Innovation and human capital policy

John Van Reenen
**Abstract**

If innovation is to be subsidized, a natural place to start is to increase the quantity and quality of human capital. Innovation, after all, begins with people. Simply stimulating the “demand side” through R&D subsidies and tax breaks may only drive up the price, rather than the volume of research activity. By contrast, increasing the supply of potential inventors can both directly increase innovation and reduce its cost. This paper examines the evidence on human capital policies for stimulating innovation such as expanding the home-grown workforce, fostering immigration, boosting universities and reducing barriers to entry into inventor careers, especially for under-represented groups. The evidence suggests targeting high ability but disadvantaged potential inventors at an early age is likely to have the largest long-run effects on growth.

Key words: innovation, R&D, intellectual property, tax, competition

JEL codes: O31; O32

This paper was produced as part of the Centre’s Growth Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

This builds on work with many co-authors, in particular Nick Bloom and Heidi Williams. I am grateful for comments by Ben Jones, Austan Goolsbee and an anonymous referee. This research was supported in part by the Sloan Foundation; Schmitt Sciences, the Smith Richardson Foundation and the Economic and Social Research Council. The content is solely the responsibility of the author and does not necessarily represent the official views of the NBER. For acknowledgments, sources of research support, and disclosure of the author’s material financial relationships, if any, please see [https://www.nber.org/books-and-chapters/innovation-and-public-policy/innovation-and-human-capital-policy](https://www.nber.org/books-and-chapters/innovation-and-public-policy/innovation-and-human-capital-policy).

John Van Reenen, MIT, NBER and Centre for Economic Performance, London School of Economics.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission in writing of the publisher nor be issued to the public or circulated in any form other than that in which it is published.

Requests for permission to reproduce any article or part of the Working Paper should be sent to the editor at the above address.

© J. Van Reenen, submitted 2021.
1 Introduction

Since the 1970s, US productivity growth has slowed—reflected in falling total GDP growth from 4 percent in the postwar years, to under 3 percent from the mid-1970s, and to under 2 percent since 2000. Moreover, slow productivity growth has been accompanied by disappointing real wage growth for most US workers, as well as rising wage inequality. At the time of writing, the COVID pandemic has damaged productivity more than any other shock in living memory.

Innovation is the only way for the most developed countries to secure sustainable long-run productivity growth. For nations further from the technological frontier, catch-up growth is a viable option, but there are limits to such a strategy for leading-edge economies such as the United States. Certainly, even in the United States, many firms are well behind the technological frontier, and helping these firms catch up—for example, by improving management practices—would be valuable (e.g., Bloom and Van Reenen, 2007). Nonetheless, innovation policy design is a key part of any solution for revitalizing America, and it will lead to large long-run increases in welfare.

The attraction of human capital policies for innovation is that they act directly on the supply side, to increase the number of potential and actual innovators. Romer (2001) emphasized the advantage of supply side policies. Demand side policies such as tax credits and direct government R&D grants can be effective in increasing firms’ incentives to do more R&D—and there is an impressive body of microeconomic research on this (Akcigit and Stantcheva 2020; Bloom, Williams, and Van Reenen 2019). However, if the supply of R&D workers is very inelastic then the risk is that the increase in demand merely drives up the equilibrium cost of R&D without increasing its volume. In other words, the incidence of the subsidy is on innovation prices rather than innovation quantities. This is what Goolsbee (1998) found in aggregate US data—scientists’ wages rose substantially with increased federal R&D spending. Microeconomic analysis might miss this, as the wage increase is a general equilibrium effect, absorbed away by the time dummies typically included in standard evaluations. Furthermore, since R&D workers are above median-pay employees, this type of demand side policy could increase inequality as well as providing little in the way of aggregate innovation.

In reality, the elasticity of supply of R&D workers is unlikely to be completely fixed, especially when we consider immigration into the United States (see below). However, in the
short run, supply could be relatively inelastic, so these concerns are real.

A supply side increase in the quantity and quality of R&D workers carries fewer of these risks. Unless the new workers are dramatically less productive than the existing stock, we would expect a direct increase in innovation. Furthermore, the increase in supply should reduce the equilibrium cost of R&D—meaning that a successful supply side policy provides a further indirect boost to the amount of innovation as firms face lower R&D costs. The work in this paper focuses on such human capital supply side policies.

The structure of the paper is as follows. I provide some background R&D and workforce statistics in section 2; in section 3, I discuss the rationale for innovation subsidies; in section 4, I discuss the evidence for four types of human capital supply policies. Section 5 offers some concluding comments.

2 Background: R&D and the Scientific Workforce

In 2015, spending on R&D performed in the United States was just under half a trillion dollars. Figure 1 shows R&D spending as a fraction of GDP for major industrialized countries. The United States spends more on R&D than any other, accounting for roughly 28 percent of global R&D spending—see National Science Board (2018).

---

2 Unless otherwise noted, all data and facts in this section—and later in the paper—are drawn from National Science Board (2018).
Over time, however, the picture is less rosy. The United States has maintained an R&D-to-GDP ratio of 2.5–2.7 percent since 1981. By contrast, other countries, particularly in Asia (Japan, South Korea, and most recently and spectacularly, China), have been devoting increasing amounts of national income to R&D. Furthermore, although US R&D intensity has been stable since the mid-1960s, the composition of R&D spending has changed dramatically, as government funding has declined and private-sector funding has increased to fill the void (see figure 2). In 2015, businesses spent more than twice as much as the federal government on R&D. R&D spending as a share of GDP grew from around 1.3 percent in 1953 to around 2.7 percent in 2015. Government tends to fund higher-risk basic research that private investors are often reluctant to take on. Therefore, public R&D investment tends to produce higher-value, high-spillover (see below) inventions over a longer period of time. Despite the decline in government R&D funding, the private sector has also invested less in basic research over time (e.g., Arora, Belenzon, and Patacconi 2018). The decline in basic research in both public- and private-sector R&D spending may be a reason why the productivity of American R&D appears to have fallen over time, as documented by Bloom et al. (2020).
In recent years, around 13 percent of US R&D is performed at colleges and universities (accounting for just under half of all basic research). Additionally, just over half of R&D expenditures at US colleges and universities is federally funded. The vast bulk of that funding goes to the life sciences, with smaller amounts going to engineering, the physical sciences, and other fields. Reflecting that distribution of federal funds across fields, the top agencies supporting federally funded academic R&D are the Department of Health and Human Services, the Department of Defense, and the National Science Foundation.

These statistics focus on R&D spending, but of course, another set of metrics of innovative activity focus on the scientific workforce. Table 1 shows that the fraction of all US workers who are researchers has grown consistently since 1981, just like the R&D-to-GDP ratio. There were about 5.3 researchers per thousand workers in 1981, 7.3 in 2001, and 9.23 in 2017. However, the growth rate was faster in other advanced economies. France, Germany, and Japan all had lower numbers in 1981, but have overtaken the United States in the most recent years. The most dramatic change over that period has been in South Korea, where the ratio of
researchers per thousand employees rose from 6.3 in 2001 to 15.3 today. China’s fraction of researchers looks less impressive than its R&D spending in figure 1, but it has still more than doubled the researcher proportion since 2001 from 1 to 2.4.

Table 1: Number of researchers per 1,000 employees, Selected Countries

<table>
<thead>
<tr>
<th></th>
<th>United States</th>
<th>China</th>
<th>France</th>
<th>Germany</th>
<th>Korea</th>
<th>Japan</th>
<th>United Kingdom</th>
</tr>
</thead>
<tbody>
<tr>
<td>1981</td>
<td>5.28</td>
<td>3.78</td>
<td>4.65</td>
<td>5.23</td>
<td>5.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>7.29</td>
<td>1.02</td>
<td>6.83</td>
<td>6.32</td>
<td>9.87</td>
<td>6.57</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>9.23</td>
<td>2.41</td>
<td>10.9</td>
<td>15.33</td>
<td>9.88</td>
<td>9.43</td>
<td></td>
</tr>
</tbody>
</table>


Another metric relevant to the size of the scientific workforce in the US is the number of temporary work visas issued in categories that cover high-skill workers: J-1 (exchange visitors), H-1B, and L-1 (intracompany transferee) visas. Between 1991 and 2015, the primary increase in these categories was in J-1 visas (exchange visitors), which increased from around 150,000 to over 330,000. The number of H-1B visas increased from around 52,000 in 1991 to nearly 175,000 in 2015; note that the H-1B cap of 65,000 was in place over that entire period, implying that the growth was driven by H-1Bs to universities, nonprofit research facilities, and government research facilities, all of which are exempt from the H-1B annual quotas.

3 The Case for Government Promotion of Innovation

Governments often want to increase innovation in an attempt to encourage economic growth; indeed, countries that have higher levels of R&D spending are typically richer (e.g., Jones 2015). However, standard economic theory also suggests that, in the absence of market failures, it would be better for the government to leave investment decisions in the hands of
private firms. There are many oft-cited government failures (e.g., the Anglo-French supersonic jet, Concorde; see Lerner 2005 for one discussion). On the other hand, there are also many examples of impressive inventions built on government-sponsored R&D, such as jet engines, radar, nuclear power, GPS, and the internet (Janeway 2012; Mazzucato 2013).

The central market failure that economists have focused on in justifying government intervention is knowledge spillovers. If one firm creates something truly innovative, this knowledge may spill over to other firms through copying or by learning from the original research—without having to pay the full R&D costs. Ideas are promiscuous; even with a well-designed intellectual property system, the benefits of new ideas are difficult to monetize in full. There is a long academic literature documenting the existence of these positive spillovers from innovations.

That said, economic theory also suggests that R&D expenditures in a market economy can be either too low or too high, depending on the net size of knowledge spillovers and business-stealing effects. The key idea behind business stealing is that product market rivalry can lead to overinvestment in R&D because innovators take market share from other firms without necessarily generating much social benefit. An oft-cited example is the case of “me-too” pharmaceuticals, where one firm may spend billions of dollars to develop a drug that is only incrementally better than a drug produced by a rival firm. However, the small improvement in therapeutic value may allow the second firm to capture nearly the entire market. In cases where me too drugs are therapeutically indistinguishable from the products they replace (and setting aside the possibility that me too drugs may generate the benefit of price-cutting competition), this dynamic potentially generates a massive private benefit for shareholders of pharmaceutical firms, with little gain for patients.

Broadly stated, three methods have been used to estimate spillovers: case studies, a production function approach, and research based on patent citations.

Perhaps the most famous example of a case study approach is Griliches (1958), who estimated the social rate of return realized by public and private investments in hybrid corn. He estimates an annual return of 700 percent, as of 1955, on the average dollar invested in hybrid corn research. Seed or corn producers appropriated almost none of these returns; they were instead passed to consumers in the form of lower prices and higher output. Although this study is widely cited, Griliches himself discusses the challenges inherent in calculating the rate of return
on something akin to a successful “oil well.” Whereas we observe an estimate that captures the cost of drilling and developing a successful well, we would ideally prefer to generate an estimate that includes the cost of all of the “dry holes” drilled before oil was struck.

The second (“production function”) approach abandons the details of specific technologies and instead relates productivity growth (or other measures of innovative output) to lagged measures of R&D investment. The key challenge is that R&D is determined by many factors that also independently affect productivity. Recent papers applying this approach have used policy experiments such as tax changes that influence research and development investments to identify the arrow of causality (for example, Bloom, Schankerman, and Van Reenen 2013).

The key idea in using patent citations to measure spillovers is that each patent cites other patents, in addition to associated publications, all of which form the basis of “prior art”—existing innovations that enabled that particular patent. Trajtenberg (1990) and Jaffe, Trajtenberg, and Henderson (1993) pioneered this approach. Although some evidence shows that citations can be strategic (and that some citations are added by patent examiners during the course of the patent examination process), the existence of patent citations provides a measurable indication of knowledge spillovers (e.g., Griffith, Lee, and Van Reenen 2011).

One challenge arising with the production function approach is how to find ways of identifying the relevant channels of influence so that “one can detect the path of spillovers in the sands of the data” (Griliches 1992, 31). Herein lies an advantage of using patent citations, which provide a direct way of inferring which firms receive spillover benefits. More generally, the trick in the search for spillovers has been to focus on defining a dimension (or dimensions) over which knowledge spillovers are mediated. Firms less distant from each other in this dimension will be more affected by the R&D efforts of their peers—for example, technological distance as revealed from past patenting classes (Jaffe 1986), geographical distance between corporate R&D labs, or product market distance (the industries in which firms operate).

As a whole, this literature on spillovers has consistently estimated that social returns to R&D are much higher than private returns, which provides a justification for government-supported innovation policy (see Jones and Summers 2020). In the United States, for example, recent estimates in Lucking, Bloom, and Van Reenen (2020) used three decades of firm-level data and a production function–based approach to document evidence of substantial positive net
knowledge spillovers. The authors estimate that social returns are about 60 percent, compared to private returns of around 15 percent, suggesting the need for a substantial increase in public research subsidies.

Given this evidence on knowledge spillovers, one obvious solution is to provide strong intellectual property rights such as patents to inventors as a means of increasing the private return to inventing. A patent is a temporary right to exclude others from selling the protected invention. Patents entail some efficiency loss because they usually enable sellers to charge a higher price markup over production costs. However, this downside could be outweighed by the gains in dynamic efficiency that arise from patents providing stronger incentives to do more R&D because potential innovators expect to be able to appropriate more of the benefits. However, in practice, as we will discuss in more detail below, the patent system is highly imperfect. For one thing, other firms can frequently invent around a patent—after all, the empirical evidence on knowledge spillovers summarized above is drawn from the United States, which has a strong system of intellectual property rights by international standards.

There are other potential justifications for R&D subsidies in addition to knowledge spillovers, related to failures in other markets. For example, financial constraints may limit the amount of innovation that firms can carry out. Because innovation is intangible, it may be hard for firms to raise funding when they have no collateral to pledge to banks in return for debt funding (Arrow 1962). This insight suggests that equity might be a better source of funding for innovation, but equity faces a different challenge: an asymmetry of information. Before innovations are patented or demonstrated in the market, the requisite secrecy about technology makes fundraising difficult. A pitch of “trust me, I have a great idea so please fund me” is rarely effective, whereas a pitch of “let me describe my not-yet-patented idea in detail” opens up the possibility of potential investors stealing an idea from the entrepreneur.

Evidence suggests that financial constraints do often hold back innovation (for a survey, see Hall and Lerner 2010). However, the presence of financial constraints around research and development funding is not necessarily a reason for government subsidies: governments often have worse information about project quality than either firms or investors, so designing appropriate policy interventions is difficult. Hence, effective government policies to address financial constraints involve not just financial support for firms but also a mechanism to accurately identify and select higher-quality investments, which is typically difficult.
4 Human Capital Innovation Policies

We now turn to consider explicit human capital policies.

4.1 Undergraduates and Postgraduates

The most commonly discussed policy here is to increase the number of individuals with training in science, technology, engineering, and mathematics, commonly known as STEM. The direct way would be to subsidize PhDs and postdocs in these subjects, increasing the generosity of support for training in these fields. Indirectly, training and subsequent careers in these fields could be made more attractive through more grants and support, especially in labs.

More generally, one can imagine support for raising educational attainment at an even younger age (undergraduates and even K through 12). There is a huge literature documenting the complementarity between human capital and new technologies (“skill-biased technical change”), so increasing human capital could have a positive effect on technical change (e.g., Autor, Goldin, and Katz 2020; Van Reenen 2011). However, this literature is usually focused on the diffusion of technologies (e.g., adoption of information and communications technology) rather than on pushing forward the technological frontier. For innovation to the economy (rather than to a firm), it is likely that postgraduate qualifications are much more important.

Much macroeconomic analysis has been conducted of the impact of human capital on growth (see, e.g., Sianesi and Van Reenen 2003 for a survey). However, the literature is rather inconclusive because of the difficulty of finding credible instruments at the macro (or industry) level. The large number of other confounders at the macroeconomic level makes it hard to infer causality. There is a vast literature looking at the impact of schooling on wages, but there is rather a paucity of work looking at more specific interventions on the STEM workforce.

4.2 University Expansion

Many papers examine the role of universities in economic prosperity in general and in innovation in particular. A major idea in these papers is that university founding and expansion increases the supply of workers with STEM qualifications and that these STEM workers increase innovation. There certainly appears to be a correlation between areas with strong science-based universities and private-sector innovation (for example, Silicon Valley in California, Route 128 in Massachusetts, Research Triangle in North Carolina, and others).
Valero and Van Reenen (2019), looking at 50 years of subnational data across more than 100 countries, find that the founding of a university increases local GDP per-capita growth in subsequent years (which also spills over nationally). The Jaffe (1989) paper was a pioneer in this area by documenting that state-level spending on university research in certain industries seems to generate higher local corporate patenting. Acs, Audretsch, and Feldman (1992) use innovation counts instead of patent data and find even stronger effects for spillovers from university research. Related findings of the positive effects of university location on patenting has been found in more recent datasets by Belenzon and Schankerman (2013), Hausman (2018), and Andrews (2020). Furman and MacGarvie (2007) studied how universities with stronger academic research profiles increased the growth of local industrial pharmaceutical labs from 1927 to 1946. They used land grant college funds under the Morrill Acts to generate some exogenous variation in the location of universities to argue that the correlation is causal. In the biotech industry, Zucker, Brewer, and Darby (1998) show that firms tend to locate near universities to take advantage of star scientists.

However, universities may also have other effects on innovation over and above the supply of graduates. Firstly, research and innovation by university faculty, possibly collaborating with private-sector firms, could also directly increase innovation. The vast literature on clustering has this as one of the mechanisms. Secondly, universities may influence local democratic participation and institutions, which may also have an effect on innovation. If universities have an effect on innovation (or growth) over and above the impact on human capital, then they are not valid instruments for human capital, as this violates the exclusion restriction. For example, Valero and Van Reenen (2019) found that university expansion was associated with more graduates, more innovation, and stronger institutions. Of course, the reduced form effect of universities on innovation is still an interesting effect if it is causal, but the mechanism through which universities raise innovation may not be solely (or even at all) through the human capital channel.

4.2.1 Graduate Supply
To make progress in isolating why universities may have an impact on innovation as key suppliers of STEM workers, Toivanen and Väänänen (2016) document that individuals growing up around a technical university (such institutions rapidly expanded in the 1960s and 1970s in Finland and offered postgraduate engineering) were more likely to become engineers. They
showed that this led to significantly greater patent activity. Establishing three technical universities caused on average a 20 percent increase in US Patent and Trademark Office patents by Finnish inventors. In a similar vein, Carneiro, Liu, and Salvanes (2018) compare municipalities in Norway where there was an upsurge in government college start-ups in the 1970s to synthetic cohorts of areas where the expansion did not take place. They document evidence for more R&D and a speed up in the rate and direction of technological progress about a decade after the colleges’ founding (if they were STEM focused).

Bianchi and Giorcelli (2019) present the most direct test of the role of universities in increasing STEM supply. They exploit a change in the enrollment requirements for Italian STEM majors, which has a significant effect of expanding the number of graduates. They document that this exogenous increase in STEM majors led to more innovation in general, with effects concentrated in fields related to chemistry, medicine, and information technology. They also document a general “leakage” problem that may accompany efforts to simply increase the STEM pipeline: many STEM-trained graduates may choose to work in sectors that are not especially focused on R&D or innovation, such as finance.

4.2.2 Research Grants to Academics (and Beyond)

One variety of government program seeking to encourage innovation is the direct provision of grant funding (e.g., through the National Institutes of Health, or NIH), either to academic researchers or more widely. Spending public R&D subsidies on universities makes economic sense as knowledge spillovers from basic academic research are likely to be much larger than those from near-market applied research in corporations.

Evaluating the effectiveness of grant funding for research and development is challenging. Public research grants usually (and understandably) attempt to target the most promising researchers, the most promising projects, or the most socially important problems. As a result, it is difficult to construct a valid counterfactual for what would otherwise have happened to the researchers, firms, or projects that receive public R&D funds. If $1 of public R&D simply crowds out $1 of private R&D that would otherwise have been invested in the same project, then public R&D could have no real effect on overall R&D allocations (much less on productivity, growth, or other outcomes). However, it is also possible that public R&D grants add to private R&D spending, or even that public R&D “crowds in” and attracts additional private R&D spending.
Jacob and Lefgren (2011) use administrative data on US grant applications to the NIH and effectively compare academic applicants that just received and just missed receiving large NIH grants. They document that these grants produce positive but small effects on research output, leading to about one additional publication over five years (an increase of 7 percent). One explanation for this modest effect is that marginal unsuccessful NIH grant applicants often obtain other sources of funding to continue their research. Consistent with that story, productivity effects are larger among researchers who are likely to be more reliant on NIH funding (for whom alternative funding sources may be less likely to be available).

Looking beyond academic output, public research and development grants may affect private firms in several ways. First, public R&D grants to academics can generate spillovers to private firms (as discussed above). Azoulay et al. (2019a) exploit quasi-experimental variation in funding from the National Institutes of Health across research areas to show that a $10 million increase in NIH funding to academics leads to 2.7 additional patents filed by private firms. Second, private firms themselves sometimes conduct publicly funded R&D. Moretti, Steinwender, and Van Reenen (2019) use changes in military R&D spending, which is frequently driven by exogenous political changes, to look at the effect of public subsidies for military R&D. They document that a 10 percent increase in publicly funded R&D to private firms results in about a 4 percent increase in private R&D, suggesting that public R&D crowds in private R&D (and also, they document, raises productivity growth). Third, private firms can directly receive public subsidies. Howell (2017) examines outcomes for Small Business Innovation Research (SBIR) grant applicants, comparing marginal winners and losers. She estimates that early-stage Department of Energy SBIR grants roughly double the probability that a firm will receive subsequent venture capital funding, and that receipt of an SBIR grant has positive impacts on firm revenue and patenting. Howell et al. (2021) find that SBIR grants in the US Air Force also have positive effects on these measures of innovation.

4.2.3 National Labs

Governments can also fund their own R&D labs—for example, SLAC National Accelerator Laboratory at Stanford University. These labs can generate more research activity and employment in the technological and geographical area in which the lab specializes. Jaffe and Lerner (2001) analyze national labs, which are often managed by universities, and document evidence of spillovers. The United Kingdom’s Synchrotron Diamond Light Source appeared to
create spillovers (Helmers and Overman 2017), but it seems in that case to have occurred mainly through relocation of research activity within the country rather than by spurring an increase in nationwide research efforts.

4.2.4 Academic Incentives

Controversy has arisen over how to design complementary policies that enable the resulting discoveries—when made at universities—to be translated into technologies that benefit consumers. The 1980 Bayh-Dole Act in the United States made some key changes in the ownership of inventions developed with public R&D support. In part because of Bayh-Dole, universities have an ownership share in the intellectual property developed by those working at their institutions, and many universities set up “technology transfer offices” to provide additional support to the commercialization of research. Lach and Schankerman (2008) provide evidence consistent with greater ownership of innovations by scientists being associated with more innovation. In addition, evidence from Norway presented in Hvide and Jones (2018) suggests that when university researchers enjoy the full rights to their innovations, they are more likely to patent inventions as well as to launch start-ups. That is, ideas that might have remained in the ivory tower appear more likely to be turned into real products because of changes in the financial returns to academic researchers.

4.3 Immigration

Migration offers an alternative lens into the effects of human capital on innovation. Historically, the US has had a relatively open immigration policy compared to other developed countries, which helped to make the nation a magnet for talent. Immigrants account for about 14 percent of the US workforce but make up 17–18 percent of college graduates and 52 percent of STEM doctorates. They also account for about a quarter of all patents and a third of all US Nobel Prizes.

Kerr and Kerr (2021) go into more detail on immigration and innovation and survey policy options around migration. Broadly, a considerable body of research supports the idea that US immigrants, especially high-skilled immigrants, have boosted innovation. For example, using state panel data from 1940 to 2000, Hunt and Gauthier-Loiselle (2010) document that a one-percentage-point increase in immigrant college graduates’ population share increases patents per capita by 9 to 18 percent. Kerr and Lincoln (2010) exploit policy changes affecting the number
of H1-B visas and argue that the positive effects come solely through the new migrants’ own innovation. Bernstein et al. (2018) use the death of an inventor as an exogenous shock to team productivity and argue for large spillover effects of immigrants on native innovation (Hunt and Gauthier-Loiselle 2010 also estimate large spillovers).

The US federal government’s introduction in the early 1920s of immigration quotas with varying degrees of strictness—for example, southern Europeans, like Italians, were more strongly affected than northern Europeans, like Swedes—has been used to document how exogenous reductions of immigration damaged innovation. Moser and San (2019) use rich biographical data to show that these quotas (perhaps inadvertently) discouraged eastern and southern European scientists from coming to the United States, and that this reduced aggregate invention. Doran and Yoon (2018) also find negative effects of these quotas. Moser, Voena, and Waldinger (2014) show that American innovation in chemistry was boosted by the arrival of Jewish scientists who were expelled by the German Nazi regime in the 1930s.

Some work pushes back against this generally positive view of the impact of immigration on innovation. Using H1-B visa lotteries, Doran, Gelber, and Isen (2015) estimate smaller effects than Kerr and Lincoln (2010). In addition, Borjas and Doran (2012) actually argue for negative effects on academic publications by Americans in mathematics journals following the fall of the Soviet Union. They do not attempt, however, to estimate aggregate effects, and they note that their findings may reflect a feature specific to academic publishing where there are (in the short run) constraints on the size of academic journals and departments. In addition, Moser, Voena, and Waldinger (2014) estimate that most of the effect of immigration on innovation comes from new entry, rather than incumbents.

In summary, my reading of the literature is that good evidence demonstrates that immigration, especially skilled immigration, raises innovation. It is a particularly attractive policy because the cost of educating immigrants has been borne by other countries rather than by American taxpayer subsidies, and, unlike many other supply side policies, the increase in human capital can occur very quickly. However, there are severe political problems with relaxing immigration policy (see Tabellini 2020).
4.4 Increasing the Quality of Inventors: Lost Einsteins

4.4.1 New Facts on Inventor Backgrounds

There has long been interest in the background of inventors, with statistical analysis of this characteristic beginning with Schmookler’s (1957) study. More recent work has documented many features of inventors in near population datasets. Bell et al. (2019a) measure inventors by those individuals who are named as inventors on the patent document (both applied and granted patents), not just those who are granted the intellectual property rights (typically the assignees will be the companies that the inventors works for, rather than the individuals themselves). Looking at about 1.2 million inventors since the mid-1990s, they find that many groups are highly underrepresented, such as women, minorities, and those born into low-income families.

Using the inventor data matched to deidentified US IRS data, Bell et al. (2019a, 2019b) are able to follow potential inventors across their life cycles. Figure 3 shows the fraction of children who grow up to be inventors by the percentile of their parents’ income. There is a strong upward-sloping relationship, showing that being born to wealthier parents dramatically increases the likelihood of becoming an inventor later in life. For example, children born into the top 1 percent of the parental income distribution are 10 times more likely to grow up to be inventors than are those born in the bottom half of the parental income distribution. This is not due to wealthier children simply producing low-value innovations; conditioning on the top 5 percent of the most highly cited patents produces nearly identical results.
Figure 3: Probability of growing up to be an inventor as a function of parental income

Notes: Sample of children is 1980-84 birth cohorts. Parent Income is mean household income from 1996-2000. Source: Bell et al (2019a); Intergenerational sample

An obvious explanation for the dramatic differences in figure 3 could be that kids in poorer families have worse innate abilities than their richer counterparts. For example, if wealthier parents are smarter, their kids are likely to be smarter and, since intelligence and inventiveness are correlated, this could explain the patterns. To examine this hypothesis, Bell et al. (2019a) match math (and English) test score results from third grade and later, which are available for a subsample of the data. There is indeed a strong correlation between third grade math scores\(^3\) and the probability of becoming an inventor in later life. However, these early test scores account for only under a third of the innovation gap; they cannot account for the vast majority of the innovation-parental income relationship.\(^4\) Figure 4 illustrates this by separating

---

\(^3\) Bell et al. (2019a, 2019b) cannot observe math scores before third grade, but it is likely that these partly reflect nurture rather than nature. As the work by Heckman and others has shown, early childhood experience has effects on cognitive and noncognitive outcomes at very young ages.

\(^4\) For example, we can statistically “give” the distribution of math test scores of rich kids to poor kids using the DiNardo, Fortin, and Lemieux (1996) reweighting technique.
the inventor-ability gradient by whether a child was born in the top quintile of the parental income distribution or bottom four quintiles. For both “rich” and “poor” children the probability of growing up to be an inventor rises with math ability and is especially strong for kids in the top 10 percent of the test score distribution. However, even for kids who are in the top 5 percent of talent for math, figure 4 shows that those from richer families are far more likely to become inventors.

**Figure 4: Relationship between math test scores and Probability of becoming an inventor**

![Graph showing the relationship between 3rd grade math test scores and probability of becoming an inventor](image)

**Source:** Bell et al (2019a); New York City sample.

Interestingly, later test scores become more informative for inventor status: eighth grade math test scores account for just under half of the inventor-parental income gradient. By the time we know which college young people attended (e.g., MIT or Stanford), the role of parental income is tiny. Of course, being born to a poor family means that the chances of going to a top college are very, very low. This suggests that an important part of the transmission mechanism between parental income and later outcomes is through the quality of schooling—something we return to below when discussing policy.

A similar story holds for gender and race (see also Cook and Kongcharoem 2010). About 18 percent of inventors born in 1980 were female, up from 7 percent in the 1940 cohort. At this rate of improvement, it would take another 118 years to achieve gender parity. Looking at the
New York City data, there is essentially no difference in the third grade math ability distribution for boys and girls (even in the right tail). With regard to race, 1.6 per 1,000 white children who attended New York City public schools become inventors compared to 0.5 per 1,000 Black children. Early ability accounts for only a tenth of these differences.\(^5\)

Rather than ability differences, an alternative explanation for the patterns in figure 3 is that it reflects a misallocation of talent. There has been a flourishing of work in recent years suggesting that large amounts of productivity are lost due to such frictions (e.g., Celik 2018; Hsieh and Klenow 2009). Hsieh et al. (2019), for example, estimate that 40 percent of the growth in US GDP per person between 1960 and 2010 is due to reductions in discrimination against women and Black people. Under this view, if disadvantaged groups were given the same opportunities as their similarly talented but more privileged peers, many more of them could have pursued an inventor career and increased the quality and quantity of aggregate human capital. For example, Bell et al. (2019b) estimate a potential quadrupling of aggregate US innovation from reducing such barriers.

Bell et al. (2019a) document that an important cause of the lower invention rate of disadvantaged groups appears to be differential exposure rates to inventors in childhood. They measure exposure by family environment, proxies for the work network of parents, and innovation rates in the commuting zones where kids grew up. They find a strong association between the probability of growing up to be an inventor and measures of childhood exposure to inventors. Figure 5, for example, shows that children growing up in a commuting zone with a high density of inventors are much more likely to become inventors as adults. About 5.5 children in 1,000 in the San Jose, California, commuting zone (which encompasses Silicon Valley) become inventors, compared to about 1 in 1,000 in Brownsville, Texas.

---

\(^5\) Cook (2014) shows that racist violence between 1870 and 1940 led to 1,100 “missing patents,” compared to 726 actual patents among African American inventors.
Figure 5: Growing up in a high innovation area, makes it much more likely you will become an inventor as an adult

The relationship between place and outcomes appears to be causal. For example, it is not simply the fact that kids who grow up in Silicon Valley are more likely to be inventors; they are more likely to invent in the detailed technology classes (relative to other classes) that the valley specializes in (say, software compared to medical devices). Girls who grow up in places where there is a disproportionate fraction of women compared to men inventors are more likely (than boys) to grow up to become inventors. Furthermore, kids who move to high-innovation areas at an earlier age are more likely to become inventors than kids who move at a later age, again suggesting a causal impact of place.

This “exposure-based” view of invention could lead to much larger welfare losses than in the standard talent misallocation models. In Hsieh et al. (2019), for example, barriers to entry into occupations (the R&D sector in this case) mean a loss of talent. However, since their model is a fully rational Roy sorting model, only the marginal inventors are discouraged from becoming inventors. Great inventors—like Einstein or Marie Curie—will never be put off. In the exposure-
based model, however, even very talented people from (say) a poor family may end up not becoming inventors because they are never exposed to the possibility. Bell et al. (2019b) show evidence in favor of this and argue for large first-order welfare losses.

4.4.2 Some Policies toward the Lost Einsteins

If we took seriously the idea that much talent is being lost because of a lack of exposure to the possibility of becoming an inventor, what are the appropriate policy responses?

A classic set of responses would focus on improving conditions in disadvantaged neighborhoods, particularly in schools. These are justified on their own terms, but the misallocation losses add to the usual equity arguments. It would make sense to target resources on those most likely to benefit: disadvantaged kids who show some early promise. Figure 4 shows that being in the top 5 percent of third grade math scores was a strong predictor of future inventor status. This suggests looking into programs that identify early high achievers from underrepresented minorities.

One example is Card and Giuliano (2016), who review the effect of in-school tracking for minorities. They look at one of the largest US school districts, where schools with at least one gifted fourth (or fifth) grader had to create a separate “gifted/high achiever” (GHA) classroom. Since most schools only had a handful of gifted kids per grade, most seats in the GHA classroom were filled with nongifted students who were high achievers in the same school grade. They served as upper-track classes for students based on past achievement. Moreover, since schools were highly segregated by race and income, the program effectively treated a large number of minority students who would typically not be eligible for standard “gifted and talented” interventions.

Card and Giuliano (2016) use a regression discontinuity design to examine the causal effects on students who are tracked since selection is based on a continuous measure of past achievement with a threshold. They found that students significantly improve their math, reading, and science when assigned to a GHA classroom, but these benefits were overwhelmingly concentrated among Black and Hispanic participants. Minorities gained about 0.5 standard deviation units in math and reading scores, a result that persisted until at least the sixth grade (where their data end). These are very substantial gains, comparable in magnitude to “high performance” charter schools evaluated by Angrist, Pathak, and Walters (2013). A concern is that the gains of the participating minorities were at the expense of those who were left behind.
To address this, the paper uses a cohort difference-in-differences design comparing schools that tracked to those that did not. They find no evidence of negative (or positive) spillovers from this analysis. The effects do not appear to be coming from teacher quality or peer quality. Rather, the authors suggest that teacher expectations may play a very important role in exposing students to the possibility of greater learning.

Changing to in-school tracking has little financial cost, as there is not an expansion of the number of teachers, classes, or school day. The in-school tracking results from a reallocation of existing resources. This suggests that such interventions could yield very large benefits in terms of growth as well as equity.

Card and Giuliano (2016) look at the short-term outcomes of within-school tracking. By contrast, Cohodes (2020) examines the long-term effects of a similar program in Boston Public Schools’ Advanced Work Class (AWC) program. Pupils who do well on third grade test scores are placed in the AWC program and receive a dedicated classroom with high-achieving peers, advanced literacy curricula, and accelerated math in later grades. While the students who participate in AWC tend to be more advantaged than Boston Public School students as a whole, about half of AWC students are Black or Latino, and two-thirds of them receive subsidized school lunch.

Cohodes (2020) estimates the effect of the program using a fuzzy regression discontinuity design (RDD) comparing those who scored just above and just below the admissions threshold. There is a large increase in high school graduation for minority students. Perhaps most importantly, AWC boosts college enrollment rates. The program increases college enrollment by 15 percentage points overall, again with gains primarily coming from Black and Latino students. This results in a 65 percent increase in college enrollment for Black and Latino students, most of it at four-year institutions. Using estimated earnings associated with colleges from Chetty et al. (2017) as a measure of college quality, AWC appears to increase college quality by about $1,750 for all students and $8,200 for Black and Latino students, though these differences are not statistically significant.6

Although attending an AWC class boosts the average test scores of peers by over 80 percent of a standard deviation, Cohodes (2020) finds little evidence to support peer effects as an explanation for AWC impacts. While AWC teachers have a higher value added, the change is not large enough to account for the gains in college attendance observed here. Instead, it appears that AWC is the beginning of a chain of events that causes

---

6 Although attending an AWC class boosts the average test scores of peers by over 80 percent of a standard deviation, Cohodes (2020) finds little evidence to support peer effects as an explanation for AWC impacts. While AWC teachers have a higher value added, the change is not large enough to account for the gains in college attendance observed here. Instead, it appears that AWC is the beginning of a chain of events that causes
Bui, Craig, and Imberman (2014) is sometimes seen as a counterexample, as their RDD analysis of a gifted and talented program found no effect. However, the paper does find an effect on science outcome, which may be the critical element for inventors. Furthermore, the paper does not look at heterogeneity of the treatment effect by parental income or minority status.

Another set of targeted policies is around mentorship. Many nonprofit foundations (such as the Lemelson Foundation and the Conrad Foundation) run “inventor education” programs targeting disadvantaged children in middle and high schools. Important parts of the program are hands-on experience of problem solving in the local community, and meeting inventors who look like the targeted groups (e.g., women scientists for girls). More generally, one can imagine internship and work exchange programs aimed at young people who would not normally be exposed to high-innovation environments.

Gabriel, Ollard, and Wilkinson (2018) have developed a useful survey of a wide range of “innovation exposure” policies focusing on school-age programs. Although there is a large number of such programs (science competitions being a leading example), they tend to be dominated by students with higher-income parents, boys, and nonminorities. Moreover, the programs are almost never subject to evaluation. One immediate priority should be devoting resources to researching their impact.

5 Conclusions

Innovation is at the heart of growth, and increasing the supply of potential inventors would seem the natural place to start to think about innovation policy. Yet the literature has tended to focus much more on policies that raise the demand for innovation through the tax system or through direct government grants, rather than policies that intervene on the supply side. At one level, this is surprising: if supply is inelastic, then demand side policies may do little to the volume of innovation and may merely increase the wages of R&D scientists. On another level, it is unsurprising: supply side policies will tend to work better in the long run, which makes them harder to empirically evaluate.

In this paper, I have looked at several different human capital policies for innovation: increasing STEM, immigration reform, university expansion, and exposure policies for the participants to stay on track for college throughout high school.
disadvantaged. Clean causal identification of policies is rarer here than in other areas, but there have been some recent and encouraging contributions. In the short run, liberalizing high-skilled immigration is likely to yield a high return. In the longer run, I suggest that exposure policies may produce the greatest effect, but much more work needs to be done in evaluating the effectiveness of such policies.

References


957–70.


<table>
<thead>
<tr>
<th>#</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1762</td>
<td>Sarah Flèche, Anthony Lepinteur, Nattavudh Powdthavee</td>
<td>The importance of capital in closing the entrepreneurial gender gap: a longitudinal study of lottery wins</td>
</tr>
<tr>
<td>1761</td>
<td>Elodie Djemaï, Andrew E. Clark, Conchita D'Ambrosio</td>
<td>Take the highway? Paved roads and well-being in Africa</td>
</tr>
<tr>
<td>1760</td>
<td>Sabrina T. Howell, Jason Rathje, John Van Reenen, Jun Wong</td>
<td>Opening up military innovation: causal effects of 'bottom-up' reforms to U.S. defense research</td>
</tr>
<tr>
<td>1759</td>
<td>Marcus Biermann</td>
<td>Remote talks: changes to economics seminars during Covid-19</td>
</tr>
<tr>
<td>1758</td>
<td>Yatang Lin, Thomas K.J. McDermott, Guy Michaels</td>
<td>Cities and the sea level</td>
</tr>
<tr>
<td>1757</td>
<td>Maria Cotofan, Robert Dur, Stephen Meier</td>
<td>Does growing up in a recession increase compassion? The case of attitudes towards immigration</td>
</tr>
<tr>
<td>1756</td>
<td>Jo Blanden, Andrew Eyles, Stephen Machin</td>
<td>Trends in intergenerational home ownership and wealth transmission</td>
</tr>
<tr>
<td>1755</td>
<td>Martin Beraja, David Y. Yang, Noam Yuchtman</td>
<td>Data-intensive innovation and the State: evidence from AI firms in China</td>
</tr>
<tr>
<td>1754</td>
<td>Rafael Dix-Carneiro, João Paulo Pessoa, Ricardo Reyes-Heroles, Sharon Traiberman</td>
<td>Globalization, trade imbalances and labor market adjustment</td>
</tr>
</tbody>
</table>
| 1753 | Niklas Gohl  
Peter Haan  
Elisabeth Kurz  
Felix Weinhardt | Working life and human capital investment |
| 1752 | Holger Breinlich  
Harald Fadinger  
Volker Nocke  
Nicolas Schutz | Gravity with granularity |
| 1751 | Bernardo Guimaraes  
João Paulo Pessoa  
Vladimir Ponczek | Non-compete agreements, wages and efficiency: theory and evidence from Brazilian football |
| 1750 | Jack Blundell | Wage responses to gender pay gap reporting requirements |
| 1749 | Andrew E. Clark  
Conchita D'Ambrosio  
Simone Ghislandi  
Anthony Lepinteur  
Giorgia Menta | Maternal depression and child human capital: a genetic instrumental-variable approach |
| 1748 | Lee Elliot Major  
Andrew Eyles  
Stephen Machin | Unequal learning and labour market losses in the crisis: consequences for social mobility |
| 1747 | Tom Kirchmaier  
Monica Langella  
Alan Manning | Commuting for crime |
| 1746 | Andrew E. Clark  
Maria Cotofan  
Richard Layard | The true returns to the choice of occupation and education |
| 1745 | Yuhei Miyauchi  
Kentarō Nakajima  
Stephen J. Redding | Consumption access and agglomeration: evidence from smartphone data |