3

Bottlenecks: Sectoral Imbalances and the US Productivity Slowdown

Daron Acemoglu, *Massachusetts Institute of Technology and NBER,* United States of America

David Autor, *Massachusetts Institute of Technology and NBER,* United States of America

Christina Patterson, *University of Chicago Booth School of Business and NBER,* United States of America

I. Introduction

One of the most enduring macroeconomic puzzles of the past several decades is the pervasive slowdown in productivity growth across industrialized nations, despite breakneck advances in information and communication technologies (ICT) and electronics. Figure 1 provides a glimpse of recent breakthroughs in ICT and electronics by plotting the distribution of patents granted over the past several decades.[1] Two patterns are evident from the figure: first, a rapid takeoff in the total number of patents in the 1980s; and second, a surge in the share of ICT and electronics patents during the same time interval. Between 1990 and 2010, the total number of patents granted rose from 99,000 to 208,000, and the combined number of ICT and electronics patents granted increased by approximately 87,000, accounting for the bulk of the increase. Figure 2 depicts the growth rate of total factor productivity (TFP) in the US economy and in the leading Organisation for Economic Co-operation and Development (OECD) economies in recent decades. Productivity growth in the United States has been minimal since the mid-’00s, and it has been slower still in many OECD countries, with the possible exception of Germany.

How can these facts be reconciled? The exponential advance of innovations in ICT and electronics has led some commentators to conclude that we are on the verge of a new age of abundance, or even “technological singularity,” driven by “superintelligent” machines (e.g., Kurzweil 2005; Diamandis and Kotler 2012; Bostrom 2014). Others looking at the

NBER Macroeconomics Annual, volume 38, 2024.
 https://doi.org/10.1086/729196

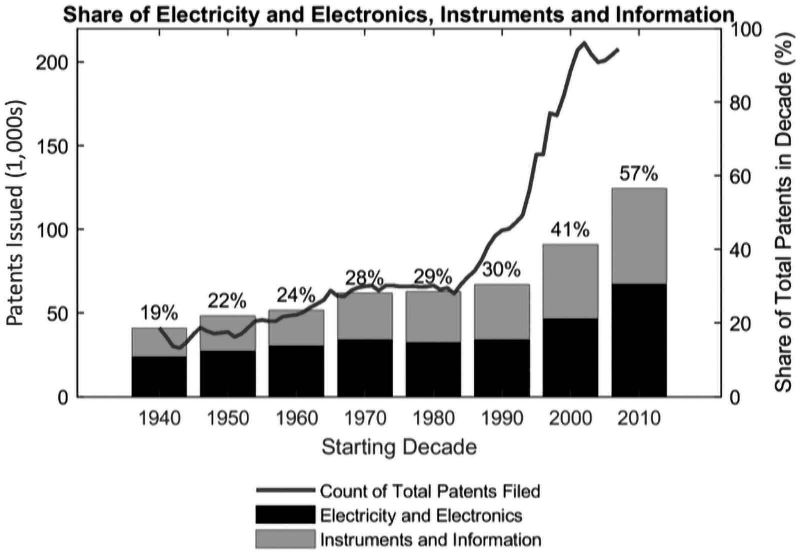

Fig. 1. Counts of US patents issued 1940–2010 and shares in (i) Electricity and Electronics and (ii) Instruments and Information. This figure plots the evolution of the counts and share (among all US utility patents) of Electricity and Electronics as well as Instruments and Information patents. Specifically, the left-hand y-axis gives the count of US utility patents issued in each year (black line), and the right-hand y-axis corresponds to the share of patents granted in each decade that are in Electricity and Electronics (black) and Instruments and Information (gray). Instruments and Information is synonymous with information and communications technologies (ICT). A color version of this figure is available online.

TFP data conclude that we have entered an age of slower growth because the most impactful technologies have already been developed and exploited (e.g., Cowen 2011; Gordon 2017).[2]

This paper offers a potential reconciliation of these trends based on the idea that technological advances over the past several decades have been unbalanced across sectors and have thus created endogenous bottlenecks, holding back aggregate productivity. We propose a simple framework in which the development of new technologies or products in a given sector requires simultaneous improvements in the quality of several inputs. For example, breakthroughs in automotive technology cannot be achieved solely with improvements in engine-management software and safety sensors but will also require complementary improvements in energy storage, drivetrains, and tire adhesion. Consequently, when some of those innovations, say batteries, do not keep pace with the rest, we may simultaneously observe rapid technological

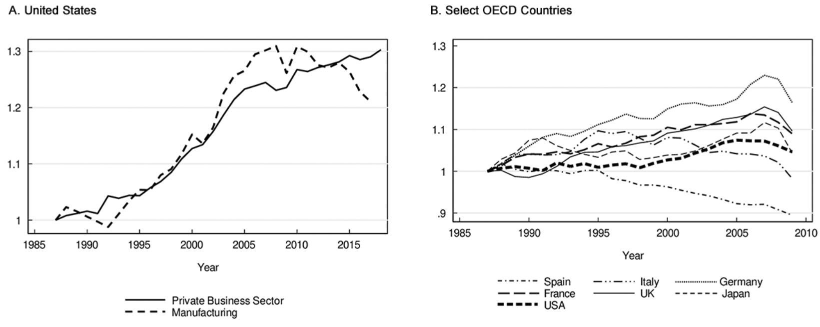

Fig. 2. Time series for aggregate total factor productivity (TFP). This figure plots the time series for aggregate TFP for the US private business sector and manufacturing (left panel) and for selected Organisation for Economic Co-operation and Development (OECD) countries (right panel). The US TFP in the left panel is normalized to 1 in 1987 and spans 1987–2017 (data from the Bureau of Labor Statistics Major Sector and Major Industry TFP database). All TFP series are normalized to 1 in 1987 in the right panel as well and span 1987–2009 (data from the 2012 release of the EU KLEMS Growth and Productivity Accounts). A color version of this figure is available online.

progress in a subset of inputs and yet slow productivity growth in the aggregate. The bottleneck created by slow progress in battery technology, in this example, is endogenous in the sense that it is the advances of nonbattery inputs that have caused batteries to become a bottleneck.

Our perspective also emphasizes how a more balanced distribution of technological progress (and research and development) can improve aggregate productivity performance. In fact, current bottlenecks may offer the potential for significantly faster aggregate productivity growth: rapid progress in these technologies could enable broader gains that are held back at present.

Several transformative technologies of the past 3 decades illustrate how bottlenecks emerge and how their alleviation can accelerate innovation and growth. High-energy-density rechargeable batteries, which power the mobile electronics and electric vehicle industries (figuratively and literally), provide a key example. Batteries were a bottleneck even prior to the 1970s, when the best available technology for rechargeable batteries (lead-acid electrochemistry) had low energy density, a slow charging rate, a short lifecycle, and an unwelcome property of releasing explosive hydrogen gas during recharging. Lead-acid batteries were succeeded in the 1970s by nickel-cadmium and nickel metal hydride (NiMH) cells, which enabled the first commercially successful gasoline-electric "hybrid" car, the Toyota Prius, introduced in 1997. However, the primary drive unit in the Prius remained a conventional gas engine;

its NiMH battery provided only supplemental electric propulsion and regenerative braking capacity. The battery bottleneck was substantially overcome by lithium-ion batteries, invented in 1973 and refined in the 1980s. The lithium-ion battery's high energy density not only enabled fully electric vehicles for mass production but also catalyzed a host of unforeseen innovations: a surge in onboard automotive processing power, enabling vehicle autonomy; battery-powered drone aircraft, now used in weather forecasting, emergency response, construction planning, filmmaking, and building inspection; and the emerging electric passenger-airplane industry. In awarding the Nobel Prize in Chemistry 2019 to John B. Goodenough, M. Stanley Whittingham, and Akira Yoshino for their invention of the lithium-ion battery, the Nobel committee observed that their work had enabled the "wireless revolution."[3]

Even more foundational to the current era is the transistor, an electronic switch that is capable of amplifying, switching, and rectifying electrical signals (Park, Steigerwald, and Walker 1976). Through the 1950s, electromechanical switches and vacuum tubes were a clear bottleneck. Though used in all kinds of electronic devices, telephone lines, radios, transmitters, audio amplifiers, and early computers, they were bulky, fragile, and slow (Sosa 2013). The transistor supplied a tiny, fast, and (ultimately) very cheap, mass-produced alternative to vacuum tubes, thus breaking the bottleneck that had impeded progress in technologies as disparate as computers, long-distance telephones, and audio amplifiers. Due to its extraordinary switching speed, the transistor also ushered in the age of digital communications. Many of the central technologies of the present—the internet, artificial intelligence (AI), mobile computing, digital imaging, autonomous vehicles—are transistor-dependent innovations that were largely unforeseen prior to digital switching. The transistor is estimated to be the most-manufactured device in history, at 13 sextillion (10^{21}) units to date, with billions more produced each day (Laws 2018; Iancu 2019). The transistor's immense footprint is also visible in figure 1, where the patenting surge in electronics and ICT would not have been feasible without this breakthrough technology.

The Global Positioning System (GPS) constitutes a third innovation that broke a technological bottleneck and enabled a suite of technologies that have become foundational to modern life. Historically, navigating an offshore or airborne vessel required either sight lines to charted objects or a combination of optical instruments, precise clocks, and detailed tables to track progress. Traditional navigation was supplemented with radio positioning systems in the 1970s, but these tools suffered from either poor accuracy or limited geographic coverage and hence did not penetrate beyond military and commercial shipping applications. GPS overcame these shortcomings and added a second crucial feature: time-keeping with atomic-level accuracy. First launched in 1978, GPS satellites now provide geolocation and date and time information to any GPS receiver on or near the earth. Although GPS was built by and for the US military, it was opened to worldwide public use in 1983, after a Korean commercial airliner inadvertently navigated over Soviet airspace and was shot down. In addition to breaking the geopositioning logjam, GPS enabled a set of highly consequential innovations that were surely not envisioned by the military planners who commissioned the system. These include precision agriculture, mining, and oil exploration; atomic-precision time information for synchronization of power transmission systems; remote surveying for geology and weather prediction; and innumerable consumer-facing services such as ride hailing, targeted advertising, and object trackers.

We first outline a simple conceptual framework that helps formalize the ideas embodied in the earlier examples. In our model, technological advances (modeled as quality improvements) in a given sector depend upon simultaneous improvements in the sector's supplier industries. Although advances in each upstream sector are potentially beneficial, these advances are complements, so that an imbalance among them is detrimental to further innovation. Our conceptual framework thus emphasizes that a balanced distribution of technological advances across sectors is important for the viability of further innovations. This mechanism is distinct from a standard Neoclassical channel where changes in input prices cause a sector to move along a fixed production possibility frontier. Our framework yields a simple estimating equation that links growth in sectoral TFP to both the average TFP and the dispersion (variance) of TFP among that sector's inputs. We estimate this equation using 462 manufacturing industries between 1977 and 2007 and also for the entire US economy between 1987 and 2007 by combining our manufacturing data with 42 nonmanufacturing industries.

Our estimates indicate that greater dispersion of TFP growth among an industry's suppliers exerts a powerful negative influence on its own growth opportunities. Our preferred specification suggests that doubling the variance of input-supplier TFP growth for a sector is associated with about 0.9 percentage points slower TFP growth for that sector.

We further document that, as conjectured, the dispersion of TFP growth among key industries has increased significantly over the past several decades. Our estimates suggest that this higher dispersion can, in an accounting sense, explain essentially all of the aggregate productivity slowdown in manufacturing between the 1970s and 2007. For example, our results imply that if the cross-industry dispersion of TFP growth in manufacturing had remained at the 1977–87 level, then aggregate TFP growth in manufacturing would have been slightly faster (rather than considerably slower) in 1997–2007 than in either of the previous 2 decades.

Our methodology also clarifies which sectors are major bottlenecks and singles out a number of industries—including pharmaceutical preparation, basic inorganic chemicals, electronic connectors, and surface active agents—as the leading bottlenecks. According to our results, a 20% decrease in the TFP growth of the 10 fastest-growing industries and a simultaneous increase in the TFP growth of each of the bottom 50% of industries—so as to keep average upstream TFP growth the same—would have led to 0.6 percentage points higher aggregate TFP growth in manufacturing. In addition, our estimates reveal that surgical and medical instruments, gas engines, and industrial valves are among the most consequentially bottlenecked sectors—meaning that they are large contributors to gross domestic product (GDP) but are inhibited by high TFP growth dispersion among their suppliers.

We confirm that these empirical patterns are broadly robust. They hold for the entire economy, and within the manufacturing sector (where TFP is better measured), they are present in weighted and unweighted specifications, in different subperiods, with varying additional controls, and with alternative measures of productivity dispersion. We also verify that these patterns are not driven by outliers, nor are they exclusively due to the rapid advances in computers and electronics sectors (though these sectors do play a central role in our results).

There is an obvious endogeneity concern in the results we present: technological trends or productivity shocks may affect supplier and customer sectors simultaneously, which could cause us to conflate the impact of sectoral linkages with correlated shocks. As a partial remedy to this threat, we exploit international (non-US) technological opportunities as an external source of identification for the variance of supplier TFP growth and obtain very similar results. We also document that it is the contemporaneous dispersion of TFP among suppliers, not the future dispersion, that predicts an industry's own TFP growth.

Another important concern relates to whether these results could be driven by relative price effects that change input intensity (e.g., less innovative inputs become more expensive and are used less intensively).[4] We show that this is unlikely to be the case. For one, we document that our results are driven by TFP, not by quantities and prices. More important, we document a similar relationship in patents: sectors with greater patenting variance across "idea suppliers" are less likely to patent themselves.[5] We also establish the same relationship at the firm level: firms facing greater variance of patenting activity across the patent classes that they cite are less likely to patent themselves.

Finally, we document analogous patterns using international data and establish that dispersion in productivity among key domestic and international supplier industries has also been a major impediment to productivity growth for several leading OECD economies.

We view our results as suggestive of a potentially important linkage between (endogenous) productivity bottlenecks and productivity growth. Although further work is needed to test whether unbalanced sectoral innovation is indeed constraining aggregate productivity growth, our evidence raises the possibility of a more nuanced explanation for the productivity slowdown experienced by industrialized nations than is available in current literature. Our analysis further suggests that, following major breakthroughs in sectors acting as bottlenecks, there should be an acceleration of both industry and aggregate productivity growth.

A conceptual issue raised by our paper is whether the dispersion of productivity growth across sectors is inefficiently unbalanced. High dispersion may result either from evolving technological opportunities or from inefficient allocation of research effort across industries. Our strategy is not geared toward identifying which allocation would be most efficient. Nevertheless, our evidence indicates that a more balanced trajectory of technological change would generate substantial aggregate gains.

Our paper is related to a small but growing literature on the causes of the productivity slowdown. Alongside the views that productivity growth is high but mismeasured or, alternatively, that good ideas are becoming increasingly scarce, several other perspectives may help to explain the productivity slowdown.[6] First, and most closely related to our work, several authors have argued that productivity growth from new technologies, especially from new general-purpose technologies, tends to lag the underlying breakthroughs substantially because the relevant sectors only slowly discover how to harness new technological capabilities. This idea was first proposed in the economics literature by David

(1990) in the context of the effects of the electrification of American industry, which David argued took place after considerable delays. It was further elaborated by Bresnahan and Trajtenberg (1995) and Helpman and Trajtenberg (1996), who proposed mechanisms for the slow emergence of productivity gains from general-purpose technologies. Closer still to our hypothesis, Brynjolfsson, Rock, and Syverson (2021) argue that productivity gains from AI and other digital technologies will trace a J-shaped curve because complementary investments and capabilities will take time to develop. Our approach, emphasizing that imbalanced innovation across sectors will act as a bottleneck, provides a specific mechanism for extensive delays in the realization of productivity gains from new technologies and platforms. Differently from these works, our paper emphasizes how the extent and duration of the productivity slowdown depend on the sectoral imbalance of innovation and the speed with which breakthroughs can take place in lagging sectors—rather than just slow adjustment in general-purpose technologies.

Second, Andrews, Criscuolo, and Gal (2016) provide evidence suggesting that, although leading firms have continued to experience steadily growing productivity, much of the aggregate productivity slowdown is related to the poor productivity performance of nonleader firms across various sectors and countries. Several other works have emphasized specific market imperfections or failures as contributing to the productivity slowdown. These include barriers to innovation and entrepreneurship (Decker et al. 2017; Aghion et al. 2019; Akcigit and Ates 2019); overinvestment in automation (Acemoglu and Restrepo 2019); insufficient government investment in research and development (Gruber and Johnson 2019); and patent rent-seeking by so-called nonpracticing entities ("trolls"), which discourages further innovation (Cohen, Gurun, and Kominers 2016). Our explanation is complementary to these ideas but distinct in its focus on productivity interactions across sectors rather than on sector-specific or aggregate factors.

Conceptually, our framework builds on models of input-output (IO) and idea linkages. Acemoglu and Azar (2020) provide a framework where innovation depends on the endogenous combinations of inputs a sector uses. Our approach here is related but emphasizes that innovation depends on the advancement of (and the balance across) the set of exogenously specified inputs. Our framework also relates to the motivating model in Acemoglu, Akcigit, and Kerr (2016), where patenting activity in a sector depends on the number of patents in "upstream" sectors that the given sector typically cites, and to the more detailed investigation of differential knowledge flows over the ideas/citation network in the recent work by Liu and Ma (2021). The key distinction between our approach and prior work is our focus on the drag that dispersion across sectors imposes on aggregate innovation and productivity growth.

The rest of the paper is organized as follows. Section II presents a motivating conceptual framework that will guide our empirical exploration. Section III overviews our data sources. Section IV presents our main results, focusing on the variance of supplier TFP growth as the measure of sectoral imbalance of innovation. This section also draws out the quantitative implications of our estimates and establishes their robustness. Section V provides several pieces of evidence that support our claim that the variance of supplier TFP growth captures the effects of imbalanced innovation across sectors. Section VI presents analogous results for a cross-country panel, and Section VII concludes. Additional information on our data, industry correspondences, and robustness checks are presented in the online appendix, http://www.nber.org/data-appendix/c14854/appendix.pdf.

II. Model

In this section, we provide a motivating conceptual framework, which will then be used to derive our estimating equations.

A. *Basic Setup*

Our starting point is the idea that new product or quality innovations in a sector depend on improvements in the quality of the inputs that they use—a point emphasized by our case studies of technological bottlenecks in the introduction. To develop this idea with minimal complexity, we consider a framework that borrows elements from existing models of IO linkages (e.g., Long and Plosser 1983; Acemoglu et al. 2012; and especially Acemoglu and Azar 2020) and also from canonical quality-ladder models (e.g., Grossman and Helpman 1991; Aghion and Howitt 1992).

Suppose that there are N sectors, denoted by $i = 1, 2, \ldots, N$. Assume also that the production function of sector i at time t is

$$Y_{it} = B_i A_{it} L_{it}^{1-\sum_{j \in S_i} \alpha_{ij}} \prod_{j \in S_i} X_{ijt}^{\alpha_{ij}}. \tag{1}$$

Here, Y_{it} denotes the output of sector i at time t, A_{it} is the productivity of this sector at time t, and B_i is a normalizing constant.[7] In addition, each sector uses labor, L_{it}, and inputs that are necessary for production, X_{ijt}, which are those in the time-invariant set S_i.[8] For simplicity, these inputs are assumed to be combined with a constant returns to scale Cobb-Douglas technology, where α_{ij} are input shares and $1 - \Sigma_{j \in S_i} \alpha_{ij}$ is the share of labor in production.

We model technological improvements by using a quality-ladder structure. In particular, we assume that $A_{jt} = \lambda^{n_{jt}}$, where $\lambda > 1$ and n_{jt} are the number of innovations this sector has experienced in the past. Each innovation, therefore, increases productivity by a factor of λ.

Our critical assumption is that the arrival rate of innovations depends on the distribution of input technologies that the sector uses:

$$\phi_{it} = H\left(\sum_{j \in S_i} \alpha_{ij} h\left(A_{jt}\right)\right), \tag{2}$$

where ϕ_{it} denotes the arrival rate of innovations at time t, h and H are monotone continuous functions, and we normalize $H(0) = 0$.[9] Different choices for these functions give different relationships between the distribution of a sector's input quality and its innovation propensity. For example, we could take $h(x) = x^{\rho}$ and $H(x) = x^{1/\rho}$ to obtain a constant elasticity of substitution (CES) aggregator. Particularly important in this context is whether the function h in equation (2) is convex or concave. The former indicates that innovation in each sector is determined by its most advanced inputs, which means that innovations across input sectors are substitutes, implying that greater (mean-preserving) dispersion of technological know-how across inputs helps innovation. Alternatively, the concave case arises when innovations across different input sectors are complements, so that greater (mean-preserving) dispersion hinders innovation. We consider the concave case to be empirically relevant because it captures the intuitive idea, highlighted by the case studies in the introduction, that new product and quality improvements require simultaneous improvements in a range of inputs, and that if some of the relevant inputs fall behind, they will act as a bottleneck, slowing technological progress.[10] In both the convex and the concave cases, because h and H are monotone, a higher level of technology for any input always helps innovation in the sector in question.

A second-order Taylor expansion of the right-hand side of equation (2) around its mean gives

$$\phi_{it} \approx H\left[\alpha_{ij} h(\bar{A}_{it}) + h''(\bar{A}_{it})\mathrm{var}(\{\alpha_{ij}A_{jt}\}_{j\in S_i})\right],$$

where $\bar{A}_{it} \equiv \Sigma_{j\in S_i}\alpha_{ij}A_{jt}$ is the (cost-share weighted) mean of the productivities of the inputs to sector i, and $\mathrm{var}(\{\alpha_{ij}A_{jt}\}_{j\in S_i})$ is the (weighted) variance of those productivities. Next, taking a first-order expansion of H around 0 and also approximating $h(\bar{A}_{it})$ by $h'(\bar{A}_{it})\bar{A}_{it}$ gives

$$\phi_{it} \approx \eta^i_{\mathrm{mean}}\bar{A}_{it} + \eta^i_{\mathrm{variance}}\mathrm{var}(\{\alpha_{ij}A_{jt}\}_{j\in S_i}), \tag{3}$$

where $\eta^i_{\mathrm{mean}} \equiv H'(0)h'(\bar{A}_{it})$ represents the effect of the mean productivity of the technological advances across inputs, which we always control for in our empirical work, while $\eta^i_{\mathrm{variance}} \equiv H'(0)h''(\bar{A}_{it})$ captures the effect of dispersion across inputs (holding the mean constant). Equation (3) will be the basis of our empirical work. The estimates of the parameter η_{variance} will show whether, in terms of our framework, the function h is convex or concave. This coefficient will also indicate the extent to which the imbalance of innovations across key input sectors in the economy may hold down aggregate productivity growth.[11]

To illustrate this point succinctly, suppose that $S_i = S$ for all i and some $S \subset \{1, \ldots, N\}$ and that $\alpha_{ij} = \alpha_j$ for all i and $j \in S$. Suppose also that h is concave, so that $\eta_{\mathrm{variance}} \equiv H'(0)h''(\bar{A}_t) < 0$, and we start with $A_{jt} = \bar{A}_t$ for all $j \in S$. Then, consider a mean-preserving spread of the A_{jt}'s so that the weighted variance, $\mathrm{var}(\{\alpha_{ij}A_{jt}\}_{j\in S_i})$, is given by σ^2. Equation (3) implies that the aggregate productivity of the economy will be reduced by $\eta_{\mathrm{variance}}\sigma^2$. So, if σ^2 and η_{variance} are both large, there will be a sizable negative impact on aggregate productivity.[12]

B. *Endogenous Innovation Effort*

It is straightforward to endogenize innovation and characterize the general equilibrium.[13] Although endogenous innovation does not play an important role in our empirical work, it is nevertheless useful to consider it to motivate our later discussion of potential inefficiencies from unbalanced innovative efforts. We add this channel to the model by modifying equation (2) to

$$\phi_{it} = \frac{1}{\gamma} H\left(\sum_{j\in S_i}\alpha_{ij}h(A_{jt})\right)^{1-\gamma} z^{\gamma}_{it}, \tag{4}$$

where $\gamma \in (0, 1)$ and z_{it} is research effort devoted to innovation in industry i at time t (e.g., overall research spending or research-related resource use, such as scientific effort). This specification implies that there are intratemporal diminishing returns to research effort in a given field, which could arise from crowding out when multiple researchers simultaneously pursue similar ideas. We include $1/\gamma$ as a constant in front of the H function, for simplicity. Note also that the H function here represents a pure knowledge externality, and thus the fact that sector i builds on the industries in the set S_i does not generate additional profits for these industries.

Suppose also that the per unit cost of research in industry i is κ_i, and the reward to an innovation in the sector at time t is π_{it}. The cost κ_i depends on the opportunity cost of research-related resources in nonresearch activities and may also include sector-specific distortions, as well as misperceptions or fads among researchers (i.e., motivations of researchers to pursue a particular field beyond its social value). We interpret the reward π_{it} as a market outcome determined by prices, market sizes, and markups (though here also, fads and misperceptions may affect rewards as well).

Given this setup, privately optimal research effort devoted to sector i at time t will be

$$z_{it}^{*} = \left(\frac{\pi_{it}}{\kappa_i}\right)^{\frac{1}{1-\gamma}} H\left(\sum_{j \in S_i} \alpha_{ij} h(A_{jt})\right),$$

and thus

$$\phi_{it}^{*} = \frac{1}{\gamma}\left(\frac{\pi_{it}}{\kappa_i}\right)^{\frac{\gamma}{1-\gamma}} H\left(\sum_{j \in S_i} \alpha_{ij} h(A_{jt})\right), \tag{5}$$

which is proportional to the exogenously specified success probability in equation (2). This ensures that equation (3) applies as before and highlights that whether the probability of successful innovation is endogenous or exogenous is not central for our empirical work.

Equation (5) emphasizes that, if the cost of research, κ_i, varies across sectors for reasons unrelated to the social cost of innovation in sector i, the unequal (unbalanced) rates of technological progress across sectors could be inefficient. In such a scenario, policies that reduce the dispersion of technological progress rates across sectors would improve the allocation of resources. For example, if the marginal cost of innovation were the same across sectors, a social planner could reduce dispersion without affecting the mean productivity of new innovations, thereby improving aggregate productivity (and welfare). Conversely, if differences in κ_i across sectors reflect differences in the social costs of innovation, then it may be infeasible to reduce the sectoral dispersion of technological progress without lowering mean productivity in the economy. Because we do not know where differences in the rate of innovation across sectors come from, these observations caution against drawing strong normative conclusions from the results that follow.

III. Data Sources

The data sources that form the backbone of our paper combine time series for industry TFP growth with IO linkage data. For manufacturing industries, we use data from the National Bureau for Economic Research and Center for Economic Studies (NBER-CES) Manufacturing Industry Database.[14] These data are sourced from the Annual Survey of Manufacturers and include annual industry-level data for 1958–2011 on output, employment, input costs, investment, capital stocks, TFP, and industry-specific prices. We include 462 manufacturing industries, corresponding to six-digit North American Industry Classification System (NAICS) codes. In accordance with the literature, TFP is defined as the residual change in real output after subtracting the (cost-share weighted) change in each of five factors: capital, production labor, nonproduction labor, energy, and nonenergy materials. We supplement the manufacturing data with annual TFP estimates for 42 nonmanufacturing industries, corresponding to three-digit NAICS codes, from the Bureau of Labor Statistics (BLS) Major Sector and Major Industry TFP database 1987–2011. As with the manufacturing data, TFP outside of manufacturing is defined as the difference between real output growth and a shares-weighted combination of growth in five inputs: capital, labor, energy, materials, and purchased services.[15]

We construct IO tables using the detailed Make and Use tables provided by the US Bureau of Economic Analysis for 1977–2007, which are available every 5 years, corresponding to the years of the Economic Census. These tables provide information on the amount that each industry produces of various commodities and the amount that they spend on each commodity, respectively. From these two tables, we construct our basic IO network, $\{\alpha_{ijt}\}$, whose entries are the dollar value of inputs that industry i uses from industry j at time t relative to the dollar value of its total intermediate costs. Because each year's release of these tables uses industry coding particular to that year's classification, we convert each table to a set of time-consistent NAICS-based industry codes, the details of which are documented in the appendix, http://www.nber.org/data-appendix/c14854/appendix.pdf. Table A1 presents summary statistics across upstream (supplying) and downstream (customer) industries. Panel *A* shows results for only manufacturing industries 1977–2007, and panel *B* depicts averages for all industries 1987–2007. In the former, we see that the average 5-year TFP growth across manufacturing sectors was 1.8 percentage points. The average TFP growth of upstream manufacturing industries is substantially higher, at 3.3 percentage points, reflecting the fact that more-productive industries are used more intensely as intermediate inputs.

To explore innovation outcomes directly, we look at patent data, starting from the Fung Institute Patent Data Project at the University of California, Berkeley, which spans the years 1976–2016. These data include every patent application and patent granted by the US Patent and Trademark Office (USPTO) during this time period. Although the data do not include patents granted outside the United States, they contain patents filed at the USPTO by non-US firms. The data include classification codes, application dates, and (importantly) cross-citations to other patents. Firm names and locations are cleaned using machine learning and natural language processing (see Balsmeier et al. 2018 for additional details on the disambiguation algorithm). The patent classification codes refer to 633 unique Cooperative Patent Classification (CPC) classes. We construct a time series that tracks the total number of patents in each CPC class by application date, as well as a similar time series for the patenting activity of each firm.[16]

We also use these data to construct the CPC-level citation network (what Acemoglu et al. 2016 refer to as the "innovation network"), which represents the knowledge flows between CPC classes. Specifically, following Acemoglu et al. (2016), we calculate a citation network, γ_{cj}, whose entries are the fraction of citations to patents in CPC class j among total citations of patents in CPC class c. To achieve greater precision and remove the time-dependent measurement error problem introduced by the increasing number of patents over time, we use the average number of citations for each class over the entire sample. We exclude all within-CPC citations, meaning citations by patents in CPC c to other patents within-CPC c.

Likewise, we construct firm-level citation networks. In this case, we calculate a citation network, ω_{kc}, whose entries are the share of citations by firm-k patents (i.e., patents that belong to firm k) to the patents of other firms within the CPC class, c. We exclude all within-firm citations (i.e., citations by firm-k patents to other firm-k patents).

Last, we supplement the domestic, US data with data for select European countries. We use data on value-added and TFP from the 2012 EU KLEMS Growth and Productivity Accounts. In this exercise, we use data from 1987 to 2007 for 30 industries in Austria, Finland, France, Germany, Italy, the Netherlands, Spain, and the United Kingdom. We combine these data with country-specific IO tables from the Groningen Growth and Development Centre (GGDC) World Input-Output Database for 2000. The relevant entry in the world IO table, $\alpha_{ik,jl}$, is the share of inputs for industry i in country k that came from industry j in country l.[17] Panel C of table A1 (http://www.nber.org/data-appendix/c14854/appendix.pdf) presents the TFP growth for this sample. Overall, the average 5-year TFP growth within this sample was 4.6 percentage points. For patenting outside of the United States, we use Google Patents global patent data from Liu and Ma (2021), which contains patents from more than 40 major patent authorities around the world. Each patent is assigned to a geographical unit using the country of residence of the inventor, the country of residence of the assignee(s), and country of the patent authority, in that order. We construct the number of patents in each country in each CPC code in each year, using the date of application for each patent. We further restrict our attention to the 20 countries with the most patenting over the sample period.

IV. Sectoral Imbalances and Productivity Growth

This section presents our main results, linking the TFP growth of an industry to the dispersion of productivity growth among its suppliers—with this dispersion representing an imbalance under our hypothesis. Concretely, we estimate a version of equation (3), derived earlier, using data on 462 six-digit NAICS-based manufacturing industries between 1977 and 2007, and 42 three-digit nonmanufacturing industries during 1987–2007. We also report the quantitative implications of these estimates and document their robustness to additional controls, different sample periods, and sources of variation in productivity growth.

A. Main Results

Our main estimating equation is the empirical analogue of (3):

$$\begin{aligned}\Delta \mathrm{TFP}_{it} = {} & \beta_{\mathrm{mean}} \sum_j \alpha_{ijt-1} \Delta \mathrm{TFP}_{jt} + \beta_{\mathrm{variance}} \mathrm{VAR}\left(\Delta \mathrm{TFP}_{jt}\right) \\ & + \mathbf{X}_{it-1}^{'} \beta_{\mathrm{other}} + \delta_t + \varepsilon_{it}\end{aligned} \tag{6}$$

where t refers to 5-year time periods, ΔTFP_{it} is the TFP growth of industry i during the 5-year time interval denoted by t,

$$\text{VAR}\left(\Delta\text{TFP}_{jt}\right) \equiv \sum_{j} \alpha_{ijt-1}\left(\Delta\text{TFP}_{jt} - \sum_{j} \alpha_{ijt-1}\Delta\text{TFP}_{jt}\right)^{2},$$

and $\Sigma_j \alpha_{ijt-1}\Delta\text{TFP}_{jt}$ is the average TFP growth among the suppliers of industry i during the 5-year time period, calculated using the α_{ijt-1}'s as weights. Recall that α_{ijt} represents the ratio of industry i's spending on inputs from industry j relative to its total intermediate spending time t. The variance of TFP growth among the suppliers of industry i is also computed using these cost shares as weights. In addition, $\mathbf{X}'_{it-1}$ denotes a vector of other (predetermined) covariates, which in some specifications includes sector fixed effects, introducing sector-specific linear trends; δ_t denotes a full set of time dummies; and ε_{it} is a heteroscedastic and (potentially) serially correlated error term, capturing all omitted factors.

This equation is comparable to our model-derived equation (3), with several operational refinements. First, we use TFP growth as our primary measure of innovation because we do not have direct measures (though we will look at patenting as well). Second, instead of relating innovation to the level of technology across inputs, as in equation (3), we link TFP growth in each sector to the TFP growth rate across inputs, because the level of TFP is not well defined. Third, we have included an error term and additional covariates. Fourth, instead of the sector-specific coefficients in front of the mean and the variance in equation (3), η^i_{mean} and η^i_{variance}, we have imposed constant coefficients, which should be interpreted as local average treatment effects.

Throughout, we always control for the mean effect of supplier TFP growth, but the main coefficient of interest for our study is β_{variance}, which captures the effect of supplier TFP growth dispersion (or innovation dispersion, in the case of our patent analyses) on a sector's productivity (innovation), holding constant the mean of supplier TFP growth (innovation). We expect this coefficient to be significantly negative if, as we hypothesize, imbalances in the rates of technological progress across an industry's suppliers impose a productivity penalty on the industry.

Table 1 reports estimates of equation (6) for 462 six-digit manufacturing and 42 three-digit nonmanufacturing industries. Panel A is for manufacturing industries, where TFP estimates are more reliable and available for a longer time period. Panel B combines the manufacturing and the nonmanufacturing industries to include the full set of sectors. All models

Table 1
Relationship between Industry TFP Growth and Supplier TFP Growth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. Manufacturing Only							
Input average	.425	.810	.653	.676	.255	1.096	.530	.187
	(.139)	(.130)	(.074)	(.170)	(.122)	(.372)	(.132)	(.156)
Input variance		−.744	−.912			−.617	−.624	
Input bottom decile		(.121)	(.118)	.059	.378	(.255)	(.198)	.367
				(.113)	(.091)			(.129)
Input top decile				−.110	−.081			−.068
				(.033)	(.032)			(.039)
Ind. fixed effects	no	no	yes	no	yes	no	yes	yes
Industry weighting	None	None	None	None	None	Nom. Sales	Nom. Sales	Nom. Sales
Observations	2,772	2,772	2,772	2,772	2,772	2,772	2,772	2,772
R^2	.108	.133	.371	.118	.361	.159	.598	.598
	B. All Industries							
Input average	.343	.915	.780	.636	.387	.708	.365	.248
	(.178)	(.161)	(.119)	(.183)	(.170)	(.399)	(.230)	(.268)
Input variance		−.905	−1.087			−.303	−.712	
Input bottom decile		(.158)	(.191)	.164	.422	(.280)	(.280)	.295
				(.099)	(.115)			(.190)
Input top decile				−.117	−.139			−.154
				(.034)	(.035)			(.058)
Ind. fixed effects	No	No	Yes	No	Yes	No	Yes	Yes
Industry weighting	None	None	None	None	None	Nom. Sales	Nom. Sales	Nom. Sales
Observations	2,016	2,016	2,016	2,016	2,016	2,016	2,016	2,016
R^2	.079	.102	.399	.090	.395	.033	.522	.531

Note: This table reports estimates of equation (6). The dependent variable is an industry's total factor productivity (TFP) growth in a 5-year period, and the two key right-hand-side variables are mean and variance of TFP growth among that industry's suppliers. Time dummies are included in all regressions, and industry dummies (corresponding to linear industry trends) are included in columns 3, 5, 7, and 8. Columns 1–5 report unweighted ordinary least squares regressions, and columns 6–8 use the industry's 1987 share of shipments as weights. Panel *A* is for manufacturing industries only 1977–2007, and panel *B* is for all industries 1987–2007. Industries are defined using 1997 North American Industry Classification System codes. Standard errors are clustered at the industry level.

include time fixed effects, and each specification includes an alternative with industry fixed effects, allowing each industry to have its own linear time trend in TFP. Odd-numbered columns include no covariates other than time dummies, whereas even-numbered columns also include industry fixed effects, thus allowing each industry to have its own linear time trend in TFP. The standard errors account for arbitrary heteroscedasticity and serial correlation at the industry level throughout. Our baseline regressions, shown in columns 1–5, are unweighted. We weight industries by their share of 1987 nominal sales in columns 6–8.

Column 1 shows the relationship between industry TFP growth and mean supplier (upstream) TFP growth, focusing only on the first term in equation (6). We detect a positive relationship between mean supplier TFP growth and downstream industry TFP growth. Adding the variance term in columns 2 and 3 strengthens the effect of mean supplier TFP growth and, more importantly, shows a precisely estimated and quantitatively large negative relationship between the variance of supplier TFP growth and industry TFP growth. For example, in our baseline specification, column 2 of panel *A*, the coefficient estimate of the variance term is -0.744 (standard error $= 0.121$). Adding linear industry trends in column 3 modestly increases this coefficient to -0.912 (standard error $= 0.118$). When we include nonmanufacturing industries in panel *B*, the point estimates are similar and only slightly larger. Figure 3 depicts the industry-level variation that produces these estimates. Specifically, we report binscatters for the regression model in column 2 of panel *A*. The left panel depicts the strong positive relationship between average supplier TFP growth and downstream industry TFP growth, and the right panel showcases the strong negative relationship between the variance of supplier TFP growth and downstream industry TFP growth.

The specification in equation (6) is a natural one, using the variance term to capture the effects from supplier TFP growth dispersion, as in our second-order approximation earlier. Nevertheless, it is useful to see whether well-performing and poorly performing supplier sectors both affect TFP growth. To investigate this question, columns 4 and 5 replace the variance term with TFP growth in the 10th and 90th percentiles of the (weighted) TFP distribution of suppliers (as we continue to control for mean supplier TFP growth). Consistent with our hypothesis, holding mean supplier TFP fixed, higher bottom-decile supplier TFP growth predicts faster own-industry TFP growth, whereas the top-decile supplier TFP growth predicts slower own-industry TFP growth

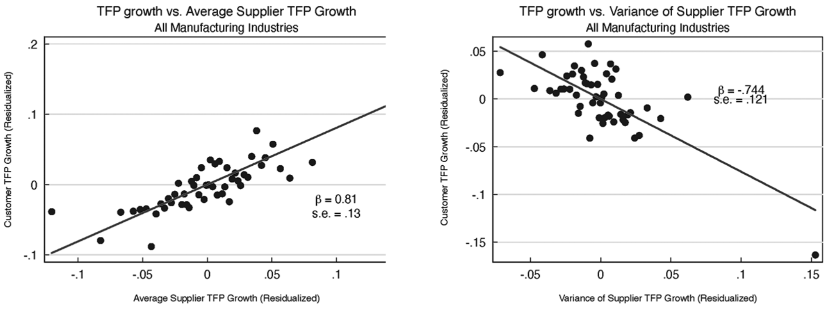

Fig. 3. Bottleneck patterns: distribution of upstream TFP growth. This figure reports binscatters (50 bins) for the regression model in table 1 from panel *A*, column 2 for the (conditional) relationship between manufacturing TFP growth and either the mean (left panel) or the variance (right panel) of supplier TFP growth. Specifically, the left panel plots the residuals from independent regressions of the x- and y-axis variables on the supplier variance of TFP growth, with time fixed effects. The right panel plots the residuals from independent regressions of the x- and y-axis variables on the supplier average of TFP growth, with time fixed effects. A color version of this figure is available online.

(with these relationships typically exhibiting statistical significance at the 5% level or below).

Last, columns 6 through 8 replicate our main specifications, but now using nominal industry sales in 1987 as weights. These weighted estimates are very similar to the unweighted specifications. For example, in column 6, the coefficient on the variance term is -0.617 (standard error $= 0.255$), which is only a little smaller than the estimate in the corresponding unweighted specification in column 2, -0.744.

Overall, the estimates in table 1 uniformly show a negative estimated impact of TFP growth dispersion across a sector's suppliers on own-industry TFP growth. In terms of our motivating conceptual framework, this suggests that productivity growth in a sector is held back when advances among its suppliers are unbalanced. In the rest of the paper, we demonstrate the robustness of these results and document a similar relationship in innovation activity. Before moving on to innovation, we draw out the quantitative implications of the productivity growth estimates in the next subsection.

B. *Quantitative Implications*

The results in table 1 imply that an imbalance in productivity growth across sectors could be a drag on aggregate growth. Temporarily deferring robustness checks, we explore whether such sectoral imbalances could be a quantitatively meaningful contributor to the productivity slowdown in the United States. For this to be the case, two conditions must be satisfied. First, the coefficient estimates in table 1 must be economically large. Second, the dispersion of sectoral TFP growth must have increased over the decades during which we witnessed the productivity slowdown.

Figure 4 confronts the latter issue by plotting the evolution of the variance of TFP across manufacturing industries. Panel *A* of figure 4 depicts the simple variance of TFP growth across all manufacturing industries, and panels *B* and *C* show the average variance of industry supplier TFP growth: for manufacturing only and for the economy overall, respectively. Both within upstream manufacturing and across all manufacturing industries, there was a striking rise in the dispersion of sectoral productivity growth in the US economy over the past several decades. This is true both overall and when weighting industries by their input share. Quantitatively, the TFP variance in manufacturing was about 0.002 before the mid-1970s and now is three times as large, around 0.006. As suggested by the patenting time series in figure 1, the electronics and computer sector accounts for a large portion of the increase in TFP variance through the 1990s. The right-hand-side plots of figure 4 document that when this sector is taken out, the rise in the variance of TFP growth is noticeably smaller—though still present—in recent decades. When we zoom out to include nonmanufacturing supplier industries, there is a similarly large increase in the variance of TFP growth from the 1980s to the present, but the pattern is not monotone, perhaps reflecting the fact that TFP is measured less reliably outside of manufacturing.

How much of the productivity slowdown can be explained by the rising variance of TFP growth? Figure 5 addresses this question by applying our (nominal sales-) weighted estimates reported in column 7 of panels *A* and *B* in table 1. We find a sizable productivity penalty from TFP growth dispersion. The estimates imply that TFP dispersion reduced manufacturing TFP growth significantly in both the 1987–97 and the 1997–2007 periods, as shown by the gray bars in the figure. If, counterfactually, input TFP variance remained at its 1977–87 value throughout the sample, then during 1987–97, instead of a 0.8 percentage point slowdown in aggregate TFP growth, we would have seen a 1.5 percentage point faster growth (as shown by the counterfactual patterned bar). Similarly, during 1997–2007, instead of the much slower 3.3% average TFP growth, the US manufacturing sector is predicted to have had only a mild TFP slowdown, to 6.5%. In

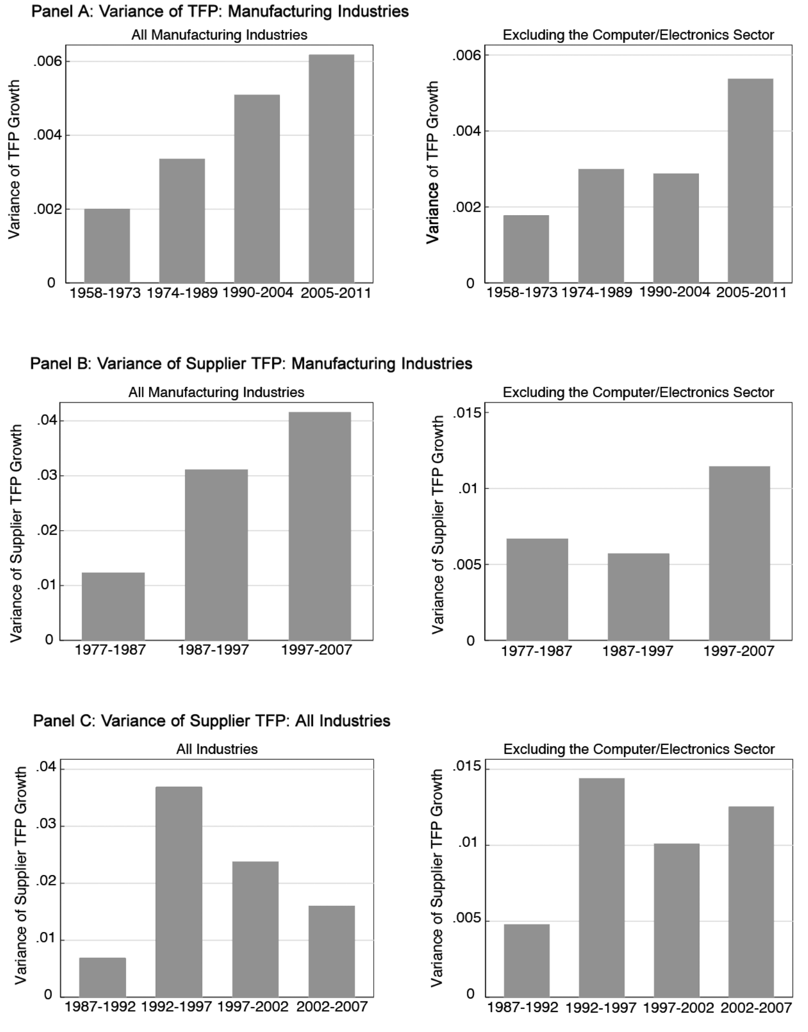

Fig. 4. Variance of total factor productivity (TFP) growth. This figure plots the variance of TFP across manufacturing industries, variance of supplier TFP across manufacturing industries, and variance of supplier TFP across all industries. Each industry observation is weighted by its share of total nominal sales. Panel *A* is for the variance of TFP growth across manufacturing industries for each 5-year period, spanning 1958–2011 (averaged into 15-year bars). Panel *B* reports the variance of supplier TFP growth across 462 six-digit North American Industry Classification System (NAICS)-based manufacturing industries, again for 1977–2007 (averaged into 10-year bars). Panel *C* reports the variance of supplier TFP growth across all industries (adding 42 three-digit nonmanufacturing industries). Figures on the right exclude the computer and electronics sector (NAICS 334). In panels *B* and *C*, the input-output network is defined at the beginning of each 5-year period. A color version of this figure is available online.

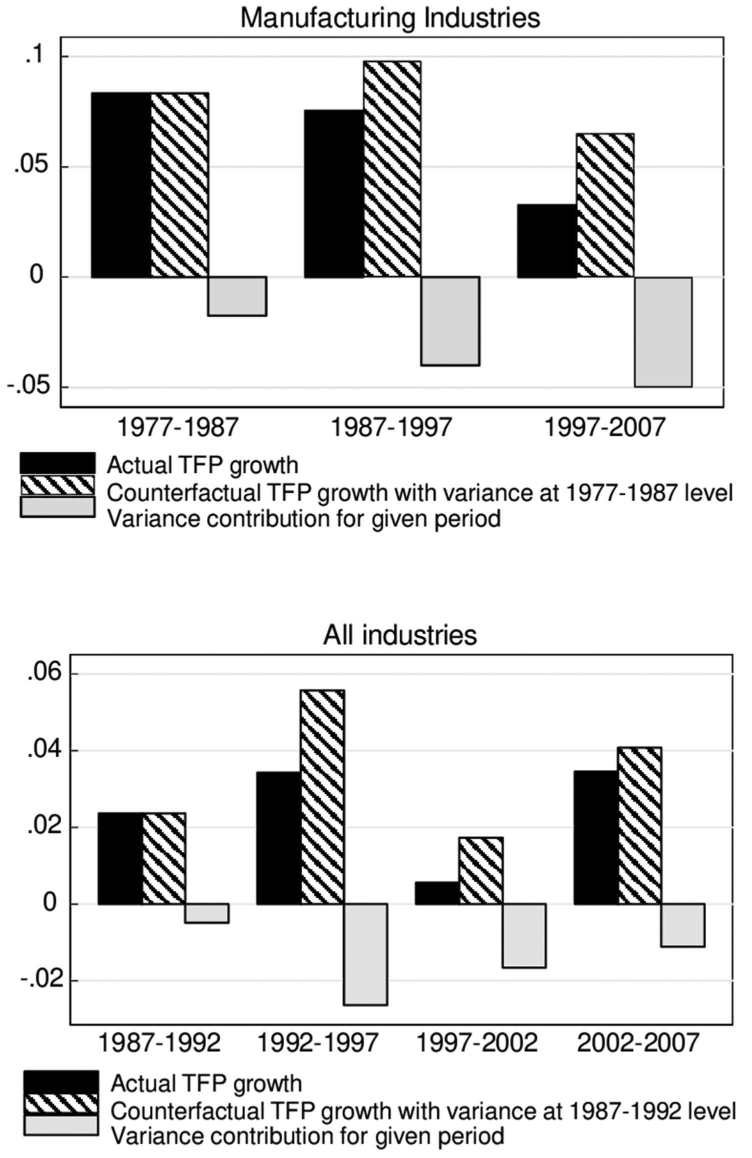

Fig. 5. Magnitude of bottleneck estimates. This figure reports actual and counterfactual total factor productivity (TFP) growth, and the contribution from supplier variance, for manufacturing and all industries for the periods 1977–87, 1987–97, and 1997–2007. The counterfactuals are based on regression estimates from the column 7 specification of table 1. Counterfactual TFP (patterned bars) is computed from the regression coefficients as the TFP growth that would have been observed in the given period if the variance of TFP growth had remained at the same level as during the initial period (1987–92). Specifically, we calculate counterfactual TFP growth by subtracting the contribution of the increase in supplier variance relative to 1977–87; the 1977–87 supplier variance is shown with the black bar, and by construction, counterfactual TFP growth in 1977–87 is equal to actual growth in this period. A color version of this figure is available online.

the lower panel of this figure, we see similar patterns for all industries. Thus, the quantitatively sizable estimates in table 1 can potentially account for the bulk of the US productivity slowdown in recent decades. We emphasize that these magnitudes are suggestive but far from definitive, given the limitations of our measurement and identification.

To provide more detailed insight into these aggregate relationships, we explored the identities of the sectors that have contributed to this quantitative effect. The variance of supplier TFP in manufacturing increased over this period both because lagging industries failed to grow and because leading industries pulled away from the rest. Panel *A* of table 2 lists illustrative examples of the fastest-growing industries, which are defined as those that have had the largest impact on supplier TFP variance between 1997 and 2007. These industries include electronic

Table 2
Examples of Limiting and Limited Industries

Panel A: List of Select Fastest-Growing Industries That Drive Rising TFP Variance
Semiconductor and related devices
Electronic computers
Iron and steel mills
Computer storage devices
Motor vehicle electrical and electronic equipment
Panel B: List of Select Bottleneck Industries
Petroleum refineries
Pharmaceutical preparation
Turbine and turbine generator set units
Printed circuit assembly
Basic organic chemicals
Panel C: List of Select Limited (Bottlenecked) Industries
Surgical and medical instruments
Relay and industrial controls
Gasoline engine and engine parts
Guided missile and space vehicles
Industrial valves

Note: Bottleneck industries (panel *B*) are defined as those for which a 10% increase in total factor productivity (TFP) would result in the largest aggregate reduction in the variance of TFP growth across all supplying industries (i.e., VAR(ΔTFP$_{ji}$) from eq. [6]). Fastest-growing industries (panel *A*) are conversely defined as those for which a 10% increase in TFP would result in the smallest aggregate reduction in the variance of TFP growth across supplying industries. Limited ("bottlenecked") industries (panel *C*) are defined as the 50 manufacturing industries with the highest variance of TFP among suppliers, after limiting to the 100 industries with the highest value-added. Sample is restricted to 462 manufacturing industries 1997–2007. See table A2, http://www.nber.org/data-appendix/c14854/appendix.pdf, for an ordered list of the top 10 industries in each category 1997–2002 and 2002–07.

computers, computer storage devices, and semiconductors. To gauge the economic leverage of these outlier industries, consider a hypothetical mean-preserving contraction of TFP growth dispersion: reduce the TFP growth of the 10 fastest-growing industries between 1997 and 2007 by 0.2 percentage points and increase the TFP growth of each of the bottom 50% of industries just enough to keep the average TFP growth constant.[18] In this scenario, the variance of supplier TFP growth between 1997 and 2007 would have been 23% lower, and aggregate TFP growth in manufacturing would have been 0.6 percentage points higher.

The remaining panels of table 2 round out the evidence on bottleneck industries. Panel *B* reports illustrative examples of slow-growing industries that became the biggest bottlenecks over the same time period. These include pharmaceutical preparation, basic organic chemicals, printed circuit assembly, and turbine generators. Panel *C* reports example industries that are most *bottlenecked*—that is, held back by the uneven innovation across their suppliers. These include surgical and medical instruments, gas engines, and industrial valves.

C. *Endogeneity Concerns*

Because the estimates in table 1 are obtained from regressions of an industry's TFP growth on the contemporaneous TFP growth of its suppliers, productivity shocks that are common across several industries might generate mechanical correlations between our right-hand-side and left-hand-side variables. In this subsection, we explore two strategies that, in net, lend support to the case that these results are informative about the effects of productivity bottlenecks.

Our first strategy is to isolate industry productivity changes that emanate from common technological developments across several advanced economies. We do this in table 3 by exploiting changes in industry TFP in major OECD countries, as reported by the 2012 EU KLEMS Growth and Productivity Accounts. For this exercise, we focus on all 504 industries (both manufacturing and nonmanufacturing), mapped to 29 EU KLEMS industries.[19] In panel *A*, we use the mean and variance of supplier TFP in France, Germany, and the United Kingdom as instruments for the corresponding variables in the United States. To purge measurement error in these instruments, panel *B* uses the rank of TFP growth by industry within-country. In both panels, columns 1 and 2 present the baseline ordinary least squares (OLS) results, and columns 3 and 4 depict two-stage least squares (2SLS) estimates.

Table 3
Country-Specific Instruments

	OLS Estimates		IV Estimates			
	(1)	(2)	(3)	(4)	(5)	(6)
			A. Average TFP Growth			
Upstream average	.951	.780	1.369	1.416	1.387	1.509
	(.232)	(.119)	(.363)	(.655)	(.378)	(.758)
Upstream variance	−.876	−1.066	−.902	−.887	−.897	−.795
	(.155)	(.135)	(.385)	(.527)	(.391)	(.588)
Estimate	OLS	OLS	2SLS	2SLS	LIML	LIML
Ind. fixed effects	No	Yes	No	Yes	No	Yes
Observations	2,478	2,478	2,478	2,478	2,478	2,478
R^2	0	0	0	0	0	0
First-stage F-stat			1.38	.63	1.38	.63
			B. Rank of TFP Growth			
Upstream average	.951	.780	.928	1.093	.928	1.094
	(.232)	(.119)	(.338)	(.348)	(.342)	(.349)
Upstream variance	−.876	−1.066	−.667	−1.480	−.664	−1.482
	(.155)	(.135)	(.445)	(.661)	(.449)	(.665)
Estimate	OLS	OLS	2SLS	2SLS	LIML	LIML
Ind. fixed effects	No	Yes	No	Yes	No	Yes
Observations	2,478	2,478	2,478	2,478	2,478	2,478
R^2	0	0	0	0	0	0
First-stage F-stat			.8	2.1	.8	2.1

Notes: This table reports instrumental-variables (IV) estimates of equation (6) for all industries 1982–2007. The dependent variable is an industry's total factor productivity (TFP) growth in a 5-year period, and the two key right-hand-side variables are mean and variance of TFP growth among that industry's suppliers. Excluded instruments are mean and variance of supplier TFP growth in France, Germany, and the United Kingdom. All columns report unweighted regressions. Time dummies are included in all columns, and industry dummies (corresponding to linear industry trends) are included in even-numbered columns. Columns 3 and 4 report two-stage least squares (2SLS) estimates, and columns 5 and 6 report limited information maximum likelihood (LIML) estimates. Panel *A* defines the upstream moments, taking the average and variance of TFP growth across industries. In panel *B*, we rank industries in each country according to their TFP growth and calculate the input-share weighted average and variance of TFP ranks. Standard errors are clustered at the aggregated KLEMS industry level. OLS = ordinary least squares.

The first-stage F-statistics are given at the bottom of panels *A* and *B* in table 3; these are somewhat low in both panels (the full first stages are reported in table A8, http://www.nber.org/data-appendix/c14854/appendix.pdf). This motivates the limited information maximum likelihood (LIML) estimates presented in columns 5 and 6, which are consistent even in the presence of weak instruments. These estimates confirm that our findings are not driven by weak instruments.

The instrumental-variables (IV) estimates of the relationship between industry TFP growth and supplier TFP mean and variance correspond closely to our earlier OLS estimates. In both panels, the OLS and IV estimates are very similar across columns 1–2 and 3–4. For example, in columns 1 and 3 of panel *A*, which do not include industry fixed effects, the OLS coefficient on the variance term is −0.876 (standard error = 0.155), and the IV estimate for the same variable in the same specification is −0.902 (standard error = 0.385). The variance term estimates are also quite close in columns 2 and 4, where we add industry fixed effects. We see a similar pattern in panel *B* when we exploit the variation in the rank of TFP growth: −0.667 (standard error = 0.445) for the IV estimate without industry fixed effects in column 3 and −1.480 (standard error = 0.661) for the IV estimate with industry fixed effects in column 4. The LIML estimates in columns 5 and 6 are also comparable. For example, in panel *B*, the variance term's coefficient estimate is −1.482 (standard error = 0.665) with industry fixed effects—similar to the 2SLS estimates in column 4 and the OLS estimates in column 2.

The congruence between the baseline OLS estimates and the IV estimates that exploit TFP changes in other leading economies bolsters our confidence that these results are not driven by shocks that are common across US industries and their suppliers. We also note that because the IV coefficient estimates are similar to the OLS estimates from table 1, the implied quantitative magnitudes are comparable as well.

Panel *A* of table 4 explores whether there is a correlation between future average and variance of TFP across suppliers and an industry's current TFP growth. Such a relationship would be concerning for the interpretation of productivity bottlenecks as a constraint on TFP growth. Across the eight columns (the first four for manufacturing industries and the last four for all industries), we do not see any evidence that future variance of TFP growth of suppliers has a negative relationship with current TFP growth of an industry. When we focus on all industries and look at the relationship between future variance and current TFP growth, the correlation is positive, but it disappears when we include our main regressors, contemporaneous average and variance of TFP across suppliers. It is also zero across all specifications for manufacturing industries. This pattern is reassuring for our overall interpretation.

Panel *B* of the same table explores whether a similar relationship exists between the variance of TFP growth among an industry's customers and its own TFP growth. Such a correlation is a distinct possibility, because many industries have customers and suppliers that are overlapping.

The general pattern is that customer variance is also negatively correlated with an industry's TFP growth, but typically only when it is entered by itself (without the variance of TFP among input suppliers). When both sets of variables are included, the coefficient on customer variance becomes less significant and smaller, whereas the variance of TFP across input suppliers remains negative and significant. This pattern is broadly supportive of our overall interpretation, even if it raises the possibility that in some specifications, there is a high enough correlation between downstream and upstream variances that we cannot rule out additional effects working from customers' TFP growth.

Panel C of table 4 investigates whether mean-reversion dynamics may be confounding our estimates. In particular, if TFP growth is serially correlated, then failing to account for this could lead to a spurious

Table 4
Relationship between Industry TFP Growth and the Distribution of TFP Growth

	Manufacturing Industries				All Industries			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. Future Supplier TFP Growth							
Future input average	.166	.154	−.006	.083	−.003	.016	−.171	−.063
	(.145)	(.123)	(.079)	(.077)	(.178)	(.168)	(.110)	(.104)
Future input variance	.065	.011	−.010	−.066	.244	.239	.343	.149
	(.098)	(.120)	(.101)	(.103)	(.133)	(.156)	(.158)	(.146)
Input average		.787		.670		.867		.752
		(.121)		(.076)		(.158)		(.116)
Input variance		−.810		−.919		−.982		−1.061
		(.124)		(.122)		(.163)		(.186)
Ind. fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
Industry weighting	None	None	None	None	None	None	None	None
Observations	2,772	2,772	2,772	2,772	2,016	2,016	2,016	2,016
R^2	.085	.137	.334	.371	.073	.106	.375	.399
	B. Customer TFP Growth							
Customer average	.626	.499	.499	.367	.460	.344	.383	.280
	(.066)	(.066)	(.074)	(.078)	(.077)	(.078)	(.102)	(.099)
Customer variance	−.503	−.302	−.796	−.529	−.443	−.264	−.907	−.662
	(.200)	(.236)	(.161)	(.151)	(.257)	(.314)	(.265)	(.254)
Input average		.466		.454		.765		.687
		(.123)		(.083)		(.173)		(.122)
Input variance		−.566		−.634		−.778		−.784
		(.127)		(.125)		(.178)		(.201)
Ind. fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
Industry weighting	None	None	None	None	None	None	None	None
Observations	2,769	2,769	2,769	2,769	2,015	2,015	2,015	2,015
R^2	.157	.173	.373	.387	.093	.115	.393	.408

Table 4
(Continued)

	Manufacturing Industries				All Industries			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	C. Lagged TFP Growth: Dependent Variable and Supplier Metrics							
Input average	.641	.637	.509	.530	.915	.921	.724	.744
	(.101)	(.097)	(.077)	(.081)	(.151)	(.153)	(.117)	(.121)
Input variance	−.678	−.715	−.753	−.776	−.923	−.939	−.889	−1.014
	(.111)	(.112)	(.121)	(.126)	(.152)	(.163)	(.173)	(.183)
Lagged input average		.056		.208		.069		.330
		(.110)		(.085)		(.137)		(.112)
Lagged input variance		.015		−.471		−.049		−.793
		(.146)		(.130)		(.195)		(.177)
Lagged dep. var.	.089	.086	−.255	−.280	.070	.066	−.362	−.391
	(.099)	(.103)	(.042)	(.045)	(.115)	(.122)	(.045)	(.050)
Ind. fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
Industry weighting	None	None	None	None	None	None	None	None
Observations	2,310	2,310	2,310	2,310	1,974	1,974	1,974	1,974
R^2	.129	.130	.418	.425	.107	.108	.464	.476

Note: This table reports estimates of equation (6). The dependent variable is an industry's total factor productivity (TFP) growth in a 5-year period, and the right-hand-side variables are mean and variance of TFP growth among that industry's suppliers, plus lead terms, mean, and variance of TFP growth among the industry's customers, and lagged dependent variables. Time dummies are included in all regressions, and industry dummies (corresponding to linear industry trends) are included in columns 3, 4, 7, and 8. Columns 1–4 are for manufacturing industries 1977–2007 and columns 5–8 for all industries 1987–2007. All columns report unweighted ordinary least squares regressions. In addition to the mean and variance of TFP growth among an industry's suppliers, panel *A* includes the 5-year lead of the same variables. Panel *B* includes the mean and variance of TFP growth among the industry's customers. Panel *C* includes the 5-year lagged mean and variance of TFP growth among the industry's suppliers and the lag of the dependent variable (the industry's TFP growth rate). Standard errors are clustered at the industry level.

relationship between supplier TFP variance and own TFP growth. Panel *C* addresses this issue by including the lag of own TFP growth, as well as lagged input average and variance terms (in some specifications). Overall the results prove quite robust, and the statistical significance and quantitative impact of the input variance term are hardly affected. For example, in specifications that include only the lagged dependent variable, the input variance has a coefficient of −0.747 (standard error = 0.115) for manufacturing industries and −0.923 (standard error = 0.152) for all industries, similar to our baseline findings in table 1 in both cases. The estimates are again similar when we include lagged input average and variance terms.

D. Robustness

Table 5 further investigates the robustness of our results to a battery of controls and specifications. For brevity, we focus on the manufacturing sample and report analogous results for all industries in the appendix (see table A7, http://www.nber.org/data-appendix/c14854/appendix.pdf). Panel *A* documents robustness for the specification without industry fixed effects, and panel *B* includes industry fixed effects that allow for linear trends in industry TFP.

For ease of reference, columns 1 and 2 report our unweighted and (nominal sales-) weighted specifications from table 1, which include only the mean and the variance of TFP as well as time dummies. The rest of the table focuses on our unweighted specification. Column 3 estimates the same models but now using 10-year periods rather than the 5-year intervals in table 1. This specification purges higher-frequency variation in TFP and focuses on longer-term variation. The results from these models are similar to the baseline estimates.

Our estimating equation (3) defines sectors that are falling behind as those that have relatively slow TFP growth in the contemporaneous 5-year period. However, if high variance in the current period reflects mean reversion following rapid growth in the recent past, this would not correspond to an imbalance but rather to a potential rebalancing. Column 4 checks this possibility by adding the covariance between the supplier TFP growth in the current and the prior periods to our specification.[20] Intuitively, this covariance term accounts for potential persistence and reversal patterns in industry-level TFP changes. We find that the covariance of TFP across periods does not meaningfully affect the relationship of primary interest. The coefficient on the variance term in panel *A* is only slightly larger, −0.640 (standard error = 0.127), and the covariance term is relatively small and imprecisely estimated. The estimate on the covariance term is larger and statistically significant in panel *B*, but the coefficient on the variance of upstream TFP growth remains unaffected by the inclusion of this covariance term. We infer from these results that the first and second moments of the upstream TFP distribution provide informative measures of sectoral imbalances.

The subsequent columns of table 5 provide additional robustness checks. Another factor that could affect measured industry TFP is changing import penetration. Column 5 controls for average imports from China by other countries in the industry and in the input-weighted average of the supplying industries (from Autor, Dorn, and Hanson

Table 5
Robustness for Downstream TFP and Upstream TFP: Manufacturing Industries

	Baseline	Weighted	10-Year	Cov.	China Shock	No Comp.	Outlier Robust	Fixed IO	All Inputs	3-Digit Leaveout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					A. Without Industry Trends					
Input average	.810	1.096	.931	.660	.565	.560	.592	.878	1.154	.076
	(.130)	(.372)	(.182)	(.116)	(.123)	(.066)	(.043)	(.115)	(.163)	(.106)
Input variance	−.744	−.617	−.477	−.640	−.724	−1.231	−.774	−.711	−.903	−.360
	(.121)	(.255)	(.094)	(.127)	(.113)	(.587)	(.059)	(.161)	(.171)	(.120)
Input covariance				−.069						
				(.170)						
Observations	2,772	2,772	1,386	2,310	1,386	2,604	2,772	2,772	2,772	2,772
R^2	.133	.159	.100	.123	.126	.122	.208	.153	.147	.080

	B. With Industry Trends									
Input average	.653	.530	.704	.482	.502	.652	.597	.716	.968	.231
	(.074)	(.132)	(.122)	(.074)	(.100)	(.072)	(.048)	(.071)	(.097)	(.099)
Input variance	−.912	−.624	−.641	−.647	−.820	−1.197	−.757	−.961	−1.149	−.449
	(.118)	(.198)	(.107)	(.131)	(.146)	(.624)	(.068)	(.139)	(.163)	(.155)
Input covariance				−.399						
				(.143)						
Observations	2,772	2,772	1,386	2,310	1,386	2,604	2,772	2,772	2,772	2,772
R^2	.371	.598	.549	.385	.472	.252	.487	.379	.378	.338

Note: This table reports estimates of equation (6) for manufacturing industries 1977–2007. The dependent variable is an industry's total factor productivity (TFP) growth in a 5-year period, and the right-hand-side variables are mean and variance of TFP growth among that industry's suppliers plus additional controls. Time dummies are included in all regressions. Panel *B* also includes industry dummies (corresponding to linear industry trends). Column 1 repeats our baseline regression from column 2 of table 1. Column 2 weights observations by the industry's share of 1987 shipments. Column 3 uses 10-year observations. Column 4 controls for the covariance between the supplier TFP growth in the current and the prior 5-year periods. Column 5 controls for the China shock, following Autor et al. (2013). Column 6 excludes the computers and electronics manufacturing sector (NAICS 334) from the regression sample and from the construction of the average and variance of TFP growth among suppliers. Column 7 runs an outlier-robust regression (rreg). Column 8 fixes the input-output (IO) table at 1987. Column 9 defines the IO network to use the share among all inputs instead of among intermediaries. Column 10 excludes the industry's own three-digit North American Industry Classification System (NAICS) code when constructing the IO network.

2013), addressing the concern that Chinese import penetration may itself affect productivity growth (e.g., Autor et al. 2020). Accounting for imports does not appreciably change the coefficient on supplier TFP variance.[21]

We noted the importance of the electronics and computer sectors earlier. Column 6 confirms that the negative relationship between industry TFP growth and supplier TFP dispersion holds even when computers and electronics manufacturing (NAICS 334) is excluded from the estimation sample as well as from the calculation of upstream metrics. With these key sectors excluded, the variance term is less precisely estimated, as expected. Nevertheless, it remains statistically significant at the 5% level or below in all of our specifications: -1.231 (standard error $=$ 0.587) in panel A and -1.197 (standard error $=$ 0.624) in panel B. These estimates reveal that our hypothesized mechanism is present even when the ICT and electronics sectors are excluded, but also that the ICT and electronics sectors showcase our mechanism and contribute substantially to its identification and quantitative implications (as corroborated by the examples in table 2).

Column 7 shows a similar relationship to our baseline results when we estimate a robust regression that downweights outliers that have a major effect on the slope of the relationship between upstream TFP variance and downstream TFP growth. Notice that in this case, the standard error of the input variance term is much smaller, highlighting that outliers were, indeed, reducing the precision of our estimates, though not affecting their magnitudes much.

Column 8 confirms that the results are again similar when we use a fixed IO matrix, rather than the time-varying IO matrix from our baseline specification. Column 9 probes robustness to our definition of input shares. Here, we define upstream shares, α_{ijt-1}, as total-cost shares rather than as intermediate-cost shares (as in our baseline specification). These two share measures will differ to the extent that the intermediate share of total costs varies across sectors. The results are once again very similar. Finally, column 10 excludes own three-digit industry when constructing the IO network. This does change the magnitude of the coefficient estimates but not their signs or statistical significance. Table A7 (http://www.nber.org/data-appendix/c14854/appendix.pdf) shows analogous estimates for the entire economy, rather than just the manufacturing sector. These results are again similar to our baseline estimates.

Because our empirical analysis is confined to the 462 manufacturing (or the 504 total) industries, our estimates will not capture any imbalances in innovation or productivity growth that happens at more-disaggregated levels. To explore whether these more-micro imbalances may also matter, and to further probe the robustness of our results, in the appendix we use estimates of within-industry, across-establishment TFP growth from the US Census Bureau's Dispersion Statistics on Productivity. These measures of dispersion have also increased during our sample period, but table A3 documents that the average upstream TFP growth dispersion among input suppliers, when added to our regression, is not statistically significant and does not change the relationship between our measure of supplier TFP growth dispersion and own TFP growth.

A final concern is that TFP growth estimates from the NBER-CES manufacturing data do not allow for fully subtracting the contribution of intermediate inputs, and this might be one reason why they are lower than estimates at more-aggregated levels that incorporate such corrections. To verify that this aspect of the data is not responsible for our results, in tables A4–A6 we repeat our main analysis (in particular, the regression models and tables 1–4) using an adjusted TFP series. This series is computed using factor shares for each disaggregated industry that are adjusted by a factor calculated to equate the factor shares at the three-digit industry level in the NBER-CES data with those in the National Income and Product Accounts data. We apply the same factor to all disaggregated industries in our data that belong to the same three-digit industry.[22] These results show very similar patterns to our main estimates.

In summary, these results confirm that the negative relationship between industry TFP growth and supplier TFP variance is statistically significant, pervasive, and largely unaffected by the inclusion of a variety of potential confounders.

E. Prices, Quantities, and Productivity

Could these patterns be explained by mismeasurement of TFP? In a standard Neoclassical setting, industries benefit when the productivity of their suppliers increases because this reduces input costs (e.g., Acemoglu et al. 2012). If TFP is measured correctly, it will be unaffected by fluctuations in employment, demand factors, and input costs that induce industries to move along (rather than changing) their production possibility frontiers. If TFP is mismeasured, however, these Neoclassical effects could erroneously spill over to TFP estimates. If, in addition, elasticities of substitution between inputs are nonunitary, as explored in Atalay (2017) and Baqaee and Farhi (2019), changes in sectoral production may affect our TFP estimates and confound our results.

We investigate the role of these Neoclassical channels in table 6. Because these Neoclassical effects work through sectoral prices or through output changes that affect sectoral outputs and prices, we do this by adding the mean and variance of supplier prices and employment levels to our baseline regressions.[23] For comparison, columns 1 and 2 restate our baseline estimates. The alternative specifications in columns 3–8 indicate that controlling for these Neoclassical channels does not qualitatively change the relationship between supplier TFP variance and industry TFP growth (and that the coefficients on these channel variables are typically insignificant). For example, when we include the mean and variance of supplier prices in column 3 (without industry fixed effects), the TFP variance term has a coefficient of −0.686 (standard error = 0.232), which is 90% of the baseline estimate, though less precisely estimated. When we include the mean and variance of supplier employment levels

Table 6
Exploring Neoclassical Effects

	Baseline		Prices		Employment		Combined	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Input TFP average	.810	.653	.815	.692	.720	.497	.602	.436
	(.130)	(.074)	(.113)	(.083)	(.134)	(.075)	(.107)	(.089)
Input TFP variance	−.744	−.912	−.686	−.527	−.703	−.786	−.655	−.424
	(.121)	(.118)	(.232)	(.243)	(.118)	(.115)	(.233)	(.245)
Input price average			.006	.077			−.141	−.069
			(.085)	(.061)			(.091)	(.065)
Input price variance			−.051	−.329			−.123	−.381
			(.204)	(.198)			(.201)	(.201)
Input employment average					.224	.244	.264	.262
					(.045)	(.056)	(.051)	(.062)
Input employment variance					.166	−.106	.117	−.213
					(.219)	(.235)	(.218)	(.246)
Ind. fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Observations	2,772	2,772	2,772	2,772	2,772	2,772	2,772	2,772
R^2	.133	.371	.133	.373	.149	.384	.152	.387

Note: This table reports estimates of equation (6) for manufacturing industries 1977–2007. The dependent variable is an industry's total factor productivity (TFP) growth in a 5-year period, and the right-hand-side variables are mean and variance of TFP growth among that industry's suppliers plus the mean and variance of supplier prices and employment. Time dummies are included in all regressions, and industry dummies (corresponding to linear industry trends) are included in even-numbered columns. All columns report unweighted ordinary least squares regressions. Industries are defined using 1997 North American Industry Classification System codes. Standard errors are clustered at the industry level.

in column 5 (also without industry fixed effects), the coefficient estimate on TFP variance is -0.703 (standard error = 0.118), which is nearly identical to the baseline estimates in column 1. The results remain similar when we include both sets of variables (prices and employment) together. When we include industry fixed effects, the estimates are once again similar to our baseline results.

The evidence in table 6 suggests that the relationship between supplier TFP and industry TFP is not a reflection of (potentially mismeasured) Neoclassical effects. Instead, the evidence suggests that it captures economic effects that work through the innovation or product-quality mechanism identified by our model. We next offer more direct evidence on this mechanism.

V. Innovation

This section investigates whether innovation, as encoded in patents, is one of the underlying mechanisms that could explain our results. For this exercise, we replace the IO network (comprised of α_{ijt} entries) with the patent citation network (corresponding to the γ_{cj}'s capturing citation patterns across CPCs). Our sectoral analysis starts at the CPC level, but we also consider firm-level results later in this section. The main question explored in this section is whether a greater imbalance of innovation across upstream sectors or firms ("idea suppliers") reduces the innovation of a downstream sector or firm. We will see that the answer to this question is a strong yes.[24]

A. CPC-Level Results

We begin the analysis at the patent-class level and estimate the following variant of equation (6):

$$\Delta\text{Patent}_{ct} = \beta_{\text{mean}}\sum_j \gamma_{cj}\Delta\text{Patent}_{jt} + \beta_{\text{variance}}\text{VAR}\left(\Delta\text{Patent}_{jt}\right) + \delta_t + \varepsilon_{ct}, \quad (7)$$

where t refers to 5-year time periods, ΔPatent_{ct} is a measure of patenting growth within-CPC c during the 5-year time interval denoted by t,

$$\text{VAR}\left(\Delta\text{Patent}_{jt}\right) \equiv \sum_j \gamma_{cj}\left(\Delta\text{Patent}_{jt} - \sum_j \gamma_{cj}\Delta\text{Patent}_{jt}\right)^2,$$

and $\Sigma_j \gamma_{cj}\Delta\text{Patent}_{jt}$ is the average patent growth during the 5-year time period among the CPCs that are upstream to c (i.e., among the CPC codes that c, the focal CPC, cites). As indicated earlier, its entries, the γ_{cj}'s, are the share of total citations over the entire sample period from patents in CPC c that go to patents in CPC j. The upstream variance of patenting growth is also computed analogously to the upstream variance of supplier TFP, though now using the γ_{cj}'s as weights.

Table 7 presents our main estimates of the patent-based version of equation (6). The first three columns measure innovation activity by log patents, which implies that sectors with zero patenting activity are dropped (this produces a sample of around 4,326 observations for our main specifications). Columns 4–6 instead focus on the Davis-Haltiwanger-Schuh (DHS) transformation (Davis, Haltiwanger, and Schuh 1998), which allows us to define growth rates when there are zero patents in a CPC in either the beginning or the end period.[25] This expands our sample slightly to 4,379 CPC observations. Throughout, we focus on unweighted specifications.

Columns 1 and 4 show a strong positive association between average patenting activity in a sector's upstream CPCs and the sector's own patenting activity. Columns 2–3 and 5–6 add the variance of patenting activity in the upstream CPCs to proxy for the imbalance of innovation activity across sectors. The latter two columns also include CPC fixed effects, which allow for linear trends at the CPC level. The estimates show

Table 7
Bottleneck Regressions Using Patenting by CPC Code

	Log Patent			DHS Specification		
	(1)	(2)	(3)	(4)	(5)	(6)
Citation average	1.291	1.335	1.473	1.284	1.307	1.402
	(.064)	(.064)	(.093)	(.065)	(.064)	(.098)
Citation variance		−.959	−.876		−1.040	−.911
		(.266)	(.314)		(.381)	(.532)
CPC fixed effects	No	No	Yes	No	No	Yes
Observations	4,326	4,326	4,323	4,379	4,379	4,376
R^2	.245	.250	.372	.202	.207	.307

Note: Standard errors are clustered at the Cooperative Patent Classification (CPC) level. All columns report unweighted ordinary least squares estimates. Year fixed effects are included in all regressions, and CPC fixed effects are included for columns 3 and 6. All regressions consider changes across 5-year averages 1975–2014. Columns 1–3 specify patent growth using the change in the log number of patents. Columns 4–6 specify patent growth as the change in patenting activity, normalized using the Davis-Haltiwanger-Schuh (DHS) transformation: $(P_{it} - P_{t-1}/\frac{1}{2}(P_{it} + P_{t-1}))$. In both cases, these specifications apply to both the dependent variable and the citation-weighted moments (i.e., the independent variables).

a powerful negative effect of upstream variance. In column 2, for example, the coefficient estimate for the variance is -0.959 (standard error $= 0.266$). The variance estimate remains essentially unchanged in column 3 when CPC fixed effects are included. The coefficient estimates are very similar in columns 5 and 6 with the DHS transformation, though standard errors are somewhat larger.

Table 8 is the patenting analogue of table 3 from our IV analysis for TFP, but now focusing on patents and exploiting variation in upstream patenting among foreign patents contained within the Google Patents global database. The estimates in table 8 are broadly supportive of the negative relationship between upstream variance and a sector's own patenting. Panel *A* depicts specifications using the change in log patenting (as in cols. 1–3 of table 7), and panel *B* shows results with the DHS transformation (as in cols. 4–6 of table 7). In each panel, columns 1 and 2 show the OLS relationship in this sample, columns 3 and 4 report 2SLS estimates, and columns 5 and 6 present the LIML estimates (which are again motivated by the weak first stages in cols. 3 and 4). Across essentially all columns, we see negative and statistically significant estimates of the impact of upstream variance.

Following the design of table 4, table 9 explores whether downstream patenting variance also matters, the possible relationship between future upstream variance and current patenting, and whether mean-reversion dynamics may be confounding our results. Reassuringly, the results in this table confirm the robustness of the estimates in table 7 to these checks. In particular, in panel *A* future citations have a smaller and often insignificant coefficient when entered at the same time as our main citation variables, and the coefficient on our citation variance measure remains similar to the baseline estimates.

In panel *B*, downstream variance—that is, variance among citing, rather than cited, patents—is negative and significant when entered by itself, which reflects the fact that, just as in the IO network, upstream and downstream measures are correlated. Nevertheless, when we also include our upstream citation variables, the downstream variance is no longer statistically significant and is in fact positive in most specifications, whereas our upstream citation variance has a similar coefficient to our baseline estimates and is statistically significant with log patents in columns 2 and 4, though it becomes less precise with DHS in columns 6 and 8.

Finally, in panel *C* we find that the inclusion of lagged patenting (the dependent variable) and the lagged citation average and variance terms

Table 8
Bottleneck Regressions Using Cross-Country Variation in Patenting as Instruments for US-Firm Patenting

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Log Patents					
Citation average	1.335	1.473	1.495	1.788	1.497	1.802
	(.064)	(.093)	(.075)	(.144)	(.076)	(.148)
Citation variance	−.959	−.876	−1.289	−2.289	−1.300	−2.477
	(.266)	(.314)	(.452)	(1.067)	(.469)	(1.201)
Year-by-CPC FEs	No	Yes	No	Yes	No	Yes
Estimator	OLS	OLS	2SLS	2SLS	LIML	LIML
Observations	4,326	4,323	4,285	4,283	4,285	4,283
R^2	.250	.372	.162	.095	.162	.092
First-stage F-stat	0	0	14.53	8.51	14.53	8.51
	Panel B: DHS Specification					
Citation average	1.307	1.402	1.446	1.679	1.447	1.687
	(.064)	(.098)	(.076)	(.150)	(.076)	(.153)
Citation variance	−1.040	−.911	−1.206	−2.176	−1.208	−2.288
	(.381)	(.532)	(.475)	(1.097)	(.483)	(1.179)
Year-by-CPC FEs	No	Yes	No	Yes	No	Yes
Estimator	OLS	OLS	2SLS	2SLS	LIML	LIML
Observations	4,379	4,376	4,325	4,324	4,325	4,324
R^2	.207	.307	.138	.078	.138	.077
First-stage F-stat	0	0	30.8	13.73	30.8	13.73

Note: Standard errors are clustered at the Cooperative Patent Classification (CPC) level. Year fixed effects (FE) are included in all regressions, and year-by-CPC fixed effects are included where indicated. All regressions consider changes across 5-year averages 1975–2014. All observations are unweighted. We use patenting growth in the 10 countries with the most patents over the sample period as instruments (Canada, China, Germany, France, the United Kingdom, Italy, Japan, Korea, Russia, Taiwan, and the United States). Specifically, we calculate the average and variance of patenting growth for each cited-CPC code in each of the five countries. Then, we use these 20 variables—the average and variance for each of the 10 countries—to instrument the average and variance of patenting growth in each cited-CPC code across US firms. Columns 3 and 4 report two-stage least squares (2SLS) estimates; columns 5 and 6 report limited information maximum likelihood (LIML) estimates. Panel *A* specifies patent growth using the change in the log number of patents; panel *B* specifies patent growth as the change in patenting activity, normalized using the Davis-Haltiwanger-Schuh (DHS) transformation: $(P_{it} - P_{t-1}/\frac{1}{2}(P_{it} + P_{t-1}))$. In both cases, these specifications apply to the dependent variable and the citation-weighted moments (i.e., the independent variables). OLS = ordinary least squares.

has very little effect on our results when focusing on the log patents measure. This is also the parent we find with the DHS transformation without fixed effects (col. 6), though with DHS and CPC fixed effects, the coefficient on the citation variance falls and becomes statistically insignificant (col. 8).

Table 10 confirms the robustness of our main CPC-level estimates to the same battery of tests we conducted in table 7 for TFP. (For brevity, we focus on log patents as the dependent variable.) We see broadly similar patterns across specifications that include or exclude CPC trends, are weighted by their share of total patenting, are at the 10-year frequency, include the covariance term, exclude the ICT and electronics sectors, leave out all citations to patents in the focal sector's three-digit CPC, limit the sample to years before 2005 to make the sample more similar to the data used for the TFP growth analyses, or focus on patents filed by US residents. The only two specifications in which the variance term is significantly weakened are columns 5 and 8. The former of these excludes the ICT and electronics sectors, and the weaker results likely reflect the factors discussed for the analogous specification in table 5—computers and electronics are emblematic of the imbalances that are our focus, so excluding these industries weakens the relevant economic forces and the precision of the estimates. The latter, column 8, excludes 54% of total patents filed by non-US residents at the USPTO, which likely accounts for the reduced precision of these estimates.

Quantitatively, these estimates suggest that upstream innovation imbalances have a major impact on overall innovation. For example, the weighted coefficient estimate in column 2 of panel *A* in table 10 suggests that a one standard deviation higher upstream variance (which is 0.03) is associated with a decline in the growth rate of patenting in a CPC code of 0.042 log points. This is a 47% reduction relative to the weighted mean of patenting across sectors, which is equal to 0.09 (weighting by the total number of patents in the CPC code in the initial 5-year period). These numbers are in the same ballpark as those implied by our TFP models.[26]

In sum, although the results in this subsection show a few specifications where the estimates are less stable than our main results reported in the prior section, they are overall supportive of a robust negative association between upstream variance of innovation activity and downstream patenting at the CPC level.

B. *Firm-Level Evidence*

We next turn to the firm-level relationship between upstream imbalances and patenting. For this exercise, we disaggregate the patents data to the firm level and allow for variation across firms: specifically, the extent to which they rely on different CPCs for their patenting. This produces our firm-level citation network, summarized by $\{\omega_{kc}\}$,

Table 9

Bottleneck Regressions Using Patenting by CPC Code: Robustness to Lags, Leads, and Citing CPCs

	Log Patent				DHS Specification			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. Future Patenting Growth among Cited CPCs							
Citation average		1.284		1.458		1.317		1.442
		(.079)		(.098)		(.087)		(.104)
Citation variance		−.733		−.697		−1.100		−1.303
		(.242)		(.299)		(.418)		(.532)
Future citation average	.927	.064	.414	.180	.907	.024	.349	.129
	(.067)	(.077)	(.089)	(.086)	(.073)	(.088)	(.113)	(.108)
Future citation variance	−.756	−.530	−.629	−.351	−.981	−.643	−.923	−.647
	(.246)	(.233)	(.290)	(.278)	(.421)	(.505)	(.589)	(.568)
CPC fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
Observations	3,712	3,712	3,709	3,709	3,753	3,753	3,753	3,753
R^2	.193	.275	.362	.424	.167	.243	.315	.370
	B. Patenting Growth among Citing Patents							
Citation average		.919		.395		.726		.252
		(.237)		(.328)		(.296)		(.368)
Citation variance		−1.003		−1.374		−1.437		−1.628
		(.442)		(.524)		(.883)		(1.169)
Citing patent average	1.276	.429	1.516	1.152	1.262	.600	1.482	1.252
	(.066)	(.234)	(.089)	(.326)	(.067)	(.292)	(.084)	(.356)
Citing patent variance	−1.204	−.111	−.897	.240	−1.159	.296	−.848	.546
	(.245)	(.425)	(.306)	(.512)	(.299)	(.808)	(.365)	(1.011)
CPC fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
Observations	4,326	4,326	4,323	4,323	4,378	4,378	4,376	4,376
R^2	.243	.252	.375	.379	.204	.211	.312	.314

	C. Lagged Patenting Growth: Dependent Variable and Citation Metrics							
Citation average	1.278	1.260	1.460	1.394	1.365	1.225	1.526	1.428
	(.061)	(.081)	(.101)	(.102)	(.068)	(.092)	(.103)	(.105)
Citation variance	−.985	−.965	−1.092	−1.017	−.793	−.797	−.581	−.488
	(.271)	(.269)	(.340)	(.341)	(.369)	(.399)	(.493)	(.491)
Lagged citation average		.032		.509		.226		.704
		(.103)		(.115)		(.121)		(.138)
Lagged citation variance		−.061		−.332		−.073		−.165
		(.257)		(.316)		(.398)		(.449)
Lagged dep. var.	.055	.052	−.097	−.132	−.068	−.084	−.197	−.238
	(.034)	(.038)	(.033)	(.036)	(.042)	(.046)	(.040)	(.041)
CPC fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
Observations	3,695	3,695	3,693	3,693	3,743	3,743	3,742	3,742
R^2	.256	.256	.399	.406	.224	.226	.377	.388

Note: Standard errors are clustered at the Cooperative Patent Classification (CPC) level. All columns report unweighted ordinary least squares estimates. Year fixed effects are included in all regressions, and CPC fixed effects are included where indicated. All regressions consider changes across 5-year averages 1975–2014. Cols. 1–4 specify patent growth using the change in the log number of patents. Cols. 5–8 specify patent growth as the change in patenting activity, normalized using the Davis-Haltiwanger-Schuh (DHS) transformation: $(P_{it} - P_{i,t-1}/1/2(P_{it} + P_{i,t-1}))$. Panel *A* includes the leading average and variance of patent growth from the 5-year period. Panel *B* includes the average and variance of citation growth. Panel *C* includes the lagged average and variance of patent growth among idea suppliers, as well as the lagged dependent variable (downstream-CPC patenting growth), from the previous 5-year period.

Table 10
Robustness for Bottleneck Regressions Using Patenting by CPC Code

	Baseline	Weighted	10-Year	Cov.	No Comp.	3-Digit Leaveout	Excluding Post-2005	US Firms Only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A. Without CPC Trends				
Citation average	1.335	1.305	1.378	1.330	1.377	1.175	1.292	1.299
	(.064)	(.084)	(.073)	(.065)	(.074)	(.078)	(.072)	(.060)
Citation variance	−.959	−1.414	−.840	−.730	−.102	−.624	−.945	−.323
	(.266)	(.393)	(.197)	(.298)	(.260)	(.219)	(.271)	(.218)
Citation covariance				−.337				
				(.454)				
CPC fixed effects	No	No	No	No	No	No	No	No
Observations	4,326	4,305	1,853	3,094	3,783	4,326	3,098	4,224
R^2	.250	.442	.347	.279	.265	.195	.230	.207
				B. With CPC Trends				
Citation average	1.473	1.433	1.555	1.466	1.474	1.211	1.326	1.363
	(.093)	(.129)	(.112)	(.107)	(.087)	(.109)	(.113)	(.091)
Citation variance	−.876	−1.215	−.594	−.765	−.310	−.574	−.724	−.209
	(.314)	(.400)	(.247)	(.356)	(.318)	(.295)	(.370)	(.284)
Citation covariance				.179				
				(.471)				
CPC fixed effects	yes	yes	yes	yes	yes	yes	yes	yes
Observations	4,323	4,304	1,846	3,090	3,781	4,323	3,096	4,221
R^2	.372	.580	.601	.453	.369	.333	.414	.303

Note: Standard errors are clustered at the Cooperative Patent Classification (CPC) level. All columns report unweighted ordinary least squares estimates. All regressions consider stacked, sequential 5-year changes 1975–2014, except for column 3, where we consider stacked, sequential 10-year changes. In all columns, we specify patent growth using the change in the log number of patents. Time fixed effects are included in all specifications, and CPC fixed effects are included in panel *B*. Column 2 weights observations by the CPC code's share of total patenting in the sample period. Column 4 controls for covariance between the idea-supplier patenting growth in the current and the prior 5-year periods. Column 5 removes patents that belong to CPC class G, which includes computers. Column 6 excludes the CPC's own three-digit CPC code when constructing the citation network. Column 7 limits the sample to years before 2005. Column 8 limits to the patents of US-based firms.

representing citations by firm k to CPC class c. We estimate the following equation:

$$\Delta\mathrm{Patent}_{kt} = \beta_{\mathrm{mean}}\sum_{c}\omega_{kc}\Delta\mathrm{Patent}_{ct} + \beta_{\mathrm{variance}}\mathrm{VAR}(\Delta\mathrm{Patent}_{ct}) + \delta_t + \varepsilon_{kt}, \quad (8)$$

where t refers to 5-year time periods, $\Delta\mathrm{Patent}_{kt}$ is a measure of patenting growth of firm k during the 5-year time interval denoted by t,

$$\mathrm{VAR}(\Delta\mathrm{Patent}_{ct}) \equiv \sum_{c}\omega_{kc}\left(\Delta\mathrm{Patent}_{ct} - \sum_{c}\omega_{kc}\Delta\mathrm{Patent}_{ct}\right)^2,$$

and $\Sigma_c\omega_{kc}\Delta\mathrm{Patent}_{ct}$ is the average patent growth in the 5-year time period among the CPCs upstream to firm k (meaning those cited-to by firm k). As indicated earlier, these are calculated using the share of total citations over the entire sample period by firm k's patents to patents in CPC c. The variance of patent growth among the cited CPCs is computed using the ω_{kc}'s as weights.

This disaggregation produces a much larger sample, consisting of almost 2 million observations at the firm level. For many firm-period combinations, however, there are no patents. Thus, in this table, we use the DHS transformation. In particular, in columns 1–3 we use the standard DHS transformation, where observations are dropped when there are two consecutive zeros. In columns 4–6, we use a modified DHS transformation, where in such cases, the transformation imputes a value of zero.[27]

Table 11 presents the main results from this exercise. The firm-level structure of the data in this table enables us to control for firm fixed effects or for CPC-times-year fixed effects, thus purging a large fraction of the variation in patenting between firms. The general pattern is a negative relationship between upstream variance at the CPC level and a firm's own propensity to patent. For example, in column 1, the coefficient estimate of the citation variance is −0.264 (standard error = 0.042). In column 3, when we include CPC-times-year fixed effects, the coefficient increases slightly, to −0.292 (standard error = 0.045). The exception to this pattern is in column 2, where we see a positive and significant coefficient when firm fixed effects are included with the standard DHS transformation. We suspect that this is driven by firms that have many zeros and thus many missing observations. Indeed, in columns 4–6, when we use the modified DHS so that all zeros are kept, the coefficients on the variances are more stable and always negative (and strongly statistically significant except in col. 5).

Table 11
Bottleneck Patterns Using Firm-Level Patenting

	DHS Specification			DHS Specification with Zeros		
	(1)	(2)	(3)	(4)	(5)	(6)
Citation average	1.089	1.048	1.039	.234	.282	.233
	(.010)	(.024)	(.015)	(.003)	(.006)	(.004)
Citation variance	−.264	.337	−.292	−.065	−.025	−.051
	(.042)	(.083)	(.045)	(.011)	(.017)	(.012)
Firm FEs	No	Yes	No	No	Yes	No
CPC × Year FEs	No	No	Yes	No	No	Yes
Observations	654,583	617,894	640,397	1,888,705	1,888,705	1,828,778
R^2	.037	.414	.044	.009	.038	.013

Note: Standard errors are clustered at the firm level. Year fixed effects (FE) are included in all regressions and firm fixed effects or year-by-CPC (Cooperative Patent Classification) fixed effects are included where indicated. All regressions consider changes across 5-year averages 1975–2014. Observations are unweighted. In all columns, we specify patent growth as the change in patenting activity, normalized using the Davis-Haltiwanger-Schuh (DHS) transformation: $(P_{it} - P_{t-1}/\frac{1}{2}(P_{it} + P_{t-1}))$. In columns 1–3, we replace missing values with 0 for these specifications. In columns 4–6, we leave missing values as is.

Table 12 provides a number of robustness checks for these firm-level results, considering analogous specifications to those we presented for the TFP and CPC-level patenting models and focusing on the standard DHS measure. Panel *A* of this table corresponds to column 1 of table 11, and panel *B* adds CPC-times-year fixed effects, as in column 3 of that table. The results are robust across specifications that are weighted by the firm's share of total patenting, control for the lagged dependent variable, change the sample period, or focus only on domestic patents. As a further robustness test, column 6 adds the mean and variance of future citations in a firm's patenting network. Future citations as well as our main measures are now statistically significant. Given the high degree of serial correlation in patenting within classes, these patterns are not surprising. They highlight, however, that future tests of our proposed mechanism should attempt to exploit shocks that affect patenting during certain discrete periods.

Quantitatively, these estimates imply that upstream firm-level imbalances have similarly sized innovation effects as we measure at the CPC level. For example, the coefficient estimate in column 2 of panel *A* of table 12, which shows the weighted regression specification, suggests that a one standard deviation higher upstream variance (which is again 0.03, as in the CPC case in the previous section) is associated with a decline in firm-level patenting of 0.13 log points. This is a sizable (73%) decline relative to a baseline of 0.18. These numbers are similar when we include

Table 12
Robustness for Bottleneck Regressions Using Firm-Level Patenting

	Baseline	Weighted	Lagged	Excluding Post-2005	US Only	Lead Horserace
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Without CPC-by-Year Fixed Effects					
Citation average	1.089	2.029	0.593	1.242	1.074	1.014
	(.010)	(.143)	(.011)	(.012)	(.012)	(.016)
Citation variance	−.264	−2.567	−.290	−.446	−.055	−.187
	(.042)	(.407)	(.038)	(.057)	(.045)	(.059)
Lagged 5-year growth			−.261			
Future citation average			(.002)			.176
						(.018)
Future citation variance						−.341
						(.059)
CPC × Year FEs	No	No	No	No	No	No
Observations	654,583	654,583	378,905	384,258	363,911	528,162
R^2	.037	.206	.076	.020	.035	.029
	B. With CPC-by-Year Fixed Effects					
Citation average	1.039	2.391	.591	1.209	1.018	1.064
	(.015)	(.133)	(.016)	(.019)	(.019)	(.022)
Citation variance	−.292	−2.687	−.104	−.462	−.088	−.286
	(.045)	(.414)	(.041)	(.060)	(.049)	(.063)
Lagged 5-year growth			−.253			
Future citation average			(.002)			.096
						(.024)
Future citation variance						−.225
						(.063)
CPC × Year FEs	Yes	yes	yes	yes	yes	yes
Observations	640,397	640,397	373,321	380,680	356,228	520,627
R^2	.044	.233	.083	.025	.043	.035

Note: Standard errors are clustered at the firm level. Year fixed effects (FE) are included in all regressions, and Cooperative Patent Classification (CPC) by year fixed effects are included in panel *B*. In all regressions, we specify patent growth as the change in patenting activity, normalized using the Davis-Haltiwanger-Schuh transformation: $(P_{it} - P_{t-1}/\frac{1}{2}(P_{it}+P_{t-1}))$. Observations are unweighted except for column 2, where we weight observations by the firm's share of total patenting in the sample period. All regressions consider changes across 5-year averages 1975–2014. Column 3 controls for the lagged dependent variable (downstream-firm patenting growth) from the previous 5-year period. Column 4 only limits the sample to years before 2005. Column 5 only includes US-firm patents, which applies to both the dependent variable and the citation-weighted moments (i.e., the independent variables). Columns 6 and 7 include the leading average and variance of patent growth from the 5-year period.

CPC-by-year fixed effects in the same column of panel *B*. Overall, these numbers are broadly comparable to those from the CPC-level analysis in Subsection V.A.

VI. International Evidence

Our primary analysis focuses on TFP growth and innovation in the United States (except when instrumenting domestic TFP growth and innovation with contemporaneous foreign development in tables 3 and 8). We supplement this evidence here by estimating a variant of equation (6) for TFP growth across European countries. As outlined in Section III, we use the GGDC World Input-Output Database to construct consistent IO linkages for 30 industries in Austria, Finland, France, Germany, Italy, the Netherlands, Spain, the United Kingdom, and the United States. We fix the global IO table at the year 2000 and focus on industry TFP growth in this cross-country sample between 1987 and 2007. These data enable us to include international IO linkages, which we exploit in our calculations of the mean and variance of supplier TFP growth.[28]

We report these cross-country estimates in table 13. We report the baseline specifications in the first four columns. These specifications are unweighted and include combinations of country effects, year effects, year-by-country effects, and year-by-industry effects, as noted at the bottom of each column. In column 1, we focus on a specification containing country and year effects. This estimate verifies that an industry's TFP growth is predicted by the average TFP growth of its suppliers. Column 2 includes the variance of supplier TFP growth. The coefficient on this measure is negative, highly significant and broadly similar to the US-based estimate, at −0.820 (standard error = 0.211).

Subsequent columns probe the robustness of this finding. Column 3 includes industry-by-year effects, so that the identifying variation is within-industry rather than cross-industry, as in the main specifications of the paper. The relationship is similar to column 2, although somewhat smaller. In particular, the coefficient on the variance term is −0.528 (standard error = 0.142). Column 4 includes both industry-by-year and country-by-year interactions, restricting to variation within-industry and within-country. In this demanding specification, the coefficient on the variance term remains negative and statistically significant, at −0.441 (standard error = 0.158).

The negative effect of supplier TFP variance is also present when we include the lagged dependent variable to control for mean-reversion

Table 13
Evidence on Bottlenecks from Cross-Country Regressions Using TFP

	Baseline				Lagged Dep. Var.	VA Weight	VA Weight	10-Year Changes	Within-Country IO
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Upstream average	.258	.270	.107	−.225	.264	.278	.266	.276	.280
	(.074)	(.078)	(.079)	(.112)	(.085)	(.103)	(.084)	(.115)	(.081)
Upstream variance		−.820	−.528	−.442	−.815	−.560	−.758	−.722	−.713
		(.211)	(.142)	(.158)	(.180)	(.579)	(.420)	(.362)	(.215)
Year FEs	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes
Country FEs	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Year × Country FEs	No	No	No	Yes	No	No	No	No	No
Year × Industry FEs	No	No	Yes	Yes	No	No	No	No	No
Lagged dep. var.	No	No	No	No	Yes	No	Yes	No	No
Observations	982	982	982	982	896	982	896	462	982
R^2	.065	.076	.364	.401	.120	.062	.192	.119	.075

Note: This table reports estimates of equation (6) for 1987–2007 using cross-country observations. The dependent variable is total factor productivity (TFP) growth of an industry in a given country in a 5-year period, and the two key right-hand-side variables are mean and variance of TFP growth among that country-industry pair's suppliers. All regressions are unweighted unless otherwise indicated. Time and country dummies are included in all regressions. The sample includes 30 industries in nine countries: Austria, Finland, France, Germany, Italy, the Netherlands, Spain, the United Kingdom, and the United States. Columns 1 and 2 are cross-country analogues of columns 1 and 2 in table 1. Column 3 includes industry-by-year fixed effects (FE), and column 4 adds country-by-year fixed effects. Column 5 includes the lagged dependent variable, and columns 6 and 7 weight each industry observation by its share of within-country value-added (VA; countries themselves are not weighted). Column 8 uses 10-year periods. Although columns 1–8 exploit variation in input shares across both countries and industries, column 9 focuses on within-country, cross-industry input-output (IO) linkages. See text for details. Standard errors are clustered at the industry level.

dynamics (col. 5). It is weaker but still present when we use 1992 within-country (nominal) value-added weights instead of our baseline unweighted specification (cols. 6 and 7, with and without controlling for the lagged dependent variable).[29] It is equally large, and (in this case) statistically significant, when we focus on a 10-year panel rather than stacked 5-year changes in column 8. In column 9, we show that the estimates are similar when we only use each country's domestic IO network, rather than the full international IO table (which incorporates inputs from each country-industry pair).[30] Finally, in table A9 (http://www.nber.org/data-appendix/c14854/appendix.pdf), we report cross-country regressions that include both own-country (domestic) values and foreign-country average values of the mean and variance of upstream (supplier) TFP growth as explanatory variables for sectoral productivity growth. These models show that own-country supplier TFP values are a far more robust predictor of sectoral productivity growth than the corresponding other-country values. This is especially the case for the variance term, where the own-country coefficient is negative and significant in all columns, whereas the other-country measure is neither significant nor consistently signed. This pattern is reassuring against the concern that our upstream TFP variance

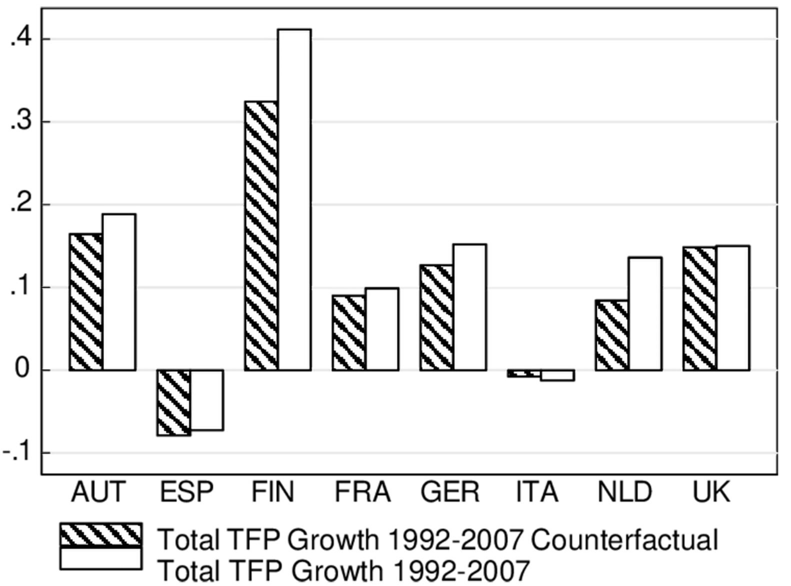

Fig. 6. Magnitude of bottleneck estimates in international data. This figure reports actual and counterfactual total factor productivity (TFP) growth between 1992 and 2007 across the countries in our international panel data (Austria, Spain, Finland, France, Germany, Italy, the Netherlands, and the United Kingdom). The counterfactuals are based on regression estimates from the column 2 specification of table 13. Specifically, counterfactual TFP (white bars) is computed from the regression coefficients as the TFP growth that would have been observed in the given country and year if the variance of supplier TFP growth had remained at the same level as during the initial period (1992–97). This is calculated by subtracting the contribution of supplier TFP variance from the actual TFP growth (patterned). A color version of this figure is available online.

Table 14
Evidence on Bottlenecks from Cross-Country Regressions Using Patenting

	Log Patent				DHS Specification			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Citation average	1.231	1.240	1.229	1.404	1.103	1.096	1.077	1.200
	(.021)	(.022)	(.022)	(.054)	(.013)	(.013)	(.013)	(.033)
Citation variance		−.031	−.030	−.025		−.048	−.034	−.056
		(.010)	(.011)	(.018)		(.020)	(.020)	(.031)
Year-by-CPC Fes	no	no	yes	yes	no	no	yes	yes
Year-by-country FEs	no	no	no	yes	no	no	no	yes
Observations	84,870	84,870	84,862	84,862	85,698	85,698	85,694	85,694
R^2	.223	.223	.278	.281	.241	.241	.306	.309

Note: Standard errors are clustered at the Cooperative Patent Classification (CPC) level. We consider the 20 countries with the greatest number of patents in the sample period (Austria, Australia, Belgium, Canada, Switzerland, China, Germany, Denmark, Spain, Finland, France, the United Kingdom, Italy, Japan, Korea, the Netherlands, Russia, Sweden, Taiwan, and the United States). Fixed effects (FE) for year and country are included in all regressions, and year-by-CPC or year-by-country fixed effects are included where indicated. Observations are unweighted. All regressions include changes across 5-year averages 1975–2014. Cols. 1–4 specify patent growth using the change in the log number of patents. Columns 5–8 specify patent growth as the change in patenting activity, normalized using the Davis-Haltiwanger-Schuh (DHS) transformation: $(P_{it} - P_{t-1}/\frac{1}{2}(P_{it} + P_{t-1}))$.

terms may be misspecified because they do not include productivity growth among important intermediates (see endnote 21).

These cross-country models also enable us to investigate whether our mechanism can account for the international slowdown in productivity growth. Figure 6, which is analogous to figure 5 for the United States, reports the results of this exercise. Across the European countries in our sample, we estimate that the rising variance of supplier TFP reduced aggregate productivity growth in eight of nine countries—all except Italy. This bottleneck effect is largest in Finland and the Netherlands, where we estimate that it reduced aggregate TFP growth 1992–2007 by 30% and 60%, respectively.

We also implement a similar specification where the outcome is patenting among firms located in different countries. We use the same CPC-citation linkages for each country (calculated using all patents within the USPTO database), but we apply a variant of equation (6) for patenting growth of firms located in the 20 largest countries within Google global patent data. Table 14 reports these cross-country estimates (which are again unweighted). Columns 1 through 4 use log patents, and columns 5 through 8 use the DHS transformation. Exploiting the cross-country variation, we see in columns 2 and 6 that, with either specification, there is a negative effect of upstream imbalances on patenting. This remains true in columns 3 and 7, where we include year-by-CPC fixed effects, thus identifying the relationship exclusively from cross-country variation in the upstream variance. The relationship is also broadly robust, though a little less precisely estimated, in columns 4 and 8, where we further include year-by-country fixed effects, thus focusing entirely on within-country variation.

VII. Conclusion

Despite the exponential pace of innovation in the ICT and electronics sectors, aggregate productivity growth in the United States and many other industrialized nations has been disappointing since the 1970s—and only more so since the early 2000s. Some have interpreted this pattern, variously, as reflecting a severe underestimation of quality and actual productivity growth, a temporary lull that proceeds a major surge in productivity, or an exhaustion of the potential supply of truly transformative innovations—leading to a long-term deceleration of productivity growth.

We proposed an alternative hypothesis that implies neither a permanent slowdown in productivity growth nor an incipient surge. We then investigated this new hypothesis empirically. The foundational idea of our approach is that innovation in any one industry relies on complementary innovations in—and subsequent productivity gains from—its input and idea suppliers. When innovation is unbalanced across industries, this holds back aggregate productivity growth by creating innovation "bottlenecks" along the IO or patent citation (idea) networks.

After presenting a simple version of this productivity bottleneck hypothesis, we explored it using data on IO linkages, citation linkages, patenting, and TFP growth. Across a variety of measurement approaches, productivity outcomes, and countries, we verify the primary prediction of this hypothesis: an industry's productivity growth is augmented by the mean productivity growth of its suppliers (measured by TFP or innovation) and, crucially, it is hampered by the variance of their productivity growth.

Our primary evidence exploits IO linkages and TFP growth to document the sensitivity of industry productivity growth to the mean and variance of supplier productivity growth. We supplemented this evidence by looking at patenting as a direct measure of innovation. This analysis suggests that there is a similarly powerful linkage between the innovativeness of a sector or firm and the imbalances it faces across its upstream (idea-supplier) sectors. For these results, we measured the upstream sectors based on industry- or firm-level citation networks.

At face value, our evidence implies that the bulk of the productivity slowdown in the United States (and several other industrialized economies) can be explained by the sizable increase in the cross-industry variance of TFP growth and innovation. For example, if TFP growth variance had remained at its 1977–87 level for the subsequent 2 decades, US manufacturing productivity would have grown twice as rapidly in 1997–2007 as it did—yielding a counterfactual growth rate that would have been close to its observed level in either of the 2 prior decades. These estimates illustrate the potential importance of our mechanism, but given the limitations of our measurement and sources of variation, they do not constitute a definitive assessment of its quantitative contribution.

We view our paper as a first step in the theoretical and empirical investigation of the interlinked nature of innovation across sectors. Based on the earlier findings, many areas of research appear fruitful. First, our hypothesis raises a critical theoretical question: Will the endogenous direction of technological progress tend to clear productivity bottlenecks, or might the market mechanism exacerbate imbalances? Second, this initial evidence highlights the need for additional empirical strategies to explore dependencies among innovating sectors and the innovations generated by their suppliers. These same relationships could be tested, for example, using firm-level IO data, where we suspect that the importance of supplier-customer linkages would be even larger. Third, another interesting context to explore is the role of global supply chains in productivity bottlenecks. On the one hand, imported intermediates and technologies can relax domestic bottlenecks. On the other hand, global supply chains may introduce more extensive technological dependencies, which could intensify bottlenecks if those trade channels become constrained. Fourth, it would be valuable to investigate the bottleneck hypothesis using historical data—focusing, for example, on major technological breakthroughs in the first half of the twentieth century. Finally, our framework makes a strong—perhaps even rash—prediction, whose verification awaits the passage of time: if and when lagging industries ultimately increase their innovation and productivity growth rates, a rapid takeoff in aggregate productivity should ensue.

Endnotes

Author email addresses: Acemoglu (daron@mit.edu), Autor (dautor@mit.edu), Patterson (christina.patterson@chicagobooth.edu). We thank Nick Bloom, Steve Davis, John

Fernald, Hugo Hopenhayn, Bill Kerr, Jennifer La'O, Valerie Ramey, and John van Reenen for their helpful comments; Rebecca Jackson, Zhe Fredric Kong, and Austin Lentsch for expert research assistance; and Ernest Liu and Song Ma for sharing international patenting data. Acemoglu and Autor acknowledge support from the Hewlett Foundation, Google, and the Smith Richardson Foundation. Acemoglu also acknowledges support from the National Science Foundation, the Sloan Foundation, and the Toulouse Network on Information Technology. Autor also acknowledges support from the Washington Center for Equitable Growth. For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see https://www.nber.org/books-and-chapters/nber-macroeconomics-annual-2023-volume-38/bottlenecks-sectoral-imbalances-and-us-productivity-slowdown.

1. The black bars correspond to the share of US Patent and Trademark Office (USPTO) patents granted in Electricity and Electronics (i.e., "electronics"), and the gray bars plot the share of patents granted in Instruments and Information (i.e., "ICT"). The black line shows the total number of patents granted.

2. Those who subscribe to the first view often highlight that growth is mismeasured, which is undoubtedly true. Nevertheless, mismeasurement does not seem to account for the broad outlines of the productivity slowdown since the 1970s. First, growth was almost surely mismeasured in the decades that followed World War II, when many new consumer goods and technologies were introduced. Second, many implications of the growth mismeasurement thesis, such as faster productivity growth in sectors with less potential for mismeasurement, do not receive support from the data (Byrne, Fernald, and Reinsdorf 2016; Syverson 2017). Third, there is no evidence for even the most basic predictions of fast, ICT-driven productivity growth; for example, industries with more-intensive use of ICT (outside of the ICT-producing industries themselves) have exhibited, if anything, slower growth of nominal and real value-added (Acemoglu et al. 2014).

3. See https://www.nobelprize.org/prizes/chemistry/2019/popular-information.

4. See Atalay (2017) and Baqaee and Farhi (2019) for such Neoclassical effects, which arise once we depart from unitary elasticities in production.

5. Specifically, using the Cooperative Patent Classification (CPC) scheme, we look at the mean and variance of patenting at the "upstream" patent classes. Upstream classes are constructed according to the citation network, which follows the approach in Acemoglu et al. (2016). We do not mix the patenting and TFP analyses, both because the idea network based on citations and the IO network are different and because the link between patents and productivity in our sample is modest, which may be due to the imperfect correspondence between industry classifications and patent technology classes.

6. The most sophisticated version of the "running out of ideas" hypothesis is developed in Bloom et al. (2020), who argue that innovations have become difficult in many fields but the rate of innovation has not declined commensurately because the amount of effort devoted to invention and innovation has increased.

7. $B_i = ((1 - \Sigma_{j\in S_i}\alpha_{ij})^{1-\Sigma_{j\in S_i}\alpha_{ij}}\prod_{j\in S_i}\alpha_{ij}^{\alpha_{ij}})^{-1}$. See Acemoglu and Azar (2020) for more details on this functional form.

8. It is straightforward to allow these sets to be time-varying, but we do not do so, to reduce notation. In our empirical work, we explore models both with and without time-varying input sets.

9. We have equated the importance of an input to its share in production, α_{ij}. This is not necessary for any of our main arguments, but it is the benchmark functional form assumption that we use in our empirical work. We also consider an alternative where the importance of an input innovation is measured by the number of citations to the innovation by patents from other industries.

10. The inputs that need to make technological advances before sector i can successfully innovate may be a subset of the inputs in S_i. Because we do not have a way to empirically determine which subset of inputs is important for innovation, we assume that all inputs in S_i are relevant, then verify robustness using other measures of industry linkages.

11. Even when $\eta^i_{\text{variance}} < 0$, an increase in the productivity of an input-supplier industry is always beneficial (and thus, the negative effect through the variance is weaker than the positive impact through the mean, η^i_{mean}) because the functions h and H are monotone.

12. As we discuss in the next subsection, it may not be possible to reduce the dispersion of technological progress across sectors without affecting the mean. In particular, such a mean-preserving dispersion reduction would require that the cost of improving technology in every sector is the same.

13. Finding the general equilibrium will also require us to solve for the wage rate (and the allocation of labor across sectors) and the interest rate (as a function of the aggregate growth rate of the economy). We do not derive these (standard) aspects of the general equilibrium.

14. See https://www.nber.org/research/data/nber-ces-manufacturing-industry-database.

15. The BLS also produces similar statistics for aggregated three-digit NAICS manufacturing industries. Although we do not use these BLS measures in our analysis, these statistics are highly correlated with the multifactor productivity measures for manufacturing in the Census data.

16. If there is more than one CPC code provided for a patent, we use the first-reported (i.e., primary) code.

17. In our baseline specification, this share includes inputs from all other countries. We explore alternative definitions and, in table A9 (http://www.nber.org/data-appendix/c14854/appendix.pdf), show that our results are mostly driven by TFP growth patterns among a country's domestically sourced inputs, with a more limited role for imported intermediates.

18. Like the fastest-growing industries, the bottom 50% industries are defined in terms of their contribution to supplier TFP variance between 1997 and 2007. Table A2 (http://www.nber.org/data-appendix/c14854/appendix.pdf) reports the full set of industries corresponding to each panel of table 2.

19. More specifically, we calculate these instruments using the US-based IO table, but taking the TFP growth across industries from each of three European countries (France, Germany, and the United Kingdom). Because the international industry data are more aggregated than our underlying NAICS data, six-digit NAICS codes are mapped to the most similar international industry code available, and the TFP growth value observed in the European instrument panel is assigned to US industries based on these mappings. To reflect this, we cluster the standard errors at the level of the 29 KLEMS industries in table 3. Throughout this exercise, we focus on our main, unweighted specifications, corresponding to cols. 2 and 3 of table 1.

20. Specifically, we calculate the covariance between the TFP growth of suppliers in the previous 5-year period ($t - 10$ to $t - 5$) and the current period ($t - 5$ to t), weighting each supplying industry by their input share in $t - 10$.

21. An additional concern is whether our variance term is misspecified, because it does not account for the productivity of offshored and imported inputs. Here we note that this concern would create attenuation toward zero and hence is unlikely to account for our findings. It also does not apply when we turn to patenting, because our analysis there will include foreign patents as well. We further discuss this issue in Sec. VI.

22. We are grateful to our discussant John Fernald for pointing out this problem and proposing the adjustment we implement here.

23. We use employment rather than output, because output numbers would be directly correlated with TFP estimates. We also focus on unweighted specifications, as in our other robustness explorations.

24. Because the mappings between CPCs and Standard Industrial Classification/NAICS classifications are imperfect, we do not explore the relationship between upstream patenting and downstream productivity growth.

25. The DHS transformation for a variable X is $X_t - X_{t-1}/(1/2(X_t + X_{t-1}))$.

26. In particular, our main TFP estimates from col. 2 of table 1 suggest that a one standard deviation increase in upstream TFP growth variance is associated with TFP growth that is 0.035 percentage points lower.

27. Recall from endnote 25 that the DHS transformation is $X_t - X_{t-1}/(1/2(X_t + X_{t-1}))$. This is undefined when both X_t and X_{t-1} are equal to zero. In the modified DHS, rather than dropping such observations, we set $0/0 = 0$.

28. Specifically, we use the world IO tables to calculate the input share $\alpha_{ik,jl}$ as the share of inputs from industry i in country k that come from industry j in country l. The shares are based only on the nine countries listed earlier.

29. We do not have nominal sales data in our international panel and hence use nominal value-added weights rather than nominal sales weights. All weights are within-country, meaning that they are relative to total GDP of the country.

30. We do not report estimates using the manufacturing sample in this case, both because manufacturing industries are not sufficiently disaggregated in this data set and because doing so would reduce our sample by about two-thirds. Finally, table A10 (http://www.nber.org/data-appendix/c14854/appendix.pdf) shows the robustness of our US results, aggregated to the 30 industries used in table 13.

References

Acemoglu, Daron, Ufuk Akcigit, and William R. Kerr. 2016. "Innovation Network." *Proceedings of the National Academy of Sciences* 113 (41): 11483–88.

Acemoglu, Daron, David Autor, David Dorn, Gordon H. Hanson, and Brendan Price. 2014. "Return of the Solow Paradox? IT, Productivity, and Employment in US Manufacturing." *AER Papers and Proceedings* 104 (5): 394–99.

Acemoglu, Daron, and Pablo D. Azar. 2020. "Endogenous Production Networks." *Econometrica* 88 (1): 33–82.

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. 2012. "The Network Origins of Aggregate Fluctuations." *Econometrica* 80 (5): 1977–2016.

Acemoglu, Daron, and Pascual Restrepo. 2019. "Automation and New Tasks: How Technology Displaces and Reinstates Labor." *Journal of Economic Perspectives* 33 (2): 3–30.

Aghion, Philippe, Ufuk Akcigit, Antonin Bergeaud, Richard Blundell, and David Hémous. 2019. "Innovation and Top Income Inequality." *Review of Economic Studies* 86 (1): 1–45.

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth through Creative Destruction." *Econometrica* 60 (2): 323–51.

Akcigit, Ufuk, and Sina T. Ates. 2019. "What Happened to US Business Dynamism?" Working Paper no. 25756, NBER, Cambridge, MA.

Andrews, Dan, Chiara Criscuolo, and Peter N. Gal. 2016. "The Best versus the Rest: The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy." https://www.oecd-ilibrary.org/economics/the-best-versus-the-rest_63629cc9-en.

Atalay, Enghin. 2017. "How Important Are Sectoral Shocks?" *American Economic Journal: Macroeconomics* 9 (4): 254–80.

Autor, David, David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review* 103 (6): 2121–68.

Autor, David, David Dorn, Gordon H. Hanson, Gary Pisano, and Pian Shu. 2020. "Foreign Competition and Domestic Innovation: Evidence from US Patents." *AER: Insights* 2 (3): 357–74.

Balsmeier, Benjamin, Mohamad Assaf, Tyler Chesebro, Gabe Fierro, Kevin Johnson, Scott Johnson, Guan-Cheng Li, et al. 2018. "Machine Learning and Natural Language Processing on the Patent Corpus: Data, Tools, and New Measures." *Journal of Economics and Management Strategy* 27 (3): 535–53. https://doi.org/10.1111/jems.12259.

Baqaee, David, and Emmanuel Farhi. 2019. "Networks, Barriers, and Trade." Working Paper no. 26108, NBER, Cambridge, MA.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb. 2020. "Are Ideas Getting Harder to Find?" *American Economic Review* 110 (4): 1104–44.

Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Cambridge: Oxford University Press.
Bresnahan, Timothy F., and Manuel Trajtenberg. 1995. "General Purpose Technologies 'Engines of Growth'?" *Journal of Econometrics* 65 (1): 83–108.
Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2021. "The Productivity J-Curve: How Intangibles Complement General Purpose Technologies." *American Economic Journal: Macroeconomics* 13 (1): 333–72.
Byrne, David M., John G. Fernald, and Marshall B. Reinsdorf. 2016. "Does the United States Have a Productivity Slowdown or a Measurement Problem?" *Brookings Papers on Economic Activity* 2016 (1): 109–82.
Cohen, Lauren, Umit G. Gurun, and Scott Duke Kominers. 2016. "The Growing Problem of Patent Trolling." *Science* 352 (6285): 521–22.
Cowen, Tyler. 2011. *The Great Stagnation: How America Ate All the Low-Hanging Fruit of Modern History, Got Sick, and Will (Eventually) Feel Better*. New York: Dutton.
David, Paul A. 1990. "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox." *American Economic Review* 80 (2): 355–61.
Davis, Steven J., John C. Haltiwanger, and Scott Schuh. 1998. *Job Creation and Destruction*. Cambridge, MA: MIT Press.
Decker, Ryan A., John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2017. "Declining Dynamism, Allocative Efficiency, and the Productivity Slowdown." *American Economic Review* 107 (5): 322–26.
Diamandis, Peter H., and Steven Kotler. 2012. *Abundance: The Future Is Better Than You Think*. New York: Simon and Schuster.
Gordon, Robert J. 2017. *The Rise and Fall of American Growth: The US Standard of Living Since the Civil War*. Princeton, NJ: Princeton University Press.
Grossman, Gene M., and Elhanan Helpman. 1991. *Innovation and Growth in the Global Economy*. Cambridge, MA: MIT Press.
Gruber, Jonathan, and Simon Johnson. 2019. *Jump-Starting America: How Breakthrough Science Can Revive Economic Growth and the American Dream*. London: Hachette UK.
Helpman, Elhanan, and Manuel Trajtenberg. 1996. "Diffusion of General Purpose Technologies." Working Paper no. 5773, NBER, Cambridge, MA.
Iancu, Andrei. 2019. "Remarks by Director Iancu." Presentation at the International Intellectual Property Conference, Seoul, June 10.
Kurzweil, Ray. 2005. *The Singularity Is Near: When Humans Transcend Biology*. London: Penguin.
Laws, David. 2018. "13 Sextillion & Counting: The Long & Winding Road to the Most Frequently Manufactured Human Artifact in History." https://computerhistory.org/blog/13-sextillion-counting-the-long-winding-road-to-the-most-frequently-manufactured-human-artifact-in-history/.
Liu, Ernest, and Song Ma. 2021. "Innovation Networks and R&D Allocation." Working Paper no. 29607, NBER, Cambridge, MA.
Long, John B., Jr., and Charles I. Plosser. 1983. "Real Business Cycles." *Journal of Political Economy* 91 (1): 39–69.
Park, John N., Robert L. Steigerwald, and Loren H. Walker. 1976. Rectifier Circuits Using Transistors as Rectifying Elements. US Patent 3,940,682, filed November 22, 1974, and issued February 24, 1976.
Sosa, Estanislao N. 2013. "Electron Tube." Encyclopedia Britannica. https://www.britannica.com/technology/electron-tube.
Syverson, Chad. 2017. "Challenges to Mismeasurement Explanations for the US Productivity Slowdown." *Journal of Economic Perspectives* 31 (2): 165–86.