firms offer a wage that leaves workers indifferent about unemployment and so no worker is willing to pay the cost of learning about wages. If the cost of sampling wages is sufficiently small, however, an equilibrium with wage and technology dispersion emerges. Some firms offer high wages and make large capital investments, while others offer low wages, make small capital investments, but earn the same profits. Workers who learn about both high- and low-wage jobs apply for the high-wage jobs, and so those firms are more likely to fill their vacancies. This outcome both rewards firms for their high wage offer and justifies the physical capital investment they made earlier.

In equilibrium, we observe that high-wage jobs are more productive. This outcome is inconsistent with a competitive labor market, where the law of one price ensures that homogeneous workers are paid the same wage even if jobs differ in their productivity. Through the lens of many models of labor market frictions, such as [1], this looks like evidence of rent-sharing and suggests investment should be inefficient. However, in a directed search model, firms that choose to invest more

have a higher opportunity cost of leaving their job vacant and so are willing to pay more to fill their job faster. Indeed, as the cost of sampling job offers converges to zero, we find that the outcome converges to a competitive search equilibrium (Moen, 1997). In our final paper, [28], we directly examine that limiting economy and again establish that the possibility of paying a higher wage to attract workers faster can ensure efficient investments before hiring occurs even in the presence of labor market frictions and despite this apparent violation of the law of one price.

It is difficult to say whether directed or random search is a more realistic assumption; reality is likely a combination of the two. Moreover, our research on directed search neglects other labor market frictions, including labor mobility costs; occupation-, industry-, or firm-specific human capital; and other turnover costs, which are probably also important for understanding human and physical capital investments and unemployment. Nevertheless, we believe these papers take an important step towards understanding how the nature of frictions affects labor market outcomes.

Directed Technical Change

Another series of papers seeks to understand how the composition of labor supply affects research and development. A theme of this research is that technical change is directed towards the activities that give innovating firms the highest return. In contrast, most of the literature on technical change either treats technological progress as exogenous or formulates the economic environment in a way that renders the direction of technical change irrelevant.

One important applied question that this research answers is why the skill premium—for example, the return to a college education—has increased during a period when the supply of skilled workers has grown so rapidly. A more theoretical concern is whether such an observation is consistent with the observed pattern of long-run balanced growth in the aggregate economy—for example, a constant growth rate and a constant real interest rate—and with shifts over time in employment and output from agriculture to industry to services.

In [3], Daron develops the basic framework (see also [5]). Two types of labor, skilled and unskilled, are imperfect substitutes in the production of final goods. Firms in the research and development sector can produce new technologies that improve the productivity of either type of worker; the reward is a monopoly on their product. In equilibrium, the productivity growth rate for each type of worker increases with the wage bill of that type of worker. If the elasticity of substitution between skilled and unskilled labor is one—that is, a Cobb–Douglas aggregate production function—a change in the share of skilled workers does not affect the share of the wage bill going to each type of worker and so does not affect the incentive to create the two types of innovations. But if the two types of labor are more easily substituted, an increase in the number of skilled workers raises the share of output accruing to skilled workers and encourages the development of

skill-complementary innovations, that is, skill-biased technical change. If skilled and unskilled workers are sufficiently good substitutes, with an elasticity of substitution greater than two, this effect is strong enough that an increase in the share of skilled workers actually raises the skill premium.

The alternative view is that all technical change is inherently skill-biased (Autor, Katz, and Krueger, 1998). Firms do not have the option to produce technologies that are complementary to unskilled workers. According to this view, the recent acceleration in the speed of technical change raised the skill premium and encouraged more workers to go to college. In [3] and [5], Daron points out a couple of shortcomings of this theory. First, much of the increase in college enrollment occurred during the 1970s, during a period when the skill premium fell; moreover, much of that increase was probably related to the Vietnam War and had little to do with the evolution of the skill premium. Second, there is nothing inherently skill-biased about technical change. Many technologies in the nineteenth century were designed to replace skilled craftsmen, while more recently, computers have reduced the skill requirements in some jobs while raising it in others. Indeed, Daron argues in [5] that a period of unskill-biased technical change in the late eighteenth and early nineteenth century corresponds to a large migration of unskilled farm workers to cities in England. A large increase in the relative supply of unskilled labor induced the development of technologies that were complementary to unskilled labor.

In [7], Daron examines the decision to develop technologies that are capital or labor augmenting. The theoretical logic is similar, although the economic issues are distinct. If there is a Cobb–Douglas aggregate production function, so the elasticity of substitution between capital and labor is one, the capital share and labor share of output should be constant over time. While this pattern holds in the long-run in a variety of countries, there are short-run fluctuations in the labor share, suggesting that the aggregate production function is not Cobb–Douglas. On the other hand, if the production function is not Cobb–Douglas and technological progress is at least partially capital augmenting, then the return to investment in physical capital and the interest rate should decline over time, which it does not do. The puzzle is therefore why firms choose to develop technologies that are complementary to labor rather than to capital.

Daron also proposes a resolution in [7]. Suppose that capital and labor are relatively poor substitutes in the production of final goods. An increase in the labor-augmenting technology raises the share of income accruing to capital, which raises the profitability of research in capital-augmenting technologies. On the other hand, capital accumulation raises the share of income accruing to labor, which makes research in labor-augmenting technology more profitable. Along a balanced growth path, these two forces cancel out. Daron proves that in the long run, all technical progress is labor augmenting; however, in response to a change in labor market policy, there may be transitory episodes of capital-augmenting technical change and transitory fluctuations in the capital share.

In work with Veronica Guerrieri [12], Daron extends this insight to an econ-

omy with multiple sectors in order to reconcile the “Kuznets facts”—constant growth, constant capital–output ratio, and a constant real interest rate—with the “Kaldor facts”—a systematic shift of production from agriculture into industry and then into services. The starting point is an assumption that some sectors are more capital intensive than others and that households are relatively unwilling to substitute across sectors. In this case, Daron and Veronica show that the growth rate of the economy is determined by the less capital-intensive sector. This result holds even when technical progress is endogenous. That is, directed technical change is consistent with balanced growth at the aggregate level but shifts across sectors at the more micro level.

A concern one might have with all these papers is the definition of the relevant market for new technologies. For example, although there has been a dramatic increase in the skill of the U.S. labor market during the last 30 years, this increase in the skill level available to firms has arguably been dwarfed by globalization, which involves the opening of markets in less developed countries which are abundant in unskilled labor. In [3], Daron argues that the relevant market for innovations is in fact a small subset of developed economies because of poor enforcement of property rights in less-developed countries. In [8], he shows that increased globalization can cause skill-biased technical change. Globalization raises the demand for skilled workers. Since most of these workers are located in countries with well-enforced property rights, this induces skill-biased technical change, which further raises the wage of skilled workers in developed countries and in less-developed countries.

In work with Fabrizio Zilibotti [33], Daron takes these arguments one step further. If new technologies are developed to take advantage of the skill composition and capital intensity of developed countries, they will be inappropriate for less-developed countries. They argue that this insight can help to explain much of the difference in measured total factor productivity and output per worker across countries (Hall and Jones, 1999). To show the empirical relevance of this finding, they use their model to prove that the biggest differences in measured total factor productivity across countries should be—perhaps surprisingly—in sectors that are the least skill intensive. Data from the U.N. General Industrial Statistics supports this hypothesis.

Finally, work with Joshua Linn [19] provides microeconomic evidence that technical change is directed in at least one industry—pharmaceuticals. The paper takes advantage of the fact that demographic shocks, notably the U.S. baby boom, lead to a systematic pattern of disease (see also Cerda, 2003). As the baby boom generation has aged, the demand for drugs that treat childhood ailments has declined while the demand for drugs that treat middle-aged ailments has increased. The authors estimate that a 1 percent increase in potential market size—the number of people in the appropriate demographic cohort—induces a 6 percent increase in the number of new drugs designed to treat that market. Although this finding does not tell us whether technical change can be directed towards skilled

or unskilled labor or towards capital or labor, it suggests that the incentive to create appropriate innovations can be powerful.

It is too early to say whether Daron's work in this area will provide the definitive explanation either for the relative supply and wages of skilled and unskilled workers, or for the Kuznets and Kaldor facts. But these papers make a convincing case that we should think about technical change as being directed and that this matters for the predictions of standard economic theories.

Investment and Growth

Daron and Fabrizio Zilibotti advance an intriguing hypothesis in [32]: low-income countries are limited in their ability to grow because the most productive activities have substantial undiversifiable risk. While Argentina and Canada may have been similar in 1900, Canadians fortuitously chose to invest in activities that eventually received high rates of return while Argentines were less fortunate. Canada grew into a large, diversified, and stable economy, while Argentina spent the next century sputtering through a series of crises.

Daron and Fabrizio demonstrate this idea in a formal model of growth. The basic idea can be understood in the following stripped-down version of their model. A risk-averse individual can invest either in a safe technology with a low but certain return, r , or in one of two risky technologies. Each of the risky technologies pays $R > 2r$ with probability $\frac{1}{2}$ and zero otherwise. Moreover, the return on the risky technologies is always perfectly negatively correlated, so if an individual could invest in both risky technologies, he would obtain a certain return of $R/2$, greater than the safe return of r . The other key assumption is that there is a minimum efficient scale for operating the risky technology within an economy and the risk cannot be diversified internationally.

Under these assumptions, it is not feasible to operate both technologies in a poor country. Individuals will choose to invest some of their wealth in safe assets and possibly some in one of the risky assets. If their gamble turns out well, they will accumulate enough wealth to invest in both assets, exiting this initial risky growth phase. But there is also a chance that the gamble and hence the level of wealth will fall. In equilibrium, poor countries are both riskier and have lower expected growth, consistent with a variety of facts that the authors summarize in the paper.

From the perspective of this essay, this paper is almost as interesting for what is not in it as for what is. The paper does not discuss the possibility that good, stable institutions may be important for long-run growth and that poor institutions may explain both the low level and the high volatility of income in many less developed countries; compare this with papers like [18]. The remainder of this essay examines this topic, which has consumed most of Daron's research energy since 2000.

Institutions and Development

To paraphrase Robert Lucas (1988), once one starts to think about economic growth, it is hard to think of anything else. Daron's work in recent years adds a data point in support of this claim. Daron and his coauthors Simon Johnson and James Robinson have examined some of the biggest questions in economics. Why are some countries so much richer than others? What determined which countries grew rich and when? I find it useful to break Daron's papers on this topic into two broad categories: The first set is largely empirical and builds a convincing case for "institutions" as the answer to these questions. The second set (which I examine in the next section) is more theoretical and tries to understand what makes for good institutions.

The basic problem with identifying whether good institutions lead to economic growth is that the choice of institutions is endogenous. For example, democratic accountability may be important for ensuring that governments try to maximize the well-being of their citizens. However, it is also likely that democracy is a luxury good; citizens demand more of it as the country grows richer. More generally, third factors like a good climate may be responsible both for the choice of institutions and for the level of income. A simple regression of gross domestic product (or other measures of economic well-being) on democracy would be plagued by these endogeneity problems.

To address this, one needs to identify cases where the choice of institutions was essentially random. The division of Korea following World War II is one such incident, and the consequences are stark; see Daron's survey article with Simon and James [16] for more on this case. There are two shortcomings with this example. First, it is just that—an example, a single data point. Second, there are now so many institutional differences between North and South Korea that it is difficult to know which institutions really matter.

In [14], Daron, Simon, and James propose an ingenious method of identifying the causal effect of institutions on growth for a broader set of countries, all former colonies. The basic idea is that the mortality rates faced by Europeans settlers affected the institutions that were imposed on colonies. When European settler mortality was high, the colonizers preferred institutions that favored the easy extraction of resources. When settler mortality was low, colonies were more likely to end up with strong protection for property rights and effective checks on government power. While the institutions persisted into the twentieth century, the sources of variation in settler mortality, mainly malaria and yellow fever, are no longer important determinants of national income.

Daron, Simon, and James use two-stage least squares to quantify the importance of this channel. In the first stage, they regress one measure of the current quality of institutions—protection against the risk of government expropriation between 1985 and 1995—on the mortality rate of settlers in the nineteenth century. They then regress a measure of current log GDP per capita on the quality of institutions predicted in the first stage. The resulting estimates show the effect

of this type of variation in institutions on income, assuming they have correctly identified the unique channel by which early settler mortality affects current GDP. The authors use a couple of strategies, such as tests of over-identifying restrictions, to argue that they have. They conclude that institutional differences are an important source of variation in income. For example, importing Chilean institutions into Nigeria would raise GDP by a factor of seven. African countries are poor not because of culture or geography but because they have worse institutions, which in turn is a legacy of malaria and yellow fever from more than 100 years ago.

In [13], Daron and Simon examine which type of institutions matter most for growth. The authors consider two types: “property rights” against government expropriation as in [14] and “contracting institutions,” which affect the ability of private agents to contract with each other. Both types of institutions are correlated and endogenous, but Daron and Simon attempt to disentangle them using instrumental variables. They argue that, while early settler mortality predicts property rights institutions, contracting institutions were determined by the legal system of the colonial power. They exploit this variation in both types of institutions to show that property rights are much more important than the contracting environment for long-run economic outcomes.

In [15], Daron, Simon, and James propose another way to observe the consequences of colonial institutions on long-run economic outcomes. The authors note a strong negative correlation between per-capita GDP in 1995 and urbanization and population density in 1500, measures of income prior to colonization. They provide a simple and plausible explanation for this pattern (p. 1235): “European colonialism led to the development of institutions of private property in previously poor areas, while introducing extractive institutions or maintaining existing extractive institutions in previously prosperous places.” Good institutions then led to sustained growth in previously poor areas while extractive institutions caused stagnation, the eponymous “reversal of fortune.”

In [17], the three authors turn from the colonies to the colonizers. They show that from 1500 to 1850, almost all of the growth in western Europe was accounted for by the subset of countries with the easiest access to Atlantic trade: Britain, France, the Netherlands, Portugal, and Spain. They argue that this was not because Atlantic trade directly led to economic growth; the magnitude of trade was too small. Instead, Atlantic trade strengthened the merchant class, particularly in countries with relatively weak monarchies like Britain, Netherlands, and to some extent France. This emergent middle class in turn pushed for political institutions that protected their property rights. These institutions later proved critical for economic growth.

Together the series of papers by Daron Acemoglu, Simon Johnson, and James Robinson make a convincing case that protecting citizens from government expropriation risk is critical for economic growth. These are not the first papers to make this case and, given the importance of this question, future researchers may cast doubt on some aspects of the interpretation of the data. But these papers have had

an enormous impact both on how economists think about the role of institutions in growth and how they approach seemingly intractable identification problems like those faced by Daron, Simon, and James.

Political Economy Theory

In a series of papers, often co-authored with James Robinson, Daron uses formal economic models to examine important issues in political economy. There is a clear connection between these papers and those I discussed in the previous section, since a common theme is why countries sometimes choose institutions that are bad for economic growth.

In [23], Daron and James explore the transition from autocracy to democracy. What makes the political elites allow the poor to vote? Daron and James argue that the elites enfranchise the poor to ward off revolution. Once the poor can vote, they enact redistributive reforms, which make revolution less attractive. The shortcoming of this answer is that the elites could simply enact the redistributive reforms directly, without enfranchising the poor. Daron and James argue that this alternative is infeasible because of a time-consistency problem. Once the threat of revolution has passed, an autarchic elite will undo the redistribution. Thus, redistribution without enfranchisement is not a credible alternative and does little to alleviate the threat of revolution. In contrast, they argue that it is more difficult to undo the enfranchisement of the poor—in [23], enfranchisement is permanent by assumption. In this way, extending the franchise solves the time-consistency problem and eliminates the threat of revolution. As evidence in support of this mechanism, the authors show that enfranchisement in Britain and other European countries was accompanied by a substantial increase in redistributive taxation and a consequent reduction in inequality.

In Europe, the assumption that enfranchisement is irreversible seems plausible, but in much of the rest of the world it is not. For example, most Latin American countries have experienced alternating periods of democracy followed by military coups and dictatorship. In [24], the authors extend their previous paper to allow for the possibility of coups, a revolution by the elite that disenfranchises the poor. Much as the threat of revolution forces the rich to redistribute and extend the franchise to the poor, the threat of coups limits redistribution in a democracy. The authors show that the occurrence of coups is dependent on the extent of inequality. In a relatively equal society, the equilibrium will be similar to that described in the authors' previous paper: there is little role for redistributive taxation and so the threat of a coup is not credible. In contrast, in an unequal society, it may be impossible to prevent coups. When the poor have power, they will opt for a lot of redistribution, which then leads to the next coup. Thus, consistent with evidence on Europe and Latin America, coups are more common and there is more volatility in the extent of redistribution in countries with more inequality.

In [25], Daron and James examine the nature of redistribution in a democ-

racy; in particular they address the vexing question of why redistribution often takes an inefficient form such as agricultural price supports rather than transfers to farmers. The authors suggest a theory built on two key assumptions: politicians cannot commit to future policies and political power increases with group size, at least over some range of parameters. They develop a two-period model. In the first period, the population is divided between farming, which is relatively inefficient, and manufacturing. Some new agents come into the economy and must decide which sector to join. Although it would be more efficient for them to enter manufacturing, the existing farmers may prefer to structure subsidies that encourage the entrants to become farmers. Having more farmers around in the second period improves their ability to extract resources from the government, partially alleviating the time-consistency problem. In [9], Daron similarly emphasizes lack of commitment as an important reason why societies choose inefficient institutions.

While these first four papers emphasize one role of the government, namely redistribution, in [31] Daron and Thierry Verdier recognize another potential role: correcting for market failures. The authors analyze the behavior of government bureaucrats who monitor polluting firms. The bureaucrats are self-interested and hence susceptible to bribes. This setting leads to a trade-off between dealing with market failures, which requires fines to align the private and social costs of pollution, and reducing corruption, since fines beget bribes. The government can use efficiency wages and monitoring to mitigate corruption, but typically will permit some corruption as the price of dealing with the market failure. Perhaps surprisingly, if it is difficult to monitor bureaucrats, it may be optimal to enlarge the bureaucracy. This increases the probability that polluting firms will be monitored, which reduces the necessary fine for pollution and hence lowers the equilibrium bribe.

The idea that government intervention may be important in a market economy carries over to [10]. In this paper, Daron notes that many countries that fail to develop are actually “weak states” with no effective central governments, rather than “strong states” in which the executive’s power is unchecked and expropriation risk is particularly high. The paper develops a formal model that illustrates the disadvantages both of strong states, where a self-interested government imposes high taxes that stifle economic activity, and of weak states, where a government that fears for its future existence is unwilling to invest in infrastructure or in good legal and contractual institutions. The paper also shows that most successful countries have relatively large governments and argues that this reflects a “consensually strong state”—one in which the government has an incentive to levy taxes and invest in public goods, exactly because of a credible threat that it would otherwise be replaced.

Conclusion

Daron Acemoglu's research has covered a tremendous amount of ground in a short period of time, and so, by necessity, has this essay. Fortunately Daron and his coauthors have also written at least four surveys of his own and related work, [21] on human capital accumulation, [6] on technological change, [16] on institutions and development, and [26] on the transition from dictatorship to democracy. These surveys are excellent entry points into these aspects of Daron's work.

As I wrote in the introduction, these papers may not be the last word on these subjects. If nothing else, Daron himself will always have more to say. I also suspect that Daron's research to date does not predict what his main focus will be in the next ten years. For example, many of his most recent working papers are on optimal routing in communications networks. But I am sure that we can expect to see Daron approach economic problems as he has in the past, using clever applied economic theory motivated by novel empirical observations.

■ *This research is supported by a grant from the National Science Foundation. I am grateful for comments from Daron Acemoglu, Simon Johnson, Kiminori Matsuyama, James Robinson, Mariagrazia Squicciarini, Fabrizio Zilibotti, and members of the editorial board of the Journal of Economic Perspectives; to James Malcomson and Christopher Pissarides for sharing the quotes in the second paragraph; and to Barbara Boavista for her able research assistance.*

References

- Autor, David, Lawrence Katz, and Alan Krueger.** 1998. "Computing Inequality: Have Computers Changed the Wage Structure?" *Quarterly Journal of Economics*, 113(4): 1169–1213.
- Becker, Gary S.** 1962. "Investment in Human Capital: A Theoretical Analysis." *The Journal of Political Economy*, 70(5, Part 2): 9–49.
- Burdett, Kenneth, and Kenneth Judd.** 1983. "Equilibrium Price Dispersion." *Econometrica*, 51(4): 955–70.
- Cerda, Rodrigo.** 2003. "Drugs, Market Size, and Population." Ph.D. thesis, University of Chicago.
- Greenwald, Bruce C.** 1986. "Adverse Selection in the Labour Market." *Review of Economic Studies*, 53(3): 325–47.
- Hall, Robert E., and Charles I. Jones.** 1999. "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics*, 114(1): 83–116.
- Lucas, Robert E. Jr.** 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics*, 22(1): 3–42.
- Moen, Espen.** 1997. "Competitive Search Equilibrium." *Journal of Political Economy*, 105(2): 385–411.
- Sicherman, Nachum.** 1991. "Overeducation in the Labor Market." *Journal of Labor Economics*, 9(2): 101–122.

This article has been cited by:

1. Emiliano Brancaccio, Fabiana De Cristofaro. 2024. In Praise of ‘general laws’ of Capitalism: Notes from a Debate with Daron Acemoglu. *Review of Political Economy* **36**:1, 289–303. [[Crossref](#)]