NBER WORKING PAPER SERIES

SCIENCE AND THE MARKET FOR TECHNOLOGY

Ashish Arora
Sharon Belenzon
Jungkyu Suh

Working Paper 28534
http://www.nber.org/papers/w28534

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2021

We would like to thank Victor Bennett, Aaron Chatterji, Wesley Cohen, Alfonso Gambardella, Sharique Hasan, Mark Schankerman, Carlos Serrano and seminar participants at the University of California San Diego and Hebrew University of Jerusalem for their helpful comments. All remaining errors are ours. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2021 by Ashish Arora, Sharon Belenzon, and Jungkyu Suh. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Science and the Market for Technology
Ashish Arora, Sharon Belenzon, and Jungkyu Suh
NBER Working Paper No. 28534
March 2021
JEL No. L24,O3,O31,O33,O34,O36

ABSTRACT

Well functioning Markets for Technology (MFT) allow inventors to sell their inventions to others that may derive more value from them. We argue that the growing reliance on science in inventions enhances MFT. In addition to higher quality inventions, reliance on science may enhance gains from trade and reduce the transfer cost of knowledge and other transaction costs. Using large scale data, we show that patents citing science are more likely to be traded, especially for novel patents and for smaller inventors. Leveraging the fall of the Berlin Wall as a source of exogenous variation in the relevant scientific knowledge to technological fields, we confirm reliance on science increases the likelihood that the invention will be traded

Ashish Arora
Fuqua School of Business
Duke University
Box 90120
Durham, NC 27708-0120
and NBER
ashish.arora@duke.edu

Sharon Belenzon
Fuqua School of Business
Duke University
100 Fuqua Drive
Durham, NC 27708
and NBER
sharon.belenzon@duke.edu

Jungkyu Suh
Duke University
Fuqua School of Business
durham, nc 27708
jungkyu.suh@duke.edu

1 Introduction

A well-functioning Market for Technology (MFT) enhances welfare by allowing inventors to sell or license inventions to innovators who may commercialize them more efficiently. The past several decades have witnessed a rise of a Market for Technology (MFT). Estimates based on Graham et al. (2018) show that patent reassignments have risen tenfold from around 2,000 to over 20,000 cases between 1980 and 2016. As well, U.S. corporations have reported a steady increase in royalty receipts and payments for industrial processes abroad, from \$1.5 billion and \$0.4 billion respectively in 1987 to \$12.8 billion and \$4.5 billion in 2017.[1] University licensing revenues have increased tenfold over an even shorter period, from \$218 million in 1991 to \$2.5 billion in 2015 (AUTM, 2015). At the same time, the reliance of inventions on science has increased, as indicated by the rise in the share of patents citing science from 4% to 28% of all U.S. utility patents between 1980 and 2015 (Marx, 2019).[2] In this paper, we relate these two phenomena theoretically and empirically. Specifically, we investigate whether reliance on science increases the likelihood that the resulting invention will be traded by the inventor, and explore the possible channels for the relationship.

We define science as knowledge that is codified in the scientific and engineering literature. Inventions can be said to rely on science if they arise from scientific discoveries, if they draw upon scientific knowledge in important ways, or if they are conceptualized in scientific terms. Reliance on science may result in higher quality inventions, but may also increase the likelihood of trade by reducing transfer costs and increasing gains from trade. Science elucidates and codifies from empirical regularities the mechanisms underlying natural phenomena (Arora and Gambardella, 1994; Mokyr, 2002). Inventions conceptualized in scientific terms may be codified more effectively and their "metes and bounds" delineated more crisply, thereby reducing contracting frictions with potential buyers. Furthermore, codification in scientific terms may facilitate the discovery of new applications, more valuable than those contemplated by the original inventor, thereby raising potential gains from trade.

We develop a simple analytical framework which incorporates three mechanisms through which reliance on science could affect the probability of trade (MFT): invention quality, transaction costs, and

[1] Excludes receipts and payments from affiliates. Data for 2017 from BEA website (https://apps.bea.gov/iTable/iTable.cfm?reqid=62step=9isuri=1&product=4); data for 1987 from the scanned issues of the Survey of Current Business. 1921-2014. https://fraser.stlouisfed.org/title/46, accessed on March 11, 2019.

[2] This rise has coincided with an increase in the amount of scientific research produced each year. In 2016, 32,246 "hard science" doctorates were awarded in the United States, which is more than twice the number in 1986 (13,914) (Thurgood et al., 2006)."Hard science" includes Science and Engineering, excluding Social Sciences, Education, Humanities and Arts. Globally, the publication of peer-reviewed scientific articles has grown at an accelerating rate, with annual growth rates of 1.8% in the 1980s rising to around 4.01% in the 1990s and 3.99% in the 2000s. In aggregate, 1.7 million articles were published in 2016, compared to just over 500 thousand in 1980. (Authors' calculations based on Clarivate Web of Science.)

gains from trade. In turn, these have separate implications for how the relationship between reliance on science and probability of trade differ, by inventor size and by invention novelty. The market level equilibrium with entry into invention shows that an increase in demand in MFT encourages the entry of smaller inventors. Insofar as reliance on science makes it easier to trade inventions, the entrants are more likely to rely on science for their inventions and to trade their inventions. This also highlights the empirical challenges in estimating the effect of reliance on science on MFT. Specifically, insofar as greater demand attracts specialized research organizations (e.g., universities and small firms) to supply inventions for trade, and insofar as we do not measure the commercialization capability of the inventor properly, this would lead to a positive association between the observed reliance on science in invention and the propensity to trade. More generally, unobserved differences in demand conditions in MFT may bias the estimated relationship away from the true one.

We mitigate this concern in several ways. First, at the patent level, we control for size using a variety of measures, and show that not only does the relationship continue to hold but that MFT-science relationship is stronger for small inventors, as predicted by our model. Second, we show that the MFT-science relationship is systematically stronger for novel inventions, making it unlikely that unobserved heterogeneity is the only cause of the relationship. Third, we aggregate up to the IPC level to study how the proportion of science-based patents is related to the proportion of patents that are traded. Consistent with our analytical framework, reliance on science is associated with entry of inventors, and a greater share of specialised inventors. However, the principal payoff from the IPC level analysis is that we can exploit changes in U.S. government funding of research following the collapse of the Soviet Union as a source of exogenous variation in the reliance on science. The identifying assumption is that the reliance on science is greater when the supply of relevant science increases for reasons unrelated to trade. We instrument for the share of inventions relying on science using changes in government funding to scientific fields relevant to the focal patent. We show that scientific output rises in response to government funding, which in turn increases the share of science-based patents, and consequently, a higher rate of trade.

We use data on U.S. patents and scientific publications between 1980 and 2016. We measure reliance on science by whether a patent cites a scientific article (Marx and Fuegi, 2020). This dataset matches front page Non Patent Literature (NPL) citations in U.S. patents to peer-reviewed scientific publications from Microsoft Academic Graph (MAG). We validate this as a measure of reliance on science in a variety of ways, and show further that our results are robust to using a textual similarity measure. We use patent reassignments from the USPTO Patent Assignment Database (PAD) to measure transactions in MFT.

We present three main findings. First, a patent that cites a scientific publication has a 23% higher probability of being traded compared to a patent that does not cite science. After we control for quality using number of claims, triadic patenting status, (and stock market value for a subset of patents matched to listed firms), the magnitude of the science-MFT coefficient is reduced by a third, but still positive and statistically significant. We interpret this as suggesting that reliance on science is related to MFT in other ways, not just through quality.

Second, consistent with the view that science increases gains from trade, we find that the relationship between science and MFT is four times larger for small firms compared to large firms. Since larger firms already have high commercialization capabilities (and hence derive lower value from selling), reliance on science will have a smaller effect on the gains they can reap from selling their invention compared to smaller firms. We also find that the science-MFT relationship is about a third larger for patents using a one standard deviation "newer" combination of technological subclasses (Fleming, 2001). Unfamiliar inventions are likely to be more difficult for buyers to evaluate, but when such inventions are based in science, buyers may gain a better understanding of otherwise unfamiliar technological components. We also find that reliance upon more recent science is more strongly associated with patent trade. This suggests that the principal channel through which science affects technology markets is not by reducing transfer costs, since mature science is more likely to provide a "common language" or background knowledge for market participants (Arora and Gambardella, 1994). Newer science, and more specialized science, on the other hand, is arguably more important in illuminating uncharted applications of inventions, thereby increasing gains from trade.

Third, we present instrumental variable results that are based on an exogenous change in the availability of science, which affects the prevalence of science-based inventions unrelated to trade conditions. After the fall of the Soviet Union, federal funding for research fell by almost 40% for Space and Defense between 1986 and 1992, whereas it doubled in Medical and Energy. These changes led to substantial variation in the supply of new knowledge, and therefore, variation in the reliance on science in different fields of invention. The geopolitical circumstances that precipitated the end of the Cold War are exogenous to MFT. We find that patent classes with a standard deviation increase in the share of science-based patents experience a 6.5% increase in the share of reassigned patents. Our IV estimates are 13% smaller than our OLS estimates, but still positive and statistically significant.

Our main contribution is to establish that reliance on science enhances technology trade. We build on the literature on MFT (Ziedonis, 2004; Serrano, 2010; Serrano et al., 2015; Serrano, 2018; Ma et al., 2017; Arque-Castells and Spulber, 2017; Arora and Fosfuri, 2000; Gans et al., 2002). Gans et al. (2008) and Galasso and Schankerman (2018) identify the effect of a reduction in transaction costs from the resolution of uncertainty around patent property rights (from USPTO allowance events and court decisions respectively) on MFT. With the exception of Marco et al. (2017), the relationship between science and MFT has not been explored. We use exogenous variation in the availability of science to provide the first estimate of the causal impact of reliance on science on MFT.

We also contribute to the literature that aims at quantifying the social returns to science (Griliches, 1957; Mansfield, 1980; Arora et al., 2020). Our results suggest that science contributes to social welfare through a separate channel – the efficient allocation of inventions, and indirectly therefore, by promoting a division of innovative labor.[3] Jones (2009) argues that the growing "burden of knowledge" implies increasing individual specialization and greater need for cooperation. The market for technology is an important mechanism for facilitating such cooperation. To the best of our knowledge, this mechanism of the social returns to science has not been previously systematically explored empirically.[4]

A secondary contribution is to elucidate the different channels through which reliance of science is related to the market for technology. Inventions based on scientific discoveries may be higher in quality. However, there are other channels as well. Our empirical results suggest that the reliance on science in invention is also associated with lower transaction costs and higher gains from trade. By clarifying the underlying mechanisms of unfamiliar inventions, and using standard terms, the relevance of the patent may be more apparent to a wider set of buyers.[5] Put differently, scientific inventions offer greater gains from trade. Realizing these gains from trade must contend with transaction costs, including the costs of transferring tacit know-how as well as opportunism by buyers or sellers. A fundamental problem of selling knowledge is that the bargaining process requires inevitably disclosing the "secret sauce" to the buyer (Arrow, 1962; Anton and Yao, 1994). A mirror problem for the buyer is that inventions often require complementary tacit know-how to exploit, and therefore require the active cooperation

[3] The evolution of Light Emitting Diode (LED) technology illustrates this point. Semiconductors that emit light were discovered as early as 1907, when Henry Round, a British radio engineer, observed a light yellow light emitting from his silicon carbide-based detector. However, the mechanisms behind this observation required a better understanding of quantum theory before the phenomenon could be applied more broadly. Therefore, early LED inventions were done in vertically integrated firms such as TI (infrared LED in 1961) and GE (red LED in the same year) (Sethi, 2013; Stevenson, 2009). Deeper scientific understing allowed specialized firms such as Universal Display Corporation (UDC) to enter by licensing their intellectual property on dopants to incumbents. See for instance UDC and BASF's patent deal in IMSExpert. "$96M in OLED Patents, "Fruitful" Purchase for 2017". National Law Review. August 12, 2016 Friday. https://advance-lexis-com.proxy.lib.duke.edu/api/document?collection=newsid=urn:contentItem:5KFJ-DRC1-F03R-N0XF-00000-00context=1516831.

[4] Our paper is also related to the growing literature that examines the use of science in inventions using patent citations to science (Narin et al., 1997; Azoulay et al., 2015; Fleming et al., 2019; Agrawal and Henderson, 2002; Belenzon and Schankerman, 2009; Fleming and Sorenson, 2004; Ahmadpoor and Jones, 2017; Veugelers and Wang, 2019)

[5] The Bessemer process, for instance, diffused rapidly after its metallurgical properties were sufficiently understood.

of the seller (Arora, 1995; Von Hippel, 1994; Polanyi, 2015; Kogut and Zander, 1992). Science can help codify a greater share of this tacit complementary knowledge and reduce the risk of hold-up and bargaining breakdowns (Galasso and Schankerman, 2014; Merges and Nelson, 1990). Scientific patents are "crisper", which helps potential buyers better understand what they are buying.[6] In sum, reliance on science reduces transaction costs because scientific inventions are easier to codify, and cheaper to transfer, and less vulnerable to contracting failures.

2 A model of trade in patents

We present a simple model to clarify the relationship between science and MFT. There are I inventors. Each inventor is endowed with an invention, whose use of science is indexed by s. In subsection 2.4 we consider entry into invention but for now, I and s are given. The inventor can commercialize the invention herself and earn $y_i = q_i(x_i + \epsilon_i)$, where q_i is the quality of the i^{th} invention, and x_i is the expected value the inventor can extract, and ϵ_i represents the idiosyncratic component of value.

For the i^{th} invention, there are N_i firms that may buy it to commercialize it themselves, earning a payoff y_{ik} (and zero otherwise), where $y_{ik} = q_i(x_{ik} + \epsilon_{ik}) - \tau_i$.[7] Here, τ_i represents transaction costs, and x_{ik} represents the systematic component of the value the k^{th} potential buyer can extract from invention i, and ϵ_{ik} is the idiosyncratic component. Transaction costs have several components. These include contracting costs, such as legal fees, and trading frictions due to imperfect contracting, as well as the cost of transferring accompanying know-how. The primary determinant of x_{ik} is commercialization capability of the buyer, reflecting how well the potential buyer commercialize the invention, although it may also reflect ability to enforce patents (Galasso et al., 2013).

We assume efficient bargaining: If trade with at least one of the N_i potential buyers offers a surplus that is at least as great as the transaction costs, the invention will be traded. This assumption sweeps away considerations of asymmetric information, or how the joint surplus is divided between the buyer and seller.

[6] A prime example is the chemical industry, where patents are more effective tools of protecting fruits of R&D thanks to their reliance on the explicit description of "formulae, reaction pathways and operating conditions" represented via Markush structures work better in protecting property rights (Arora and Fosfuri, 2000; Levin et al., 1987).

[7] We assume that the inventor and potential buyers are equidistant from each other in the product space, so that any rent dissipation is the same, regardless of who commercializes the invention. Hence the baseline payoffs can be normalized to zero for all.

2.1 Probability of trade

The invention is not traded if $y_i \geq \max_k^{N_i}\{y_{ik}\} \iff x_i + \epsilon_i - \max_k^{N_i}\{x_{ik} + \epsilon_{ik}\} \geq -\frac{\tau_i}{q_i}$. We assume that ϵ_i and ϵ_{ik} are distributed *iid* as Type I extreme value distribution (Gumbell distribution).[8] The probability the invention is not traded, P = Pr(No Trade)

$$P(\text{No Trade}) = \frac{\exp(x_i + \frac{\tau_i}{q_i})}{\exp(x_i + \frac{\tau_i}{q_i}) + \sum_k^{N_i} \exp(x_{ik})} = \frac{1}{1 + \sum_k^{N_i} \exp(x_{ik} - x_i - \frac{\tau_i}{q_i})} \tag{1}$$

Equation 1 illustrates, as discussed earlier, that reliance on science can increase trade through gains from trade, $x_{ik} - x_i$, and through a reduction in transaction costs relative to the quality of the invention, $\frac{\tau_i}{q_i}$.[9] The combined effect is to increase the probability of trade. To minimise clutter, we feature the transaction cost channel in the algebra below. Formally,

Result 0 Science based inventions are more likely to be traded.

2.2 Role of science

2.2.1 Inventor capabilities

Recall that x_i represents the ability of the inventor to derive value from her own invention. If we call "small" inventors as those with lower x_i, these inventors are more likely to sell.[10]

$$\begin{aligned} \frac{\partial P(\text{no trade})}{\partial x_i} &= \frac{\exp(-x_i - \frac{\tau_i}{q_i} + \bar{x}_i)N_i}{(\exp(-x_i - \frac{\tau_i}{q_i} + \bar{x}_i)N_i + 1)^2} = \frac{a}{(a+1)^2} \geq 0, \quad \text{where} \\ a &= N_i \exp(-x_i - \frac{\tau_i}{q_i} + \bar{x}_i) \end{aligned} \tag{2}$$

2.2.2 Inventor capabilities and science

$$\frac{\partial^2 P(\text{no trade})}{\partial x_i \partial s} = \frac{1-a}{(a+1)^3}\frac{\partial a}{\partial s} \quad \begin{cases} \leq 0 & \text{if } a > 1 \\ \geq 0 & \text{if } a < 1 \end{cases} \tag{3}$$

[8]We normalize the scale parameter to unity and the location parameter to zero. This normalization eases notation and does not affect the results.

[9]Through out we normalize transaction costs by invention quality. The gains from trade channel is explicitly featured when considering novel inventions.

[10]For ease of exposition, we show the proof with homogeneous buyers, where $x_{ik} = \bar{x}_i \forall k$ where $\bar{x}_i$ is the value that the buyer can derive from the invention, the probability of No Trade can be rewritten as $\frac{1}{N_i \exp(\bar{x}_i - x_i - \frac{\tau_i}{q_i}) + 1} = \frac{1}{a+1}$.

The probability of no trade, $P(\text{no trade }) = \dfrac{1}{1+a} \approx 0.9$ in the sample.[11] Therefore, $a \approx \frac{1}{0.9} - 1 \leq 1$ so that the science will accentuate the increase in the probability of trade as firm size falls. Intuitively, smaller firms have greater net surplus from trade $\bar{x}_i - x_i - \tau$. Their probability of trade will be more responsive to a decrease in τ or an increase in $\bar{x}_i$, as long as the *pdf* of ϵ is monotonically increasing. This will typically be the case when the probability of trade is low. Formally,

Result 1 Smaller firms are more likely to trade their inventions than larger firms, and the difference increases for science-based inventions.

2.3 Novelty of invention

Novel patents are likely to have greater variation across buyers in their ability to extract value. We show greater heterogeneity among buyers increases trade, and this effect is stronger for science based novel inventions. Novelty may also affect transaction cost, which we analyze later.

2.3.1 Buyer heterogeneity

We analyze how the no trade condition in equation 1 is affected by an increase in the heterogeneity of buyers. We compare two cases, one where all external buyers have the same $x_{ik} = \bar{x}_i$, and the other where external buyers vary but $\mathrm{E}_k[x_{ik}] = \bar{x}_i$, and we set $q_i = 1$ to ease notation.

$$P(\text{No Trade}|\text{heterog}) = \frac{1}{1+b} < P(\text{No trade}|\text{homog}) = \frac{1}{1+a},\ b \ = \sum_k^{N_i} \exp(x_{ik} - \tau - x_i) \tag{4}$$

where the inequality follows from recognizing that the exponential is a convex function and applying Jensen's Inequality, so that $b > a$. Further, greater heterogeneity among buyers in terms of commercialization capability, x_{ik}, reduces the probability of No Trade.

$$\left.\frac{\partial P(\text{No Trade})}{\partial s}\right|_{heterog} = \frac{\partial \tau}{\partial s}\frac{b}{(b+1)^2},\ \left.\frac{\partial P(\text{No Trade})}{\partial s}\right|_{homog} = \frac{\partial \tau}{\partial s}\frac{a}{(a+1)^2} \tag{5}$$

Because $1 > b > a$, the probability of trade increases faster with reliance on science when buyers are

[11] Note that $\dfrac{\partial a}{\partial s} = -\dfrac{\partial \tau}{\partial s}\dfrac{1}{q_i}\exp(-x_i - \dfrac{\tau_i}{q_i} + \bar{x}_i) = -\dfrac{\partial \tau}{\partial s}\dfrac{1}{q_i}a \geq 0$.

heterogeneous. That is

$$\left.\frac{\partial P(\text{Trade})}{\partial s}\right|_{heterog} - \left.\frac{\partial P(\text{Trade})}{\partial s}\right|_{homog} = \frac{\partial \tau}{\partial s}\left(\frac{a}{(1+a)^2} - \frac{b}{(1+b)^2}\right) \geq 0 \tag{6}$$

Result 2 The probability of trade increases with science faster when buyers are heterogeneous.

The foregoing argument also has implications for the relationship between novelty and reliance on science. Insofar as novelty implies greater heterogeneity in valuation, novel inventions are more likely to be traded than other inventions, and reliance on science enhances the effect of novelty. However, it is likely that novel inventions involve greater transaction costs. For instance, novel inventions are more likely to require the transfer of tacit knowledge for successful implementation. Therefore, buyer heterogeneity and higher transaction costs cut in opposite directions, and the net effect on the level of trade is ambiguous. Nonetheless, if reliance on science is more effective in reducing transaction costs for novel inventions, then even if the effect of novelty is ambiguous, we show that reliance on science will enhance the likelihood of trade for novel inventions.

Science is likely to be more effective in reducing transaction costs for novel inventions. The scientific basis of such inventions will make their potential value more apparent to buyers, and it should also reduce the tacit knowledge required for successful transfer. It is therefore plausible that the fall in transaction cost due to science is greater for novel inventions than for incremental inventions. Formally, letting σ represent novelty of invention, we assume that $\frac{\partial^2 \tau}{\partial \sigma \partial s} \leq 0$. Consider two inventions, $i\&j$, and let $\sigma_i > \sigma_j$, so that i is more novel than j. The probability of no-trade (normalizing $q_i = q_j = 1$) is given by $\frac{1}{1+b_m}, m = i, j$, where $b_m = \sum_k^N \exp(-x_m + x_{mk} - \tau(\sigma_m, s)$. Evaluated at the point where the probability of trade is the same, so that $b_j = b_i = b$, we have

$$\frac{\partial P_i(\text{no trade})}{\partial s} - \frac{\partial P_j(\text{no trade})}{\partial s} = \frac{b}{(1+b)^2}\left(\frac{\partial \tau(\sigma_i, s)}{\partial s} - \frac{\partial \tau(\sigma_j, s)}{\partial s}\right) \leq 0 \tag{7}$$

Result 3 Reliance on science increases the probability of trade more for novel inventions than for incremental inventions.[12]

[12]Buyer heterogeneity is related to gains from trade. Result 3 shows that transaction cost reductions are more important when gains from trade are higher, pointing to the sense in which these are synergistic.

2.4 Equilibrium: Specialization and division of innovative labor

We now relax the assumption that inventions are exogenously assigned, and analyze entry into invention. The prospect of being able to trade an invention increases its value. We show that an increase in the number of potential buyers would result in a higher share of inventions that rely on science.

Suppose the inventor captures a fraction λ of the surplus if the invention is traded. The expected payoff from invention is $\Pi = \lambda \ln (\exp(x_i) + N(\exp(\bar{x} - \tau(s)) - R$, where R is the investment required to produce an invention.[13] The marginal inventor, who is indifferent between inventing and not inventing, is characterized by $x^*(s)$ such that

$$\Pi(x^*(s)) = \lambda \ln (\exp(x*) + N(\exp(\bar{x} - \tau(s))) - R = 0 \tag{8}$$

Equation 8 highlights the different ways in which reliance on science and MFT are related. First, as Adam Smith noted, trade encourage specialization: $\dfrac{\partial x^*}{\partial N} = -\dfrac{P(\text{trade})}{1 - P(\text{trade})} \leq 0$. Second, the equilibrium size of marginal inventor relying on science is smaller than of the marginal inventor not relying on science.

$$\frac{d\Pi(x^*)}{ds} = 0 \implies \frac{\partial x^*}{\partial s}\frac{\partial \tau}{\partial s} = \frac{P(\text{trade})}{1 - P(\text{trade})} \implies \frac{\partial x^*}{\partial s} \leq 0 \tag{9}$$

Third, reliance on science implies a direct increase in trade, as well indirectly because the marginal inventor relying on science is smaller, i.e., $\dfrac{dP(\text{no trade})}{ds} = \dfrac{a}{(1+a)^2}\left(\dfrac{\partial \tau}{\partial s} + \dfrac{\partial x*}{\partial s}\right) \leq 0$. The share of science-based inventions is correlated with the number of potential buyers, creating a potential source of bias. To see this, suppose that s is a binary variable, and the probability an invention uses science is p, which is the same for all inventions. Further, suppose inventor types, x, are distributed uniformly on the unit interval. The equilibrium entry condition (equation 8) implies that only inventions that satisfy $x > x^*(s)$ are realised. Let $x_1 = x^*(s = 1), x_0 = x^*(s = 0)$. Equation 9 implies that $x_1 < x_0$. The observed share of inventions, $\tilde{p}$, that use science is given by

$$\tilde{p} = \frac{p(1 - x_1)}{p(1 - x_1) + (1 - p)(1 - x_0))} > p \iff x_1 < x_0. \tag{10}$$

Equation 10 implies the observed share of science-based patents, $\tilde{p}$, is greater than p, the true share of

[13] For a proof, see Anderson et al. (1992), pages 60-61. For simplicity we assume that $\tau_i = \tau(s)$ so that transaction costs depend only the use of science but not on the identity of the inventor, and assume homogeneous buyers.

science-based patents. The share of traded patents in equilibrium is

$$\text{share traded } = \underbrace{\int_{x_0}^{1} \mathrm{P}(\text{trade}|x, s=0)dx}_{\alpha} + p \underbrace{\left(\int_{x_1}^{1} \mathrm{P}(\text{trade}|x, s=1)dx - \int_{x_0}^{1} \mathrm{P}(\text{trade}|x, s=0)dx\right)}_{\beta} \tag{11}$$

Equation 11 highlights the empirical challenge in estimating the structural relationship between the reliance on science and MFT, represented here by β, at the market level.[14] We observe $\tilde{p}$ rather than p. Unobserved factors, such as an increase in the number of potential buyers, N, will directly increase the probability of trade, as well as increase $\tilde{p}$. This will bias the OLS estimate upward. Patent level regressions may also not be free from bias if we imperfectly measure x_i, the commercialization capability of the inventor. Equation 9 implies that x_i is negatively related to the reliance on science, and also, by assumption, negatively related to the probability of trade.

To summarize, patent level estimations estimate the direct relationship (i.e., the reduction in transaction costs and increases in gains from trade), but not the indirect effects due to entry. Moreover, there is potential upward bias if commercialization capability is measured inadequately. At the market (IPC) level, estimates combine the direct and indirect (entry of specialized inventors), but there is potential upward bias because the observed reliance on science is measured with bias.

We present estimates at both the patent and the IPC level. In addition, we develop a source of exogenous variation in p at the IPC level to purge $\tilde{p}$ of the bias. Concretely, suppose that $p = p(K), p' > 0$ where K is the stock of relevant science. That is, we assume that the share of science based inventions is increasing in the stock of available science. As discussed below, we use changes in government support for science that are unrelated to conditions in the market for technology, as a source of exogenous variation in K, and therefore, in p. The identifying assumption is that these changes in the government funding ΔG, and the resulting changes in the stock of knowledge, ΔK, are orthogonal to N, the unobserved number of potential buyers in the MFT.

3 Data

We combine data on patents and peer-reviewed scientific publications. Our patent data is from the 2016 publication of PatStat and encompasses around 5.2 million utility patents granted by the USPTO

[14] The structural relationship represented by β has both a direct component (represented by $\mathrm{P}(\text{trade}|x, s=1) - \mathrm{P}(\text{trade}|x, s=0)$), and an indirect component, represented by the differences between x_1 and x_0, as seen in the lower limits of the integrals. The bias is related to the indirect component, because $x_1 = x_0 \implies \tilde{p} = p$

from 1980 to 2016. Patent reassignment (transaction date, identity of buyers and sellers) are from the USPTO Patent Assignment Database (PAD) (Graham et al., 2018), which records details on the transfer of ownership between patent assignees. To account for sample truncation, we limit our sample to patents granted on or before 2011 (for which we observe reassignments until 2015).[15] The final sample consists of about 3.9 million patents, of which 6.3% are reassigned at least once. We describe next the main steps taken to construct the sample and main variables.

Science-based inventions — We define science-based inventions as those that make at least one citation to a scientific article (Narin et al., 1997; Arora et al., 2020; Roach and Cohen, 2013; Sampat, 2010). We use data from Marx and Fuegi (2020), which matches US patents to scientific publications in Microsoft Academic Graph (MAG) to identify pairs of citing patents and cited scientific publications (see appendix A.1.1 for more details). We identify 724,395 patents out of 3,883,777 that cite at least one scientific article in MAG on their front page.[16] However, there could be still be measurement error. Science-based patents may not cite science, and conversely, some citations to science may be "pro-forma", not really reflecting reliance on science. Accordingly, we also analyze market (patent class) level regressions because aggregation should reduce classical measurement error: the share of science-citing patents in a patent class is arguably a more reliable indicator of the extent to which inventions in that class rely on science. Our results are robust to an alternative measures of reliance on science, in-text citations, and the textual similarity between patents and scientific publications, as shown below.

Patent reassignments — We measure MFT by the patent reassignments in the USPTO Patent Assignment Dataset (PAD) from 1980 to 2015 for patents granted on or before 2011 (Marco et al., 2015).[17] The USPTO records transfers of ownership that occur between patent assignees. While the reporting of transfers is voluntary, firms that acquire patents have an incentive to report transfers, particularly in enforcing the acquired patents. We build on prior researchers, who have cleaned reassignments data to obtain those related to MFT transactions (Marco et al., 2015; Serrano, 2010).[18]

Invention quality — We use three methods to measure the quality of the patented invention. First, we use data from Patstat to count the the number of *forward patent citations* a patent has received and normalize this by the average number of citations received by all patents in the focal patent's publication year. Our second measure is whether the patent is a *triadic patent* i.e., is registered in the three largest

[15] About 58% of reassignments are within five years after patent grant.

[16] See appendix A.1.1 for further details on variable construction.

[17] We focus on patent trade, which, unlike licensing, entail an exclusive transfer of property rights. Reliable large scale licensing data are not easily available, particularly for private firms.

[18] 6.3% of our sample's patents are reassigned at least once. See appendix A.1.2 for details on the cleaning procedure.

patent jurisdictions - the European, Japanese, and U.S. patent offices (Dernis and Khan, 2004). That the same invention is patented in all three offices implies that the value to the inventor is high. Third, for firms listed in American stock exchanges, we use the *stock market valuation* of patents from Kogan et al. (2017), which uses the excess stock returns for patenting firm on the date of the patent's issuance date recorded in the USPTO official gazette. In addition, we use the *number of claims*, and the *length of the first claim* as other measures for the quality of the patent.

Invention novelty — We use two alternative measures of novelty. The first uses *patent textual similarity* to prior patents. Building on Arora et al. (2018), for each focal invention, we calculate its textual similarity score for all previous patents (all USPTO patents with an earlier priority date than the focal invention). We normalize the proximity scores vector of the top 100 closest citation pairs for each focal patent by dividing each score by the corresponding maximum pairwise textual score for the focal patent. We average the standardized scores to derive a single textual proximity score for each focal invention.

The second measure of novelty is the technology *combination familiarity* measure from Fleming (2001). We count the number of times the same combination of patent sub-classes had appeared before the focal patent's publication date. The assumption is that combinations of sub-classes that appear more often should be more familiar.[19]

Size — We measure the commercialization capability of an inventor by its declared size in USPTO maintenance fee payment records. Firms with less than 500 employees are classified as small, and pay 50% lower filing and maintenance fees. Second, we use initial patent assignee names matched to public company names in Compustat from the DISCERN project (Arora et al., 2020), and use inclusion in Compustat as another indicator of size.

Buyers' heterogeneity — We measure buyer heterogeneity as the top four-assignee concentration ratio by patent class-years. The higher the share of patents assigned to the top four patentees in a patent class, the more unequal the distribution of valuations of inventions. We first extract the assignee names that are disambiguated in the HBS inventor dataset (Lai et al., 2011).[20] We then calculate a four-assignee concentration ratio by dividing the patent stock of the four most frequent assignees by the patent stock of all assignees in a 4-digit IPC-year.[21]

[19] In our sample, the combination familiarity score ranges from 0 (first combination of its kind) to 174 (appeared 174 times before) with a mean of 76.8. In the regression analysis, we refine this measure so that the count exponentially decays with time at an annual rate of 18%. That is, a previous patent subclass combination from five years ago is weighted by $exp(-\frac{1}{5}) = 37\%$. This time decay allows for even old technological combinations to exhibit higher novelty if sufficient time passes by between patenting activities.

[20] Available from `https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/5F1RRI`

[21] By constructing this measure, we implicitly assume that the assignees approximate the potential buyers in a technology

Marginal inventor characteristics — In empirical analysis at the patent class level, we proxy the capability of the marginal inventor by the average size of patent holders, the total number of unique sellers, and the share of entrants to patenting in a 4 digit IPC-year. Average size of patent holders is defined as patents granted to "small" assignees based on application and maintenance fee payment divided by patent stock in each patent 4-digit IPC-year. To identify the number of unique sellers, we cluster similar assignee names by using string distance measures. Assignee name pairs that are sufficiently similar to each other are then treated as a single name.[22] The share of new entrants is defined as the number of entrants as a share of total assignees for a 4 digit IPC-year.

Table 1: Summary Statistics for Main Variables

	Obs.	Mean	Std. Dev.	10%	50%	90%
Patent Publication Year	3883777	1998.886	8.725	1986	2000	2010
Reassignment Dummy	3883777	0.063	0.243	0	0	0
Cite Science Dummy	3883777	0.179	0.383	0	0	1
ln(IPC Combination Familiarity+1)	3878063	1.198	1.694	0	0	4
Textual Similarity	2425203	0.224	0.130	0	0	0
Small Entity Dummy	3689237	0.226	0.418	0	0	1
Compustat Patent Dummy	3883777	0.265	0.441	0	0	1

Notes: Reassignment is a binary variable equal to one if the patent has ever been reassigned in the USPTO PAD dataset. *Cite Science* is equal to one if there has been a citation to Microsoft Academic Graph (MAG), and zero otherwise. *Combination Familiarity* of a patent is constructed by counting the number of times a patent's IPC sub-class combinations have appeared in the past (details in Fleming (2001)). *Small Entity* is equal to one if an assignee is classified as a small entity by section 41 of the U.S. patent act, and zero otherwise. *Compustat Patent* is equal to one if an initial assignee is matched to a Compustat firm, and zero otherwise.

4 Econometric framework

For a given patent, we confirm that reliance on science increases the probability of trade, especially when the invention is novel, the inventor is small, and faces heterogeneous buyers. These relationships hold as we aggregate up to the IPC level. In addition, we confirm that reliance on science is associated with greater entry into invention, especially of small inventors. Finally, we develop sources of exogenous variation in the reliance on science at the market level to estimate the structural relationship between reliance on science and MFT.

market, and that 4-digit IPC classes are appropriate delineators of technology markets.

[22]This prevents misspellings or differences in legal nomenclature (Corp, Inc, Ltd etc.) from classifying a single assignee into two different entities. To a limited extent, this strategy also allows us to identify and unify technology licensing arms or divisions of companies, provided the name of the company is long enough. We define entrants as assignees that are patenting for the first time since the beginning of our sample in 1980.

4.1 Baseline trade equation (OLS)

We estimate a patent level equation for the likelihood a patent is traded :

$$Reassignment_i = \beta_1 s_i + \boldsymbol{Z_i'}\boldsymbol{\gamma} + \boldsymbol{\xi_t} + \boldsymbol{\psi_c} + \upsilon_i \tag{12}$$

Reassignment is equal to one for if the patent is reassigned at least once during its term and zero otherwise.[23] Reliance on science, s_i, is equal to one for patents with at least one NPL citation to MAG and zero otherwise, and Z_i contains a variety of controls for quality. As well, we include complete set of dummies for the patent grant year (ξ_t) and its 4-digit IPC (ψ_c). υ_i is unobserved patent level characteristics. We expect $\hat{\beta}_1 > 0$.

We examine Results 1, 2, and 3 and interact s_i with the inventor size, market concentration, and invention novelty, respectively. Per Result 1, we expect the level relationship between size and patent reassignment to be negative, as science increases the gains from trade more for small inventors than large. Result 2 implies that the interaction term between reliance on science and market concentration should be positive, while Result 3 implies that the interaction term between science and novelty should be positive, as science reduces the higher transaction cost due to newness.

We argued that reliance on science can lower transfer costs as well as increase gains from trade. To disentangle the relative importance of these two mechanisms, we use the characteristics of the science being cited in our patents. For example, if transfer cost reduction (τ_i) is the sole mechanism, we may expect patents building on older science to be more likely to trade. Mature and established theories and empirical results are likely to have weathered more frequent and rigorous tests of validity (and attempts at falsification). However, for gains from trade ($\bar{x}_i - x_i$), we would expect patents using younger science to trade more often. Patents building on "cutting edge" science have greater uncertainty relating to their commercialization value.

We also explore whether patents that draw more specialized science are more likely to be traded compared to patents that are more narrowly specialised. Jones (2009) has argued that scientific progress has increased the returns to specialization, necessitating greater collaboration among such specialists. Jones (2009) himself focused on collaboration in the production of knowledge. However, it is plausible that the application of knowledge may also require collaboration. MFT is an important mechanism for

[23] A patent that is reassigned multiple times gets the same *Reassignment* value of one as a patent that has been reassigned once.

such vertical specialization and collaboration. In other words, inventions that draw upon specialized science probably require more follow-on invention, and therefore, offer greater gains from trade.

4.1.1 Entry of small inventors

To see whether reliance on science favors small inventors, we estimate a specification at 4-digit IPC c, publication year t level,

$$Share\ of\ Small_{ct} = \beta_0 + \beta_1 Share\ of\ Science\ Citing\ Patents_{ct} + \boldsymbol{Z'_{ct}}\gamma + \boldsymbol{\xi_t} + \boldsymbol{\psi_c} + \upsilon_{ct} \tag{13}$$

We proxy the commercialization capability of the marginal seller by (i) the share of small patentees (ii) the number of sellers and (iii) entrant share, and regress these against the share of science citing patents out of all granted patents at the 4 digit IPC-year level. We control for the share of triadic patents, average number of claims, and average length of the first claims because technological advances may encourage the entry of new sellers. We also include IPC and year fixed effects to exclude the effect of any year or technology class-specific differences.

4.2 Instrumental variable strategy and other robustness analyses

4.2.1 Measurement error

It is plausible that we measure the reliance on science with error. We directly probe the robustness of our measure in three ways. First, we add a dummy for whether the citations to the scientific article also appears in the body text of the patent, in addition to the front page NPL section. Second, we weight the front page citations to science by the number of forward citations received from MAG articles by the cited science. Third, we calculate a measure of textual similarity between focal patent text to scientific articles published in the Web of Science, using data from Arora et al. (2018).

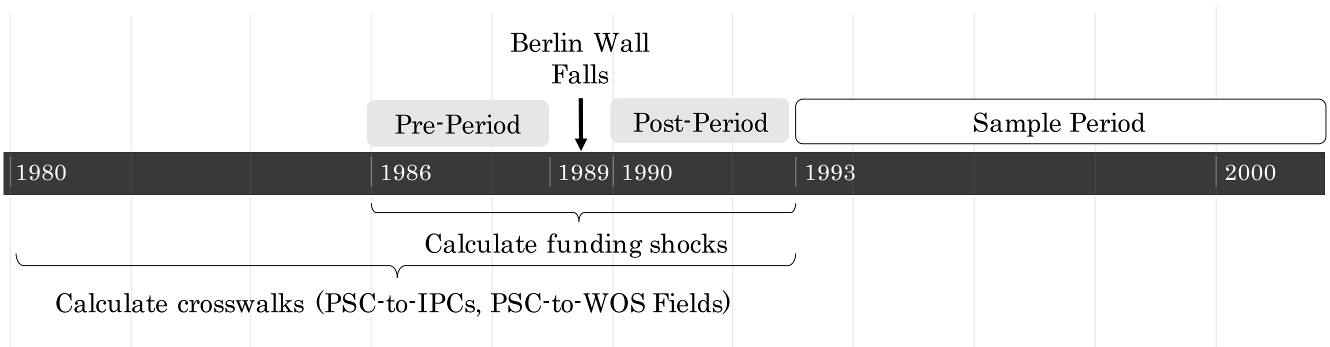
4.2.2 Unobserved heterogeneity and instrument variable analysis

The association between reliance on science and MFT may also reflect other factors. Equation 11 showed that unobserved differences in the demand for inventions, N, would result in a spurious correlation between $\tilde{p}$, the observed share of science citing patents, and the share of patents that are traded.[24] To

[24] At the patent level as well, equation 9 implies that unobserved variation in commercialization capability is likely correlated with reliance on science, potentially creating biased estimates.

address this concern, we exploit a quasi-experiment where the cost of relying on science falls due to an exogenous rise in the relevant knowledge for some inventions, but not others. Specifically, we use the reallocation of federal funding for R&D around the end of the Cold War as a source of exogenous change in the availability of relevant scientific knowledge. We use the changes in the predicted stock of scientific knowledge as an instrument for the share of patents citing science to purge the effect of unobserved demand factors, as well as purge it of measurement error. This procedure allows us to estimate the causal effect of science on MFT.

Figure 1: Timing Structure of Instrumental Variable Estimation

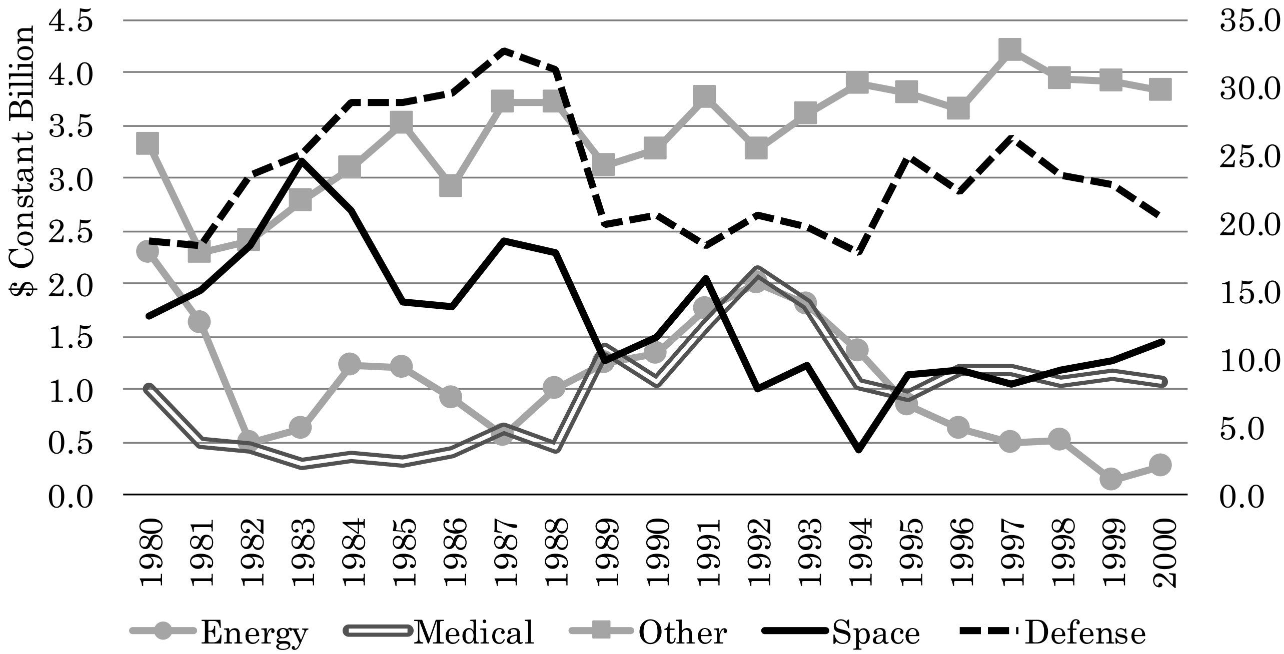


Government funding for R&D accounted for close to half of all research and development in the United States between 1980 and 1995 (see figure 2). Merrill (2018) shows that the end of the Cold War has resulted in a significant reallocation of the federal government's research portfolio, both within and between agencies. Between agencies, the dominant position of the Department of Defense (DoD) yielded to increased support for the Department of Health and Human Services. Within DoD, funding shifted out of physics, chemistry, and electrical engineering into computer sciences, oceanography, and biology during the 1990s. In a dataset of federal contract R&D data collected from USAspending.gov, we find that annual obligations of the U.S. government to contractors has decreased for defense and space technology, while increasing for medical and energy related items around the years after the end of the Cold War (see figure 2).[25]An increase in public funding for a scientific discipline increases the stock of knowledge, and hence increases the share of inventions relying on the focal science (Azoulay et al., 2019; Fleming et al., 2019; Moretti et al., 2019).

Our instrument is defined at the 4-digit IPC level and measure the i) the predicted number of relevant scientific papers to a 4-digit IPC and ii) the difference in average federal R&D funding between

[25]Appendix table A2 in section A.2 summarizes this resource reallocation

Figure 2: Government Contract R&D Funding, During and After the Cold War



Notes: This figure plots the aggregate value of the contracts signed by the federal government for research and development, separated by main 2-digit Product and Services Codes: Energy (AG), Medical (AN), Space (AR), Defense (AC&AD). Figures are adjusted to 2012 dollars using GDP deflators from Louis Johnston and Samuel H. Williamson, "What Was the U.S. GDP Then?" MeasuringWorth, 2020.

the post (1990-1992) and pre (1986-88) periods. Our first instrument is created in three steps. First, we create estimates of the stock of scientific knowledge (number of papers) relevant for each 4 digit IPC, using the share of NPL citations to different scientific disciplines by patents in an IPC. Second, we run an OLS regression at the Web of Science Field-year level to predict the number of papers published in a given field and year between 1992 and 2000. Third, we weight the predicted stock of knowledge due to the shocks to federal R&D funding in each scientific discipline by its relevance to a focal IPC to create the predicted shock to scientific knowledge for each IPC due to the end of the Cold War.[26]

We implement our IV estimation with two-stage-least-squares. In the first stage, we predict the share of science-citing patents in a 4-digit IPC-year using our instrument, predicted relevant science. In the second stage, we estimate the share of traded patents as a function of the predicted share of science-citing patents. The first stage of this IV specification confirms that the reliance on science is a function of available knowledge stock, which we have shown in column 1 of table 9 to be positively related to post Cold War federal R&D funding shocks. The first stage regressions in table 9 shows that the relevant F statistics are over 100. Thus, changes in federal procurement spending due to the fall of the Soviet Union

[26]Please see appendix section A.2.2 for details.

had an appreciable effect on the stock of scientific knowledge relevant to invention as well as on the share of patents relying on science.[27]

An alternate instrument is simply the difference in average government funding relevant to IPCs. For each fiscal year, we calculate the average federal R&D contract obligations for each Product and Services Code (PSC)[28] and match it to 4-digit IPCs based on the share of patents filed in each IPC by DISCERN firms contracting in the focal PSC, between 1980 and 1992.[29] We average these R&D dollars "relevant" to each 4-digit IPC, before (1986-88) and after (1990-92) the end of the Cold War. The difference of the natural logs of these values is our instrument that is used to predict the share of science-citing patents at the 4-digit IPC-year level after the Cold War (1992-2000).

A potential concern is that the end of the Cold War also changed demand conditions across technology fields. The federal government spent on average $315 billion a year on procurement between 1986 and 1992, where R&D contracts average $34.9 billion, accounting for 11% of total procurement. Specifically, non-R&D spending cuts may reduce demand in select fields, and thereby reduce the rate of invention in the field, possibly affect the share of science-based inventions as well as trade in technology. That is, changes in government procurement spending may also have affected N, the number of potential buyers of inventions (Lichtenberg, 1987).

We address this concern in three ways. First, the correlation between spending differences for R&D and non-R&D is positive but relatively low (r=0.257).[30] For instance, patent class "C07H" (SUGARS; DERIVATIVES THEREOF; NUCLEOSIDES; NUCLEOTIDES; NUCLEIC ACIDS) experiences a four-fold increase in its R&D funding between 1986 and 1992, while only a 3% decrease is observed for non-R&D funding. Pharmaceutical firms that receive medical research contracts also sell drugs (PSC 6505) and medical equipment (6515) to the federal government, which tends to be stickier (e.g. in VA medical centers). Second, we directly control for non-R&D federal contract spending. Third, we add robustness checks in appendix table A3 that derives funding shocks net of number of patents, share of reassigned patents and patent forward citations.

[27] We allocate federal R&D funding to 198 Web of Science Fields by weighting the number of MAG publications matched to contracting DISCERN firms in each PSC. We then regress the number of scientific publications with at least one author affiliation in the United States against the difference in logged funding for each Web of Science field between the pre (1986 to 1988) and post (1990 to 1992) period. Please see appendix A.2 for further details. In unreported robustness checks we verify that the results are unchanged when we weight the number of papers with forward citations received by the papers within five years of their publication..

[28] "The Product and Service Codes (PSC) Manual provides codes to describe products, services, and research and development (R&D) purchased by the federal government. These codes indicate "what" was bought for each contract action reported in the Federal Procurement Data System (FPDS)." 2015 Edition of the Federal Procurement Data System's Product and Service Codes Manual (Available from `https://www.fpds.gov/downloads/top_requests/PSC_Manual_FY2016_Oct1_2015.pdf`)

[29] Please see appendix section A.2.2 for details on the crosswalk.

[30] See figure A3 for a visual comparison

5 Estimation results

5.1 Reliance on science and trade

Table 2 contrasts the reassignment probability of patents that cite science and those that do not. We find that science-based patents are 1.4% (or 22% relative to the sample mean) more likely to be traded than those that are not based in science. In section 2, we argued that science leads to more trade because of its ability to reduce transfer costs (τ_i) and affect the gains from trade ($x_i - x_{ik}$). However, it is also possible that science-based inventions have higher quality (q_i). Indeed, table 2 shows that patents citing science have higher forward patent citations, are likely to be triadic patents, and have more claims and higher stock market values. Consistent with our model, reassigned patents also tend to exhibit higher values of these proxies of quality compared to those not reassigned.

Table 2: Mean Comparisons of Patent Characteristics, by Science and Reassignment Status

	T-Test		Cite Science = 0			Cite Science = 1		
	Diff.	Std. Error	Count	Mean	SD	Count	Mean	SD
Reassignment (%)	1.3879***	0.0316	3159382	6.023	23.791	724395	7.411	26.195
5-year Forward Patent Citations	3.4266***	0.0140	3159382	5.455	9.328	724395	8.882	15.512
Triadic Patent	0.1832***	0.0006	3159382	0.270	0.444	724395	0.453	0.498
Number of Claims	4.7358***	0.0164	3158923	14.454	11.470	724387	19.190	16.503
Stock Market Value of Patent (KPSS)	5.7256***	0.0827	945391	11.332	35.157	272502	17.057	46.715
	T-Test		Reassignment = 0			Reassignment = 1		
	Diff.	Std. Error	Count	Mean	SD	Count	Mean	SD
5-year Forward Patent Citations	2.5321***	0.0226	3639800	5.935	10.422	243977	8.467	15.592
Triadic Patent	0.0465***	0.0010	3639800	0.301	0.459	243977	0.348	0.476
Number of Claims	2.3334***	0.0265	3639335	15.191	12.497	243975	17.524	15.217
Stock Market Value of Patent (KPSS)	-1.5637***	0.1576	1156249	12.692	38.362	61644	11.128	33.312

Notes: Comparison of means at the patent level. Standard errors are in parentheses. Significance levels are annotated as $^{*} < 10\%$ $^{**} < 5\%$ $^{***} < 1\%$

5.2 Baseline trade equation

Table 3 presents the Linear Probability Model (LPM) estimates. Controlling for year and 4-digit IPC fixed effects, column 1 shows that citing a scientific article is associated with a 23% higher probability of a patent being traded, relative to the sample mean.[31] Patents that cite science may represent higher quality inventions and better crafted patents. We therefore control for triadic patent status, number of claims, and the length of the first claim. More independent claims and shorter independent claims are related to broader patent scope (Kuhn and Thompson, 2019). In addition we measure whether a patent

[31] Unless stated otherwise, the percentage magnitudes reported here are relative to the sample mean in each specification.

has been filed in multiple patent jurisdictions (U.S., Europe, and Japan). Such triadic patents tend to be of high private value, and hence, tend to be of higher quality.[32] Column 2 shows that number of claims and triadic patenting status are positively correlated with patent trade. The coefficient on science citation decreases in magnitude by 28%, but remains positive and statistically significant. For a subset of patents that are issued to U.S. listed firms, we measure their market valuations based on excess stock price returns of inventing firms on their grant dates (Kogan et al., 2017). Citing science continues to have a positive and statistically significant relationship with reassignment.[33]

Splitting the sample by technology classes (Columns 4-7, Table 3), we find that the *Cite Science Dummy* coefficient in the Life Sciences is over three times as large as in ICT. Value chains in the ICT sector tends to be more complex, possibly muting the gains from having a clearer scientific grounding for one invention. The life sciences, in contrast, have clearly delineated targets and therapeutic areas that are tackled by clearly structured molecular compounds. Therefore, transfer cost reduction from clarifying

[32]In columns 2 and 3 of appendix table B2, we also add dummies for quintiles and deciles of five-year forward patent citations. More highly cited patents are traded more, but the positive relationship between reassignment and science citation remains significant.

[33]Interestingly, the relationship between stock market valuations and patent trade is negative, perhaps because this captures the value idiosyncratic to a firm (x_i in the model) rather than common quality (q_i).

the underlying mechanisms by referencing science in an invention may be larger for the life sciences.

Table 3: Markets for Technology and Reliance on Science (OLS)

	Baseline Science			By Technological Sector			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	Reassignment = 1						
	Baseline	Quality Controls	Compustat Sample	Life Sci	Chem	ICT	Other
Cite Science Dummy	1.424**	1.026**	0.360**	1.134**	0.831**	0.869**	1.019**
	(0.038)	(0.038)	(0.128)	(0.122)	(0.092)	(0.058)	(0.083)
Triadic Patent Dummy		0.914**	0.532**	2.493**	0.306**	1.226**	0.559**
		(0.029)	(0.158)	(0.100)	(0.077)	(0.058)	(0.045)
Number of Claims		0.082**	0.020**	0.081**	0.067**	0.067**	0.096**
		(0.001)	(0.004)	(0.004)	(0.003)	(0.002)	(0.002)
Length of First Claim		-0.000**	0.000	-0.000**	-0.000	-0.000	-0.000**
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mkt Value of Patent (KPSS)			-0.002				
			(0.001)				
Avg of DV	6.283	6.283	5.589	8.680	6.564	5.678	6.295
4-digit IPC Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Fixed Effects	No	No	Yes	No	No	No	No
R^2	0.010	0.012	0.156	0.010	0.009	0.014	0.010
N	3,883,776	3,882,632	797,912	357,124	473,782	1,020,997	1,704,044

Notes: Unit of analysis is the patent. Patent reassignment is a binary variable equal to 100 if the patent has ever been reassigned in the USPTO PAD dataset and zero otherwise. *Cite Science Dummy* is one if there has been a citation to Microsoft Academic Graph (MAG), and zero otherwise. *Triadic Patent Dummy* is one if the patent shares a prior art in the USPTO, EPO, and JPO, and zero otherwise. *Number of Claims* counts the number of independent and dependent claims in a patent. *Length of First Claim* counts the number of words in the first claim of the patent. *Mkt Value of Patent (KPSS)* is the value of a patent (in million dollars) based on the cumulative abnormal returns in the firm's market value at the issuance event of the patent Kogan et al. (2017). Standard errors are robust to arbitrary heteroscedasticity for all columns except for column 3, whose standard errors are clustered at the Compustat firm level.

5.2.1 Measurement error and validation

We introduce three robustness checks per section 4.2.1. We include "in-text" citations to MAG publications that appear in the body text of the patent. In-text citations do not affect the patentability of an invention as much as front page NPL citations. Therefore, their inclusion may signal a greater reliance on science by the inventor. We find in column 1 of table 4 that in-text citations to science are positively correlated with patent trade.

Table 4: Markets for Technology and Reliance on Science, Alternative Measures (OLS)

	DV: Reassignment=1		
	(1)	(2)	(3)
Cite Science Dummy: In-text	0.878**		
	(0.044)		
ln(Cite Science: FwdCitation Weighted)		0.162**	
		(0.006)	
Textual Similarity to Science			0.114**
			(0.007)
Triadic Patent Dummy	0.945**	0.913**	0.974**
	(0.028)	(0.029)	(0.029)
Number of Claims	0.083**	0.082**	0.084**
	(0.001)	(0.001)	(0.001)
Length of First Claim	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)
Avg of DV	6.283	6.283	6.283
4-digit IPC Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
R^2	0.012	0.012	0.012
N	3,882,632	3,882,632	3,882,632

Notes: Unit of analysis is at the patent level. "Cite Science Dummy: In-text" is equal to one when a scientific article is cited in the body of the patent text as an in-text citation. "ln(Cite Science: FwdCitation Weighted)" takes the natural log of the total number of citations papers cited by a patent receive from other scientific papers in Microsoft Academic Graph. "Textual Similarity to Science" measure the textual overlap of a patent's text with abstracts of articles in Clarivate Web of Science's Science Citation Index - Expanded. All other variables definitions are identical to table 3. Standard errors are robust to arbitrary heteroscedasticity.

On the other hand, the science surrounding an invention may be too established or canonical such that too few citations are made. For example, GPS technology relies on Einstein's general theory of relativity for its time dilation corrections between satellites and ground receivers, but few GPS patents cite his papers. We therefore weight the science citation dummy by the quality of the cited science in column 2 of table 4 and find that patent reassignment is still positively correlated with the number of citations a cited paper receives from other papers within MAG. We also measure citations to science that are linked through another patent citation (Jones, 2009). The distance to the "citation frontier" is zero for patent with a citation to MAG ($D = 0$), while a patent that cites such a patent has a distance of one ($D = 1$). We calculate this distance up to 10 citation links and find, as shown in appendix table B1 that the positive association between citing science and patent reassignment persists. In appendix table B1, we find that the estimated coefficient of science citation increases from 1.026 when a patent cites science directly to 1.190, when a patent cites a patent that cites science once removed from the citation frontier ($D = 1$). The coefficient decreases as the distance from the citation frontier increases.

We offer three additional validity checks of our citation-based measure of reliance on science. First, Arora et al. (2020) show that patents that cite scientific publications are more likely to respond in the

Carnegie Mellon Survey that public research findings (from government and academia) are important for their inventions. They also find that firms tend to recognize the importance of the field they cite in their patents. Firms that cite scientific articles in patents also tend to operate a greater basic scientific research program as a share of total R&D.

Table 5: Reliance on Science, for Inductees of the National Inventors Hall of Fame

Group	Variable	Obs.	Mean	Std. Dev.	10%	50%	90%
Patents by National Inventor Hall of Fame (NIHF) Inductees	Cite Science Dummy	97	0.567	0.498	0	1	1
	Textual Simil. to Science	97	4.634	2.002	1.609	4.875	6.978
also Nobel Laureate	Cite Science Dummy	8	0.750	0.463	0	1	1
	Textual Simil. to Science	8	6.539	1.473	3.871	6.228	8.388
also Lasker Laureate	Cite Science Dummy	1	1	.	1	1	1
	Textual Simil. to Science	1	9.183	.	9.183	9.183	9.183
also Turing Laureate	Cite Science Dummy	4	1	0	1	1	1
	Textual Simil. to Science	4	4.292	2.126	1.386	4.718	6.347
also Franklin Institute Award Laureate	Cite Science Dummy	11	0.636	0.505	0	1	1
	Textual Simil. to Science	11	5.546	1.656	3.526	6.054	7.309
also NMTI Laureate	Cite Science Dummy	18	0.556	0.511	0	1	1
	Textual Simil. to Science	18	4.482	2.244	0.693	5.454	6.347
non-NIHF Patents	Cite Science Dummy	3,883,680	0.179	0.383	0	0	1
	Textual Simil. to Science	3,883,680	3.541	2.037	0.693	3.555	6.213

Notes: This table shows summary statistics on reliance on science for patents published between 1980 and 2011 that are also matched to inductees of the National Inventors Hall of Fame (NIHF). Nobel Laureates include those in physics, physiology or medicine, and chemistry only. The NMTI refers to the National Medal of Technology and Innovation. Cite Science Dummy is equal to one if the patent cites at least one scientific article from MAG. Textual Similarity to Science is a continuous measure of textual overlap between a patent and scientific articles from Clarivate Web of Science between 1990 and 2015.

Second, we focus on a subset of patents whose inventors were inducted to the National Inventor Hall of Fame (NIHF) and examine whether high-caliber inventors that are also scientists are more likely to cite science, compared to similarly high-caliber inventors. The NIHF inductees are recognized for "great technological advances that make human, social and economic progress possible," which may range from Post-it Notes to blue LEDs.[34] The NIHF lists descriptions of the nature of the invention and the contribution of the inductee, together with the USPTO patent number for the most representative invention of that inductee.[35] We are able to link 97 patents for 111 inductees for our sample period.[36] We manually searched for whether each inductee had been awarded any one of four prestigious awards in science: the Nobel Awards in physics, medicine, and chemistry, the Lasker Award (for medicine), the Turing Award (for Computer Science), and the Franklin Institute Award (Arts et al., 2020). We find that 57% of the patents by NIHF inductees cite science, which is more than three times the sample average.

[34] Spencer Silver, Patent Number: US3691140A, and Shuji Nakamura, Patent Number: US5290393A.

[35] https://www.invent.org/NIHF-hall-of-fame-inductees-list-alphabetical

[36] While the inventions are patented between 1980 and 2011 to fit with our sample, inventor years of birth range between 1914 and 1975.

However, we find that this probability increases to 75% for patents by Nobel Laureates, while Lasker and Turing Award winners always cite science in their most representative patents listed by the NIHF. We also link these inductees to the winners of National Medals of Technology and Innovation (which does not specifically require a scientific contribution) and find that they cite science slightly less often (56%).

Third, we calculate an alternative measure of reliance on science based on the textual overlap of patent text with abstracts from scientific articles. We leverage data from Arora et al. (2018), which calculates a weighted cosine similarity measure between the full text of patents in our sample and the abstracts of scientific articles in Clarivate Web of Science's Science Citation Index - Expanded. Using this textual overlap measure between patent-paper pairs, we keep the top 100 patents most similar to the focal paper. We count the number of times a given patent is classified within this "top 100" set, which we use as an alternative measure for reliance on science. We also tried the patent level averages of the ranks, as well as the similarity scores with respect to scientific publications. We replicate the baseline results using this new measure in column 3 of table 4, finding that a standard deviation increase in textual similarity to science is associated with a 3.7% greater reassignment probability.

Table 6: SCIENCE AND MFT, BY SELLER SIZE, BUYER HETEROGENEITY, AND PATENT NOVELTY (OLS)

	Seller Size		Buyer Hetero	Patent Novelty		Science Characteristics	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	Reassignment = 1						
	Maint.Fee	Compustat	C4 Share	ABCL	Fleming	Recency	Specialization
Cite Science Dummy	0.605**	-0.410**	0.900**	1.557**	1.314**		
	(0.040)	(0.053)	(0.064)	(0.080)	(0.046)		
Avg(Lag to Cited Science)						-0.021**	
						(0.004)	
1-Normalized Field Counts							0.795**
							(0.102)
Small Entity Dummy	0.040						
	(0.041)						
Small Entity Dummy × Cite Science Dummy	2.531**						
	(0.101)						
Non Compustat Dummy		0.769**					
		(0.033)					
Non Compustat Dummy × Cite Science Dummy		2.374**					
		(0.069)					
C4 Share			0.010				
			(0.203)				
C4 Share × Cite Science Dummy			1.848**				
			(0.499)				
Textual Similarity				-2.001**			
				(0.133)			
Textual Similarity × Cite Science Dummy				-1.845**			
				(0.258)			
ln(IPC Combination Familiarity+1)					0.065**		
					(0.008)		
ln(IPC Combination Familiarity+1) × Cite Science Dummy					-0.225**		
					(0.018)		
Triadic Patent Dummy	0.996**	0.899**	1.016**	1.166**	0.933**	1.823**	1.794**
	(0.030)	(0.029)	(0.033)	(0.034)	(0.030)	(0.062)	(0.062)
Number of Claims	0.081**	0.085**	0.094**	0.055**	0.082**	0.060**	0.059**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Length of First Claim	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**	-0.000*	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Avg of DV	6.265	6.283	6.938	4.877	6.274	7.401	7.420
4-digit IPC Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.012	0.012	0.012	0.011	0.012	0.015	0.015
N	3,689,008	3,882,632	3,287,801	2,424,715	3,877,819	830,868	830,868

Notes: Unit of analysis is at patent level. *Textual Similarity* averages the patent-to-patent pairwise text similarity scores from Arora et al. (2018) for each focal patent, normalized by its maximum score. *Combination Familiarity* of a patent is constructed by counting the number of times a patent's IPC sub-class combinations have appeared in the past (details in Fleming (2001)). *Small Entity* is equal to one if an assignee is classified as a small entity by section 41 of the U.S. patent act, and zero otherwise. *Non Compustat* is equal to one if an initial assignee is not matched to a Compustat firm, and zero otherwise. *Avg(Lag to Cited Science)* is defined as the average difference in the grant year of a patent and the publication year of a cited scientific article. The rest of the variable definitions are identical as table 3. Standard errors are robust to arbitrary heteroscedasticity.

5.2.2 Inventor commercialization capability

We investigate the empirical support for Result 1 that patents owned by smaller firms are more likely to be traded, especially for patents that cite science. We measure seller commercialization capability (x_i) through patent ownership by "small" firms recorded in USPTO maintenance fee payments, and by Compustat companies. We expect that smaller patentees are more likely to sell their invention than larger ones, and the association is stronger for patents that cite science. Table 6 column 1 shows that small firm patents that cite science are 50% ($= \frac{3.13\%}{6.28\%}$) more likely to be sold, compared to small firm patents not citing science.[37] Column 2 shows that patents owned by non-Compustat owners are on average 0.05% more likely to be sold relative to Compustat patents, but the gap for patents citing science is larger by two orders of magnitude. We replicate this result using textual similarity measures to science in columns

[37] In unreported checks, we confirm that the level effect of regressing patent reassignment against small *Small Entity Dummy* without interactions with *Cite Science Dummy* in column 1 is positive and significant.

1 and 2 of table 7. The predicted reassignment gap between small and large firms is around 5 times larger for patents in the 75th percentile of similarity scores (4.997) compared those in the in the 25th percentile (2.079). The gap between non-Compustat and Compustat firms is around 48% larger for 75th percentile similarity scores compared to 25th percentile.

5.2.3 Buyer heterogeneity

Result 2 predicts that industries with more heterogeneous buyers will exhibit more trade, and reliance on science magnifies the gap. We measure industry concentration by the share of C4 patentee-owned patents in a focal patent's 4-digit IPC. Because C4 patentees are likely to be large entities, simply regressing reassignment against the share of C4 patentees will measure the effects of patentee size rather than buyer heterogeneity. Therefore, we limit our sample to patents owned by non-C4 patentees. Column 3 of table 6 tests these predictions. For non-C4 patentees, being in a concentrated market where 90% of all patentees are C4 patentees leads to a 4% gain in reassignment probability compared to one where only 10% of all patentees are C4 patentees. However, patents that cite science in those concentrated markets are 38% more likely to be traded than those that do not cite science. We also replicate these findings in column 3 of table 7 replacing citations to science with textual similarity to science.

5.2.4 Invention novelty

Recall that more novel inventions may have higher gains from trade but also higher transaction costs, implying that the relationship between novelty and trade is theoretically ambiguous. Result 3 implies, however, that novel patents that rely on science have a higher probability of trade than novel patents that do not cite science. We test Result 3 by interacting science with two measures of invention novelty: textual similarity to prior patent art (Arora et al., 2018) and patent subclass Combination Familiarity (Fleming, 2001). The interaction terms in columns 4 and 5 of table 6 are negative and statistically significant, consistent with our prediction. Patents whose subclass combinations are in the first decile of Combination Familiarity scores (in other words, novel patents) are 14% less likely to be traded compared to those in the tenth decile (not novel patents). However, novel patents that are based in science are 21% more likely to be traded than novel patents not based in science, whereas the same difference for not novel patents is 7%: the effect of science on reassignment is close to three times larger in novel patents.[38]

[38] We confirm in unreported robustness checks that similar results hold with other text-based patent similarity measures (Kuhn, 2016).

Table 7: Textual similarity to science and MFT, by seller size, buyer heterogeneity, and patent novelty (OLS)

	Seller Size		Buyer Hetero	Patent Novelty	
	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Reassignment = 1				
	Maint.Fee	Compustat	C4 Share	ABCL	Fleming
Textual Similarity to Science	0.074**	-0.092**	0.088**	0.149**	0.160**
	(0.007)	(0.012)	(0.011)	(0.014)	(0.008)
Small Entity Dummy	-0.296**				
	(0.066)				
Small Entity Dummy × Textual Similarity to Science	0.217**				
	(0.015)				
Non Compustat Dummy		0.170**			
		(0.059)			
Non Compustat Dummy × Textual Similarity to Science		0.294**			
		(0.014)			
C4 Share			-0.494		
			(0.301)		
C4 Share × Textual Similarity to Science			0.206**		
			(0.073)		
Textual Similarity				-2.168**	
				(0.209)	
Textual Similarity × Textual Similarity to Science				-0.057	
				(0.050)	
ln(IPC Combination Familiarity+1)					0.149**
					(0.013)
ln(IPC Combination Familiarity+1) × Textual Similarity to Science					-0.038**
					(0.003)
Triadic Patent Dummy	1.056**	0.968**	1.079**	1.244**	0.990**
	(0.030)	(0.029)	(0.033)	(0.034)	(0.029)
Number of Claims	0.083**	0.087**	0.096**	0.057**	0.084**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Length of First Claim	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Avg of DV	6.265	6.283	6.938	4.877	6.274
4-digit IPC Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
R^2	0.012	0.012	0.011	0.011	0.012
N	3,689,008	3,882,632	3,287,801	2,424,715	3,877,819

Notes: Unit of analysis is at patent level. *Textual Similarity to Science* takes one plus the natural log of the number of times a patent is classified as one of 100 closest patents to a scientific article in Clarivate Web of Science's Science Citation Index-Expanded. The rest of the variable definitions are identical to those in table 6.

5.2.5 Cited science recency and specialization

The empirical results thus far indicate that reliance on science is associated with greater trade, even after controlling for the quality of the patented invention. This suggests that reliance on science may also lower transfer costs and increase gains from trade. Patents using recent science may be more likely to trade if gains from trade are at work, but if patents citing older, "textbook", science are more likely to be traded, then reductions in transfer costs may be at work. Column 6 of table 6 tests these predictions on 830,868 patents that cite at least one scientific article in MAG and finds that patents citing more recent science are more likely to be traded. A one standard deviation decrease in the average citation lag (7.7 years) is associated with about a 2% higher trade probability relative to the sample mean. Column 7 explores whether patents drawing on specialized science are more likely to be traded. We measure specialization by one minus the number of Web of Science fields found in the NPL citations of a patent divided by the number of papers cited. We find that a patent citing a standard deviation increase in specialization (.320) is associated with a 3.4% higher probability of trade relative to the sample mean. This supports the view that science increases gains from trade.[39]

In short, our results strongly suggest that the science-MFT relationship is not simply a consequence of unobserved differences in the quality of the patented invention. Our evidence indicates that the science-MFT relationship is stronger for inventions that are based in newer and more novel science, and for more novel inventions. Since transaction costs are likely higher for such inventions, these findings suggest that science based inventions also have higher potential gains from trade. Put differently, the science-MFT relationship is multifaceted, with science potentially reducing transaction costs as well as enhancing gains from trade.

5.3 Entry and market structure (OLS)

We turn to examining IPC level results. From Equation 9 we expect the commercialization capability of the marginal seller (inventor) to decrease as the use of science in an IPC-year increases. Table 8 presents the estimation results. We find that IPC-years that have a higher share of patents citing science tend to have a higher share of "small" patentees (column 1), larger number of sellers (column 3), and more first-time patentees (column 5).[40] Our estimates imply that a one standard deviation increase in science-citing patent share from the sample mean translates to a 18% gain in the share of small entities. A similar gain is observed for number of sellers: there are 0.068 sellers per patent on average, but a standard deviation higher citations to sciene have 0.077 sellers per patent. The share of entrants predicted for IPC-years in the 10th percentile of science-citing patent share is 28%, while it is 32% for patents in the 90th percentile.

[39] In appendix table B3, we redefine the citing patent-cited paper pair to be between a citing patent and a cited paper's cited paper. In the language of Ahmadpoor and Jones (2017), the results in table 6 use characteristics of "1^{st} degree" connections between patent and paper, while the results in table B3 use those for "2^{nd} degree" connections, reaching into deeper connections into the scientific literature. We find that the magnitude of the coefficients are smaller, but the direction and significance of the results hold.

[40] Patentees that are patenting for the first time since 1980, divided by number of patents. We exclude the first five years of our panel (1980-1985) to mitigate concern that the early years of the panel will have more entrants.

Table 8: SCIENCE, MFT, AND ENTRY INTO INVENTION

Dependent Variable:	Small Entity Share		No. of Sellers		Entrant Share	
	(1) Baseline	(2) Novelty	(3) Baseline	(4) Novelty	(5) Baseline	(6) Novelty
Avg(Cite Science Dummy)	0.315**	0.320**	0.057**	0.048**	0.134**	0.121**
	(0.034)	(0.033)	(0.010)	(0.011)	(0.023)	(0.023)
log(Avg(MAG Combination Familiarity)+1)		-0.003**		-0.000		-0.002*
		(0.001)		(0.000)		(0.001)
Avg(Triadic Patent Dummy)	-0.367**	-0.332**	-0.006	-0.009	-0.081**	-0.075**
	(0.031)	(0.032)	(0.008)	(0.008)	(0.021)	(0.020)
Avg(Number of Claims)	-0.004**	-0.004**	0.002**	0.001**	-0.002*	-0.002*
	(0.002)	(0.002)	(0.000)	(0.000)	(0.001)	(0.001)
Avg(Length of First Claim)	0.000	0.000	-0.000	0.000	0.000*	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log(Number of Patents+ 1)	-0.026**	-0.024**	-0.008**	-0.008**	0.006	0.006
	(0.005)	(0.004)	(0.002)	(0.002)	(0.004)	(0.003)
Avg of DV	0.281	0.260	0.068	0.260	0.291	0.279
IPCs	334	327	337	327	328	327
Years	31	31	32	31	26	31
R2	0.928	0.928	0.559	0.576	0.939	0.945
N	6,913	6,290	7,173	6,493	5,899	5,563

Notes: Unit of analysis is at the 4 digit IPC-year level. Observations with fewer than 100 patents are dropped. *Avg Cites to Science* is the count of patents in a 4 digit IPC-year that have made a citations to science. *Small entity share* is the number of small entity (<500 employee) patents. *Number of sellers* equals the number of unique patent sellers that have been identified for each 4 digit IPC-year. *Entrant Share* is calculated by dividing the number of new assignees (entrants) by the total number of assignees in each 4 digit IPC-year. The first five years of the panel are excluded for columns 5 and 6. All columns include fixed effects for 4-digit IPC and patent publication years. Standard errors are clustered at the 4-digit IPC level.

We expect less mature science to be more strongly associated with inventor entry if it affects gains from trade and more mature science if it affects transfer costs. In columns 2, 4 and 6, we include the average of the Scientific Combination Familiarity score calculated at the patent level for each IPC-year. We find that patent classes that are populated by inventions that use more novel (less familiar) scientific combinations are more likely to have smaller entities, more sellers and entrants. This supports the idea that the entrants to scientific invention are capitalizing on gains from trade to sell their inventions.

5.4 Instrumental variable estimation

Our results that science-based inventions have higher rates of trade, especially for novel inventions, smaller inventors, and heterogeneous markets, is consistent with the view that science lowers knowledge transfer costs and increases gains from trade. This relationship can be confounded by unobserved factors. For instance, an increase in the (unobserved) number of potential buyers would imply an increase in the observed reliance on science. In this section, we obtain causal estimates of the effect of reliance on science on MFT by instrumenting the use of science by changes in U.S. federal contract R&D caused by the end of the Cold War. In the first stage, we predict the share of patents citing science in an IPC-year using changes in federal contract R&D around the end of the Cold War. The predicted values are used in the second stage.

Table 9: POST COLD WAR FEDERAL R&D SHIFTS AND MFT (IV ESTIMATES)

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	1st Stage IV	2nd Stage IV	1st Stage IV	2nd Stage IV
Dependent Variable:	ln(Papers)	ln(Share of Reassigned Patents)	ln(Avg Cites to Science)	ln(Share of Reassigned Patents)	ln(Avg Cites to Science)	ln(Share of Reassigned Patents
Δ ln(Gov. R&D Contracting) (WOS Field)	1.257**					
	(0.019)					
ln(Gov. R&D Contracting (Pre, $1Bn)) (WOS Field)	0.364**					
	(0.003)					
ln(Avg Cites to Science)		1.040**		0.905**		0.666**
		(0.056)		(0.145)		(0.105)
Number of Papers (Predicted, 1000s)			0.079**			
			(0.003)			
Δ ln(Gov. R&D Contracting)					0.175**	
					(0.002)	
ln(Gov. R&D Contracting (Pre, $1Bn))		-0.476**	-0.007	-0.478**	0.081**	-0.481**
		(0.086)	(0.028)	(0.106)	(0.016)	(0.081)
ln(Gov. non-R&D Contracting (Pre, $1Bn))		0.314**	-0.059**	0.307**	-0.034**	0.296**
		(0.031)	(0.011)	(0.044)	(0.002)	(0.032)
ln(Number of Patents)		-0.042**	-0.003	-0.037*	0.021**	-0.027**
		(0.007)	(0.004)	(0.015)	(0.004)	(0.007)
Share of Small Assignees		0.396**	-0.245**	0.364**	-0.253**	0.308**
		(0.052)	(0.011)	(0.054)	(0.005)	(0.050)
Avg of DV	5.016	2.059	0.109	2.059	0.109	2.059
SD of Science		0.128		0.128		0.128
Cragg-Donald F-Stat			700.301		360.217	
Year Fixed Effects	Yes	Yes		Yes		Yes
R^2	0.227	0.119				
N	1,373	1,928	1,928	1,928	1,928	1,928

Notes: Analysis for column 1 is at the Web of Science Field-paper publication year level. Analyses for Column 2-6 are at the 4-digit IPC-patent publication year level. Sample period is 1992 and 2000 inclusively for all columns. *Avg Cites to Science* range from zero to one and averages the Cites Science Dummy at the unit of analysis. *Share of Reassigned Patents* ranges from zero to hundred and averages the Reassignment dummy at the unit of analysis, which at the patent level is equal to 100 if a patent is reassigned and zero otherwise. $\Delta log(Gov.\ R\&D\ Contracting)$ is calculated as the difference logged values of R&D contracting obligations between pre (1986-88) and post (1990-92) periods. $Number\ of\ Papers(1000s)$ refer to the number of papers (in thousands) relevant to a 4-digit IPC. $Number\ of\ Papers(Predicted, 1000s)$ is the predicted value from equation A4. *Gov. R&D Contracting (Pre)* is the average government contract R&D funding for the pre-period (1986-1988). *Gov. non-R&D Contracting (Pre)* is the average government non-R&D contract value for the pre-period (1986-1988). Column 1 includes paper publication year fixed effects, while columns 2-6 include patent grant year fixed effects. Standard errors are clustered at the year level.

Table 9 columns 3 and 4 present the results using the predicted paper instrument; columns 5 and 6 present results using the funding differences instrument. Consistent with our prediction of an upward bias, the OLS coefficient (column 2) is larger in magnitude than the second stage IV coefficients in columns 4 and 6. F-statistics for all first stage regressions are above 104.7, which recent work argues to be the appropriate critical F-value for valid inference using a second stage critical t-statistic of 1.96 (Lee et al., 2020).[41] Patent classes that cite science a standard deviation more due to federal research funding shocks experience a 5.6% increase in patent trade probability relative to the sample mean.

[41]Results are also robust to bootstrapped standard errors with 1,000 samples in appendix table B4

6 Discussion and Conclusion

This paper aims at advancing our understanding of how science affects the rate and direction of innovation. The use of science in invention can enhance the commercialization of inventions if it facilitates trade in inventions to those that are best able to commercialize them. Such trades support a division of labor between upstream inventors and downstream commercializers. Science generalizes phenomena into universal categories and unravels the mechanisms that underpin them. This may directly lead to higher quality inventions. Furthermore, conceptualizing inventions in scientific terms makes them easier to codify, reducing search costs for potential buyers, and enables buyers to evaluate and integrate inventions. This should reduce transaction costs that are thought to afflict trade in technology, as well as enhance the potential gains from trade.

Our main contribution is to establish that science-based inventions are more likely to be traded. Patents that reference a scientific article are 16-23% more likely to be traded than patents that do not reference science. Trade also depends on the demand for inventions. We derive and test three predictions of the reliance on science increasing gains from trade and reducing transaction costs. First, we find that patents invented by smaller firms are more likely to be traded, and that science magnifies this contrast with larger firms. Second, concentrated industries with an unequal distribution of potential patent buyers exhibit more patent trade, especially when the focal patent uses science. Third, the science citation effect on reassignment is up to three times larger for novel patents compared to not-novel patents. We also find that reliance upon science is associated with greater share of small inventors, and with entry of new inventors. Conditional upon citing science, we find that the positive relationship with MFT is especially strong for cited science that is more recent, suggesting that science also increases gains from trade, in addition to reducing transfer costs. Finally, we exploit a variation in the amount of scientific funding available from the federal government in the immediate aftermath of the Cold War to confirm the causal relationship between science and the market for technology.

Our findings imply that enhancing scientific understanding can increase social welfare over and above its role in generating valuable inventions: by encouraging the expansion of markets for technology, which allocates ownership rights to the most efficient user of existing inventions, and indirectly, by supporting a division of innovative labor.

References

Agrawal, A. and Henderson, R. (2002). Putting patents in context: Exploring knowledge transfer from MIT. Management science, 48(1):44–60. Publisher: INFORMS.

Ahmadpoor, M. and Jones, B. F. (2017). The dual frontier: Patented inventions and prior scientific advance. Science, 357(6351):583–587.

Anderson, S. P., De Palma, A., and Thisse, J.-F. (1992). Discrete choice theory of product differentiation. MIT press.

Anton, J. J. and Yao, D. A. (1994). Expropriation and inventions: Appropriable rents in the absence of property rights. The American Economic Review, pages 190–209.

Arora, A. (1995). Licensing tacit knowledge: intellectual property rights and the market for know-how. Economics of innovation and new technology, 4(1):41–60.

Arora, A., Belenzon, S., Cohen, W., and Lee, H. (2018). Are ideas getting harder to find, or have firms just become too big?

Arora, A., Belenzon, S., and Sheer, L. (2020). Back to Basics: Why do Firms Invest in Research? American Economic Review.

Arora, A. and Fosfuri, A. (2000). The market for technology in the chemical industry: causes and consequences. Revue d'économie industrielle, 92(2):317–334.

Arora, A. and Gambardella, A. (1994). The changing technology of technological change: general and abstract knowledge and the division of innovative labour. Research policy, 23(5):523–532.

Arque-Castells, P. and Spulber, D. F. (2017). The Market for Technology: Harnessing Creative Destruction.

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In The rate and direction of inventive activity: Economic and social factors, pages 609–626. Princeton University Press.

Arts, S., Hou, J., and Gomez, J. C. (2020). Natural language processing to identify the creation and impact of new technologies in patent text: Code, data, and new measures. Research Policy, page 104144.

AUTM, U. (2015). Licensing Activity Survey, FY2015. Association of University Technology Managers. AUTM.net, Accessed, 1.

Azoulay, P., Fons-Rosen, C., and Zivin, J. S. G. (2015). Does science advance one funeral at a time? Technical report, National Bureau of Economic Research.

Azoulay, P., Graff Zivin, J. S., Li, D., and Sampat, B. N. (2019). Public R&D Investments and Private-sector Patenting: Evidence from NIH Funding Rules. The Review of Economic Studies, 86(1):117–152.

Belenzon, S. and Schankerman, M. (2009). University knowledge transfer: private ownership, incentives, and local development objectives. The Journal of Law and Economics, 52(1):111–144.

Dernis, H. and Khan, M. (2004). Triadic patent families methodology. OECD Science, Technology and Industry Working Papers.

Figueroa, N. and Serrano, C. J. (2019). Patent trading flows of small and large firms. Research Policy, 48(7):1601–1616.

Fleming, L. (2001). Recombinant uncertainty in technological search. Management science, 47(1):117–132.

Fleming, L., Greene, H., Li, G., Marx, M., and Yao, D. (2019). Government-funded research increasingly fuels innovation. Science, 364(6446):1139.

Fleming, L. and Sorenson, O. (2004). Science as a map in technological search. Strategic Management Journal, 25(8-9):909–928.

Galasso, A. and Schankerman, M. (2014). Patents and cumulative innovation: Causal evidence from the courts. The Quarterly Journal of Economics, 130(1):317–369.

Galasso, A. and Schankerman, M. (2018). Patent rights, innovation, and firm exit. The RAND Journal of Economics, 49(1):64–86.

Galasso, A., Schankerman, M., and Serrano, C. J. (2013). Trading and enforcing patent rights. The

RAND Journal of Economics, 44(2):275–312.

Gans, J. S., Hsu, D. H., and Stern, S. (2002). When Does Start-Up Innovation Spur the Gale of Creative Destruction? The RAND Journal of Economics, 33(4):571–586.

Gans, J. S., Hsu, D. H., and Stern, S. (2008). The impact of uncertain intellectual property rights on the market for ideas: Evidence from patent grant delays. Management Science, 54(5):982–997.

Graham, S. J., Marco, A. C., and Myers, A. F. (2018). Patent transactions in the marketplace: Lessons from the uspto patent assignment dataset. Journal of Economics & Management Strategy, 27(3):343–371.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. Econometrica, Journal of the Econometric Society, pages 501–522. Publisher: JSTOR.

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? The Review of Economic Studies, 76(1):283–317.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological innovation, resource allocation, and growth. The Quarterly Journal of Economics, 132(2):665–712.

Kogut, B. and Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. Organization science, 3(3):383–397.

Kuhn, J. M. (2016). Property rights and frictions in the sale of patents.

Kuhn, J. M. and Thompson, N. C. (2019). How to measure and draw causal inferences with patent scope. International Journal of the Economics of Business, 26(1):5–38.

Lai, R., D'Amour, A., Yu, A., Sun, Y., and Fleming, L. (2011). Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975 - 2010). Edition: V5 Section: 2011-06-16 11:45:01.239.

Lee, D. L., McCrary, J., Moreira, M. J., and Porter, J. (2020). Valid t-ratio Inference for IV. arXiv preprint arXiv:2010.05058.

Levin, R. C., Klevorick, A. K., Nelson, R. R., Winter, S. G., Gilbert, R., and Griliches, Z. (1987). Appropriating the returns from industrial research and development. Brookings papers on economic activity, 1987(3):783–831.

Lichtenberg, F. R. (1987). The effect of government funding on private industrial research and development: a re-assessment. The Journal of industrial economics, pages 97–104. Publisher: JSTOR.

Ma, S., Wang, W., and others (2017). Selling innovation in bankruptcy.

Mansfield, E. (1980). Basic research and productivity increase in manufacturing. The American Economic Review, 70(5):863–873.

Marco, A. C., Myers, A. F., Graham, S. J., D'Agostino, P. A., and Apple, K. (2015). The USPTO patent assignment dataset: Descriptions and analysis.

Marco, A. D., Scellato, G., Ughetto, E., and Caviggioli, F. (2017). Global markets for technology: Evidence from patent transactions. Research Policy, 46(9):1644–1654.

Marx, M. (2019). Patent Citations to Science. type: dataset.

Marx, M. and Fuegi, A. (2020). Reliance on science: Worldwide front-page patent citations to scientific articles. Strategic Management Journal. Publisher: Wiley Online Library.

Merges, R. P. and Nelson, R. R. (1990). On the complex economics of patent scope. Columbia Law Review, 90(4):839–916.

Merrill, S. A. (2018). Righting the Research Imbalance. Technical report, The Center for Innovation Policy at Duke Law.

Mokyr, J. (2002). The gifts of Athena: Historical origins of the knowledge economy. Princeton University Press.

Moretti, E., Steinwender, C., and Van Reenen, J. (2019). The Intellectual Spoils of War? Defense R&D, Productivity and International Spillovers. Technical report, National Bureau of Economic Research.

Mowery, D. C. (1998). The changing structure of the US national innovation system: implications for international conflict and cooperation in R&D policy. Research Policy, 27(6):639–654. Publisher: Elsevier.

Narin, F., Hamilton, K. S., and Olivastro, D. (1997). The increasing linkage between US technology and public science. Research policy, 26(3):317–330.

Pece, C. (2016). Putting the Cart Before a Lame Horse: A case study for future initiatives to automate the use of administrative records for reporting government R&D. Technical report, National Center for Science and Engineering Statistics, National Science Foundation.

Png, I. P. (2019). US R&D, 1975–1998: A new dataset. Strategic Management Journal, 40(5):715–735. Publisher: Wiley Online Library.

Polanyi, M. (2015). Personal knowledge: Towards a post-critical philosophy. University of Chicago Press.

Roach, M. and Cohen, W. M. (2013). Lens or prism? Patent citations as a measure of knowledge flows from public research. Management Science, 59(2):504–525.

Sampat, B. (2010). Do Applicant Patent Citations Matter? Implications for the Presumption of Validity Christopher A. Cotropia 2 Mark Lemley 3. PhD Thesis, Columbia University.

Serrano, C. J. (2010). The dynamics of the transfer and renewal of patents. The RAND Journal of Economics, 41(4):686–708.

Serrano, C. J. (2018). Estimating the gains from trade in the market for patent rights. International Economic Review, 59(4):1877–1904.

Serrano, C. J., Holmes, T., Kortum, S., Aguirregabiria, V., Eckstein, Z., Figueroa, N., Ge, S., Gragera De Leon, F., Hopenhayn, H., Laincz, C., Lande, K., Mcmillan, R., Mitchell, M., Omer, V., Pakes, A., and Skrzypacz, A. (2015). Estimating the Gains from Trade in the Market for Patent Rights. pages 25–27.

Sethi, A. (2013). The Business of Electronics: A Concise History. Palgrave Macmillan US.

Stevenson, R. (2009). The LED's dark secret. IEEE Spectrum, 46(8).

Thurgood, L., Golladay, M. J., and Hill, S. T. (2006). US Doctorates in the 20th Century. Arlington, VA: National Science Foundation.

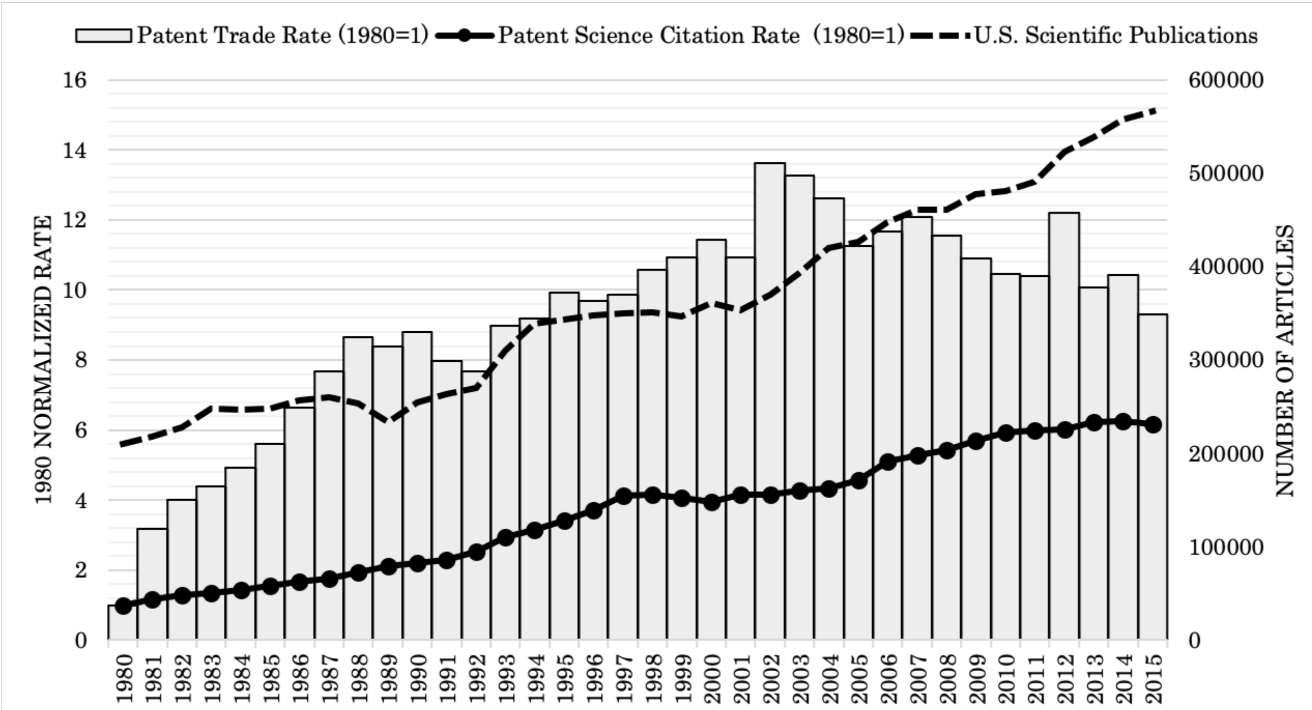
Veugelers, R. and Wang, J. (2019). Scientific novelty and technological impact. Research Policy, 48(6):1362–1372. Publisher: Elsevier.

Von Hippel, E. (1994). "Sticky information" and the locus of problem solving: implications for innovation. Management science, 40(4):429–439.

Ziedonis, R. (2004). Don't fence me in: Fragmented markets for technology and the patent acquisition strategies of firms. MANAGEMENT SCIENCE, 50(6):804–820.

Appendix A Data

Figure A1: Markets for Technology and Production and Use of Science



Notes: This graph plots time trends of patent trade, patent citations to science, and U.S. scientific publication output over our sample period. *Patent Trade Rate* for a given year is defined as the ratio between the number of patents reassigned over the number of patents in force in that year. *Patent Science Citation Rate* for a given year is defined as the number of citations to scientific articles in Microsoft Academic Graph (MAG) divided by the number of patents published in that year. Both rates are normalized by their levels in 1980 in this graph. *U.S. Scientific Publications* refer to the total number of scientific publications with a U.S. author in a given year from Clarivate Web of Science's Science Citation Index-Expanded (SCI-EXPANDED) and Conference Proceedings Citation Index-Science (CPCSI-S).

A.1 Sample construction

Our patent data is from the 2016 publication of PatStat and encompasses around 5.2 million utility patents granted by the USPTO from 1980 to 2016. We collect information on patent reassignment (transaction date, identity of buyers and sellers) by linking them to the USPTO Patent Assignment Database (PAD) (Graham et al., 2018), which records details on the transfer of ownership between patent assignees. To account for sample truncation, we limit our sample to patents granted on or before 2011 (for which we observe reassignments until 2015).[42] We construct measures of invention novelty based on textual similarity and on technological combinations, following Fleming (2001).

A.1.1 Patent Citations to Science

We employ a publicly available dataset from Marx and Fuegi (2020), which matches NPL citations to scientific articles available in Microsoft Academic Graph (MAG). The dataset assigns confidence scores for matches between a patent's NPL citation and a MAG article (1 being the lowest and 10 being the highest). We take the "PCS (Patent Citations to Science)" file and first exclude matches with under a 90% confidence score. We further exclude cited articles in the social sciences or humanities, leaving us

[42]About 58% of patents that are reassigned are done so within five years of being granted.

with OECD subject fields in "Natural Sciences", "Engineering and Technology", "Medical and Health Sciences", "Agricultural Sciences".

Textual similarity to science — We take the pairwise textual similarity measure from Arora et al. (2018), which calculates cosine similarities on the text of U.S. patents and scientific articles from the Science Citation Index - Expanded collection of Clarivate Web of Science. For each paper published between 1990 and 2015, we sort the patents in descending order of their similarity scores to the focal paper. We then rank the top 100 patents in terms of similarity scores to each publication ("similar patents"). The *Textual Similarity to Science* variable is equal to the natural log of one plus the number of publications for which a patent is classified as a "similar patent". In unreported robustness checks, we calculate two additional metrics. First, we normalize the the similarity scores by the maximum similarity score pair for a publication, and take the average of this score for each patent. Second, we take the average of the similarity rank each patent receives with respect to each publication. The direction and statistical significance of the column 6 result in table 3 are not sensitive to these alternative calculations of patent textual similarity to science.

Journal Impact Factor — Journal impact factor for a journal in year t is calculated as the number of forward citations in years $t-1$ and $t-2$ received by the number of papers published in years $t-1$ and $t-2$ by the focal journal. The average journal impact factor for patent i averages this value for all articles cited by a patent.[43]

Citation Lag — We measure how recent the science being cited is in relation to a patent by measuring the average year difference between the grant year of the patent and the publication year of the paper. For patent i citing $j \in J_i$ articles,

$$Avg(Lag\ to\ Cited\ Science_i) = \frac{\sum_{j \in J_i} Grant\ Year_i - Publication\ Year_j}{|J_i|}$$

The lower this value, the more recent ("younger") the science being used in relation to the patent.

Number of Fields — we calculate how specialized a patent's scientific citations are by counting the number of unique WoS Fields of cited scientific papers per patent and dividing it by the number of cited papers (the measure ranges between 0 and 1). For patent i citing J articles published in K fields,

$$Normalized\ Field\ Counts_i = \frac{|K_i|}{|J_i|}$$

The lower (higher) this value, the more specialized (interdisciplinary) the patent is in terms of its scientific citations.

Scientific Combination Familiarity — we calculate the novelty of the combination of Web of Science scientific fields cited by counting how many times the same scientific combinations have been cited by the paper cited by a focal paper since 1790.

$$\begin{aligned} &Combination\ Familiarity\ for\ WOS\ Fields\ (Decayed)_i = \\ &\qquad \sum_{patents\ k\ before\ patent\ i} 1\{k\ cites\ identical\ combination\ of\ WOS\ Fields\ as\ i\} \\ &\qquad \times exp(\frac{grant\ date\ of\ patent\ k\ -\ grant\ date\ of\ patent\ i}{time\ constant\ of\ knowledge\ loss}) \end{aligned}$$

Where the *time constant of knowledge loss* is set to 5 years such that a previous Web of Science combination from five years ago is weighted by $exp(-\frac{1}{5}) = 37\%$. This is an analog of the Technological Combination Familiarity measure by Fleming (2001), which is calculated for patent classes. There are 175 Web of Science (WoS) Fields assigned to 1,740,815 articles cited by patents in our sample. Intuitively,

[43] We source journal impact factors from Marx and Fuegi (2020)

the more often the same combination appears (the higher the *Familiarity* score), the less novel are the patent's scientific combinations.

A.1.2 Identifying market transactions for patents

We download the 2016 version of the USPTO Patent Assignment Dataset and identify patent reassignments that may classify MFT transactions. Our framework follows methods pioneered by Serrano (2010) and refined by Ma et al. (2017) and Figueroa and Serrano (2019).

We define MFT transactions as transfers of technology between two independent entities. This excludes ownership transfers within firms and purchases of capabilities rather than technology (e.g. M&As that transfer lab personnel and capital equipment along with patents). The USPTO records each received patent transfer in a "Reel Frame" (RF) ID, and has classified the conveyance types of these transfers into assignment of assignor's interest, name changes, government interest agreements, security agreements, and release by secured parties. We exclude all other conveyance types than assignments of assignors' interest. The USPTO also identifies employer assignment as the first recorded transaction for a patent where the patent is transferred alone with an execution date prior to the patent application disposal date (Graham et al., 2018, p.27). These RF IDs are also removed.

We add several additional checks. First, we exclude assignments whose date is before the grant date of a patent. While it is possible that a transaction has occurred before the patent was granted, it is also possible that the patent's initial assignment was mistaken with a reassignment to a buyer. Without a way to positively identify pre-grant patent application purchases, we decide it is safer to exclude these cases to reduce false positives.

Second, we exclude cases where the assignee ("buyer") names in the PAD records are similar to assignee names in the USPTO PATSVIEW. The assignee names in PATSVIEW record the initial assignee name(s) on the granted patent document. Therefore, if the assignee name in the PAD records are similar to the original owner's (assignee on patent document), we can rule out an MFT transfer between two independent entities.

Third, we exclude cases where the assignor ("seller") of an assignment is similar to the inventor of the patent from USPTO PATSVIEW. These cases are likely to be corporate employees transferring their patent rights to their firms per terms in their employment contract (it has been common practice among large corporations such as Du Pont, IBM, and Google to automatically transfer patent rights from employees to employers by such contracts).

Fourth, we download all completed acquisitions recorded in SDC Platinum between 1980 and 2015 and match the "Target Name" and "Acquiror Name" in SDC to patent assignor and assignee names in PAD. If the buyer-seller pair of companies in SDC correspond to the buyer-seller pairs in PAD, we exclude them.

Fifth, we also measure the string distance between assignor-assignee pairs so that intra-corporate reassignments (from, say, a company's headquarters to its licensing subsidiary) are dropped. For the second to fifth steps, we judge that names are similar based on Jaro-Winkler, Jaccard, and a normalized Levehnstein edit distance (python package available from https://github.com/seatgeek/fuzzywuzzy) after standardizing common suffixes such as "CORP", "LTD" and prefixes such as "LEGAL REPRESENTATIVE". Specifically, we take one minus the maximum value of the distance measures (which range between zero and one) and classify those pairs with larger than an appropriate threshold as similar to each other. We conduct extensive human checks around these thresholds to reduce classification error.

Sixth, we exclude RF IDs with more than 25 patents being transferred, because these are likely to be part of M&A deals between large firms.

A.1.3 Technological sector classifications, By 2 digit IPCs

Table A1: TECHNOLOGICAL SECTOR CLASSIFICATIONS, BY 2 DIGIT IPCS)

	Technological Sector				
2digit IPCs	Life Sciences	Chemicals	ICT	others	Total
A0	0	21,088	0	33,988	55,076
A2	0	19,613	0	8,672	28,285
A4	0	0	0	80,474	80,474
A6	260,352	13,368	0	55,149	328,869
B0	0	97,147	0	15,620	112,767
B2	0	0	0	156,854	156,854
B3	0	19,394	0	5,506	24,900
B4	0	0	0	72,225	72,225
B6	0	0	0	260,836	260,836
B8	0	0	0	1,486	1,486
C0	15,111	229,919	0	59,285	304,315
C1	53,768	29,920	0	43	83,731
C2	0	28,317	0	16,968	45,285
C3	0	5,038	0	0	5,038
C4	0	625	0	0	625
D0	0	1,971	0	27,061	29,032
D2	0	0	0	9,512	9,512
E0	0	0	0	68,229	68,229
E2	0	0	0	29,611	29,611
F0	0	0	0	104,165	104,165
F1	0	0	0	110,113	110,113
F2	0	4,985	0	66,463	71,448
F4	0	0	0	15,259	15,259
G0	26,549	0	489,134	273,734	789,417
G1	0	0	51,524	8,890	60,414
G2	0	0	0	9,437	9,437
H0	1,511	2,592	480,584	214,912	699,599
Total	357,291	473,977	1,021,242	1,704,492	3,557,002

Notes: This table tabulates the four technological sector classifications (Life Sciences, Chemicals, ICT, and other) that we use in our main sample of 3.5 million USPTO patents published between 1980 and 2011.

A.2 Instrumental variable construction

A.2.1 Data collection

We collect federal procurement contracts for research and development services by manually downloading government contracts by year and agency from `https://www.usaspending.gov` and `https://beta.SAM.gov`. The former is mandated by the Federal Funding Accountability and Transparency Act of 2006 and is maintained by the Office of Management and Budget and includes all procurement activities of the U.S. federal government since 2000. The latter is run by the General Services Administration and contains procurement data as early as the 1970s. We keep data from 1980 onwards for our analysis.

Table A2: Government Funding of Procurement Contracts for R&D Servicese

		Contract Value		
Rank	PSC	1986-88	1990-1992	Description
1	AN41	$ 249,686	$ 99,000,000	R&D- Medical: Health Services (Basic Research)
2	AH96	$ 143,921	$ 25,000,000	R&D- Environmental Protection: Other (Management/Support)
3	AE21	$ 253,278	$ 27,800,000	R&D- Economic Growth: Product/Service Improvement (Basic Research)
4	AN12	$ 4,530,456	$ 315,000,000	R&D- Medical: Biomedical (Applied Research/Exploratory Development)
5	AN11	$ 26,600,000	$ 1,420,000,000	R&D- Medical: Biomedical (Basic Research)
6	AN46	$ 47,000,000	$ 1,970,000,000	R&D- Medical: Health Services (Management/Support)
7	AG94	$ 851,749	$ 32,600,000	R&D- Energy: Other (Engineering Development)
8	AN15	$ 313,241	$ 9,018,358	R&D- Medical: Biomedical (Operational Systems Development)
9	AE33	$ 2,653,725	$ 63,200,000	R&D- Economic Growth: Manufacturing Technology (Advanced Development)
10	AE35	$ 17,300,000	$ 386,000,000	R&D- Economic Growth: Manufacturing Technology (Operational Systems Development)
...	...	...	...	...
212	AR12	$ 11,200,000	$ 553,805	R&D- Space: Aeronautics/Space Technology (Applied Research/Exploratory Development)
213	AS21	$ 1,865,031	$ 65,330	R&D- Modal Transportation: Surface Motor Vehicles (Basic Research)
214	AH32	$ 2,703,422	$ 69,106	R&D- Environmental Protection: Water Pollution (Applied Research/Exploratory Development)
215	AR22	$ 54,400,000	$ 1,283,886	R&D- Space: Science/Applications (Applied Research/Exploratory Development)
216	AG73	$ 3,824,674	$ 54,971	R&D- Energy: Solar/Photovoltaic (Advanced Development)
217	AG55	$ 17,900,000	$ 144,088	R&D- Energy: Nuclear (Operational Systems Development)
218	AN40	$ 10,600,000	$ 62,824	R&D- Health Services
219	AD54	$ 50,600,000	$ 189,912	R&D- Defense Other: Fuels/Lubricants (Engineering Development)
220	AR94	$ 881,000,000	$ 775,323	R&D- Space: Other (Engineering Development)
221	AZ10	$ 368,000,000	$ 39,265	R&D- Other

Notes: The observations are sorted in descending order by the logged difference between the pre (1986-88) and post (1990-92) period.

Our procurement data covers all contracts signed by the Department of Defense, Energy, Health and Human Services, and Veterans Affairs between 1980 and 2020. As of FY2019, these agencies accounted for more than 72% of all procurement contracts and constitute four of the five largest spenders on "contractual services and supplies" (the omitted agency is the Office of Personnel Management, which primarily deals with health benefits and life insurance funds and unlikely to contract out R&D services). We collect the signing date, action obligation[44] in current dollars, vendor names, contracting agency and the relevant 4-digit Product and Service Codes (PSC) for these four agencies between 1980 and 2019. We match the vendor names to the firm names found in the DISCERN database of American public, R&D performing firms between 1980 and 2015, using a combination of automated string distance metrics

[44] Action obligations are "intentions" backed by a contractual agreement. The government does not release actual dollar amounts spent on contracts. After initial contract signing, the actual expenditure can increase, decrease, or stay the same. Agencies enter a "contract action" into the database whenever they know that what was initially obligated has changed. These corrections may sometimes (5%) lead to negative obligations.

and manual cleaning. We limit procurement contracts to services related to R&D only (1st digit PSC corresponding to "A").

A.2.2 Crosswalk definition

We crosswalk the values of the contracts from PSC codes to patent classes (IPCs) and publication fields (WoS fields) using the following methods:

Crosswalk from PSC to IPC — We calculate the level of federal scientific spending relevant to an IPC by multiplying a dyadic weight based on patent class-level patenting distributions of vendor firms winning contracts in a 4-digit PSC. For firm i patenting in IPC k and contracting in product code j, the 4-digit PSC-to-4-digit IPC weight is defined as:

$$weight_{jk} := \sum_i Contract\_Value_{ij} \times \frac{patents_{ik}}{patents_i} \tag{A1}$$

for all DISCERN patents and contracts published and signed between 1980 and 1992. The post Cold War funding shock relevant for each 4-digit IPC (used in the first stage regression of table 9 column 5 and appendix table A3) is calculated as the logged difference between the average contract value for a "pre" period from 1986 to 1988 and a "post" period from 1990 to 1992.

$$\Delta ln(Gov.\ R\&D\ Contracting_k) := ln\left(\frac{\sum_j weight_{jk} \times \frac{\sum_{t=1990}^{t=1992} Contract\_Value_{jt}}{3}}{\sum_j weight_{jk} \times \frac{\sum_{t=1986}^{t=1988} Contract\_Value_{jt}}{3}}\right) \tag{A2}$$

Crosswalk from PSC to WoS Field — We calculate the level of federal scientific funding relevant to a Web of Science field by multiplying a dyadic weight based on scientific field-level publication distributions of contracting vendor firms. For firm i publishing in WoS field l and contracting in product code j, the 4-digit PSC-to-WoS field weight is defined as:

$$weight_{jl} := \sum_i Contract\_Value_{ij} \times \frac{papers_{il}}{papers_i} \tag{A3}$$

for all DISCERN papers and contracts authored and signed between 1980 and 1992.

Number of Predicted Papers — Using the above crosswalk, we run an OLS regression to predict the number of papers as a function of the funding shocks around the end of the Cold War:

$$\begin{aligned} Number\ of\ Papers_{lt} = \beta_0 + \beta_1 \Delta ln(Government\ R\&D\ Contracting)_l \\ + \beta_2 ln(Government\ R\&D\ Contracting\ (Pre))_l + \boldsymbol{\xi_t} + \nu_{lt} \end{aligned} \tag{A4}$$

where $\boldsymbol{\xi_t}$ are paper publication year fixed effects. The WoS field level funding shock in equation A4 is defined as:

$$\Delta ln(Gov.\ R\&D\ Contracting_l) := ln\left(\frac{\sum_j weight_{jl} \times \frac{\sum_{t=1990}^{t=1992} Contract\_Value_{jt}}{3}}{\sum_j weight_{jl} \times \frac{\sum_{t=1986}^{t=1988} Contract\_Value_{jt}}{3}}\right) \tag{A5}$$

Crosswalk from WoS Field to IPC — To construct the instrument for columns 3 and 4 of table 9, we crosswalk the number of predicted papers ($\widehat{Number\ of\ Papers_{lt}}$) at the web of science-year level to the patent class-year level by multiplying a dyadic weight based on the NPL citations from patents the scientific literature. For patents in 4-digit IPC k citing papers in WoS field l, the number of scientific papers relevant for each IPC is defined as:

$$Number\ of\ Papers_{kt} = \sum_{l} weight_{kl} \times Number\ of\ Papers_{lt} \tag{A6}$$

where the weight is defined as:

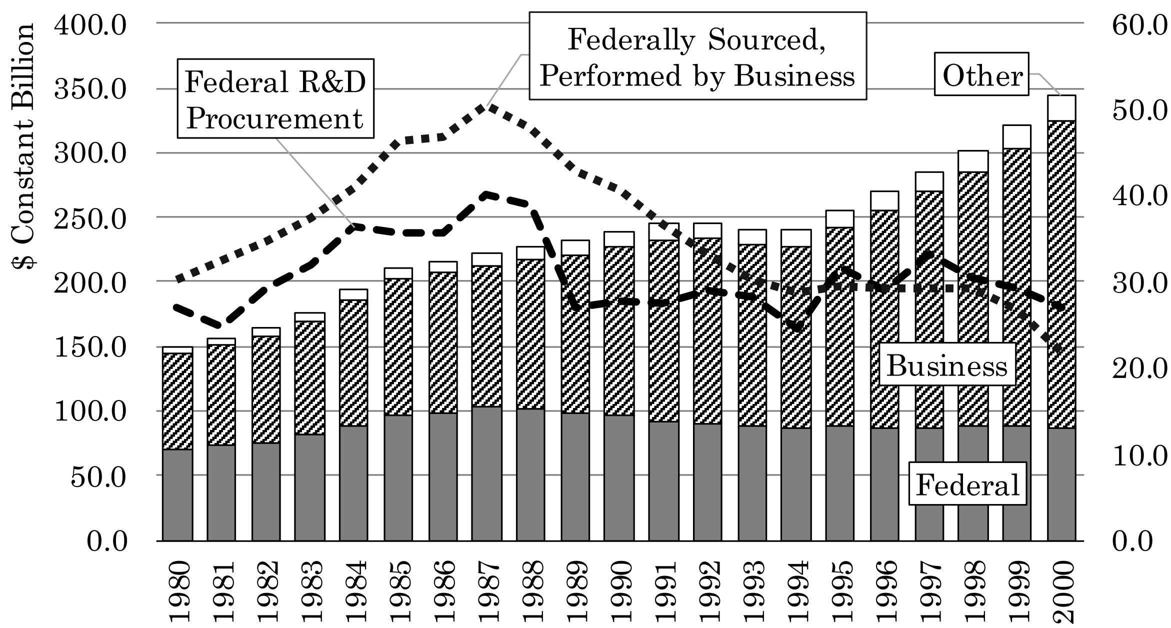
$$weight_{kl} = \frac{NPL\ Citations_{kl}}{Total\ NPL\ Citations\ Received_k} \tag{A7}$$

for all patents granted between 1980 and 1992.

A.2.3 Comparison to other sources of R&D expenditure data

We use the National Science Foundation's "National Patterns of R&D" data series to compare the relative magnitude of federal R&D funding to other sources and verify that the procurement data we use for the construction of the instrument. Figure A2 shows that the federal government was responsible for funding around 43% of all R&D expenses between 1986 and 1992. Moreover, the R&D procurement data we are using for the instrumental variable covers around 76% of total federal R&D spending performed by industry.

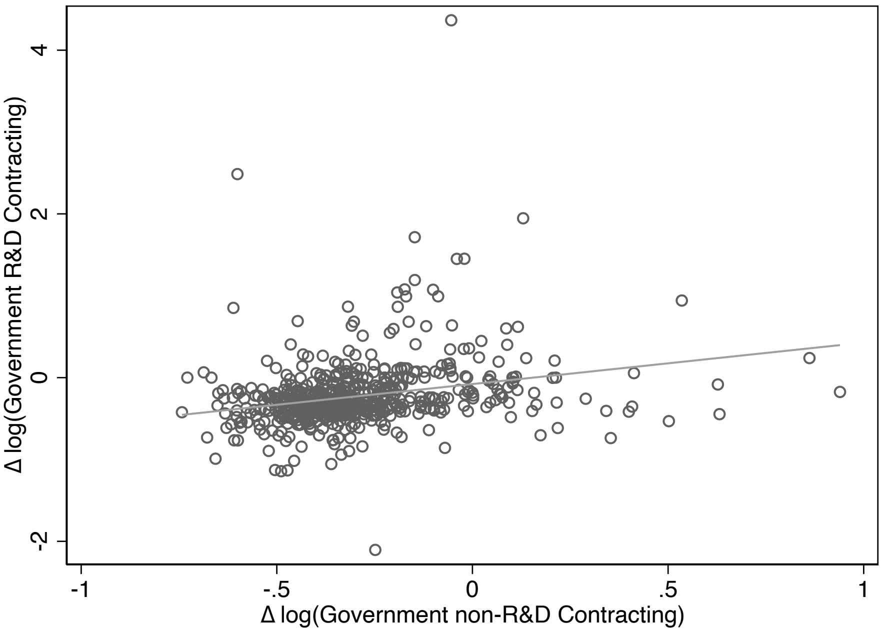
Figure A2: U.S. R&D Funding, by Source



Notes: The bar graph plots the aggregated annual research and development expenditure by source of funds from the *NSF National Patterns of R&D Resources (2014-15)*, tables 8 and 9 (the "Other" category aggregates non-federal government, higher education, and other non-profits). The line "Federally Sourced, Performed by Business" is a subset of the federal spending on R&D that is performed by the business sector (the others performing sectors are federal intramural, FFRDC, non-federal government, higher education, and other non profits). The line "Federal R&D Procurement" plots the Federal R&D procurement dollars used in the construction of the instrumental variable in section 5.4. Figures are adjusted to 2012 dollars using GDP deflators from Louis Johnston and Samuel H. Williamson, "What Was the U.S. GDP Then?" MeasuringWorth, 2020.

There are three likely sources of this discrepancy. First, we collect data for the four largest spenders on R&D, while the NSF data covers all federal agencies. Second, reporting behaviors between the procurement data and the survey data tend to differ (Pece, 2016). That is, even for the same agency and subcategory of spending, there are significant gaps in the dollar values reported between the procurement data and the NSF survey data. In FY2016, for instance, the Department of Defense reported $24.6 billion

Figure A3: Funding Shocks for R&D vs non-R&D Procurements



Notes: This plots the difference in federal funding before (1986-88) and after (1990-92) the fall of the Berlin Wall at the 4 digit IPC for R&D funding (y-axis) against non-R&D funding (x-axis) (r=0.257).

on procurement contract obligations for R&D services. In addition, the agency reported approximately $6 billion in grants (for all types of grants, including but not limited to R&D grants). Therefore, we expect to see a figure of under $31 billion for the DoD in the Federal Funds for R&D Survey data for FY2016. However, the FFS reports $42 billion, which results in a difference of $11 billion. It is possible that some data was omitted because the contracted amounts were below the reporting thresholds (ranging between $2,000 and $10,000 depending on item specific requirements), or due to national security concerns. Pece (2016) also points out that discrepancies between accounting systems maintained by the respective agencies make collection of consistent data difficult.

Third, the NSF survey data includes grants, which we have not collected. The NSF "National Patterns" report relies on the Survey of Federal Funds for Research and Development (FFS) and the Census Bureau's Business R&D and Innovation Survey (BRDIS). The questionnaires in these surveys collect data on not only contracts but also grants.[45]

We address the concern that this omission may systematically undercount certain scientific disciplines. For instance, the life sciences relies heavily on grants from the National Institutes of Health (NIH). We therefore compare the R&D procurement spending for the life sciences against the R&D outlays for the Department of Health and Human Services (under which the NIH is classified). The NSF data indicates a 35% increase in real dollar terms (from $9.9 to $13.5 billion 2012 dollars). Our data on the HHS also shows a 5.7 times increase for the same period, while R&D contracts for medical (2 digit PSC code: "AN") purposes increase 5.3 times. Therefore, the direction of the change is same, if the magnitudes are different.

A.3 IVE robustness check

We calculate a new instrument that calculates funding differences after controlling for differences in MFT demand conditions. This instrument predicts funding shocks net of controls for patenting quantity, quality, and propensity to trade from an OLS specification. We estimate the following OLS specification for 4-digit IPC k and = year t for years 1986 through 1992:

$$\begin{aligned} Government\ R\&D\ Contracting_{kt} &= \beta_0 + \beta_1 Post1989_t + \beta_2 Post1989_t \times IPC\_Dummy_i \\ &+ IPC\_Dummy_k + Year\_Dummy_t + \boldsymbol{Z'_{kt}} + \ \epsilon_{kt} \end{aligned} \tag{A8}$$

where controls $\boldsymbol{Z'_{it}}$ consist of number of patents, forward patent citations, and share of patents traded. *Government R&D Contracting*$_{kt}$ is the amount of government R&D funding relevant to each 4-digit IPC-year.[46] For each 4-digit IPC, the logged difference in predicted government funding due to the end of the Cold War net of the controls is $\Delta log(\widehat{Government\ R\&D\ Contracting}) := log(\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 IPC\_Dummy_i + IPC\_Dummy_i) - log(\hat{\beta}_0 + IPC\_Dummy_i)$. This is the instrument used in columns 3 and 4 of table A3.

Table A3: Post Cold War Federal R&D Shifts and MFT (Predicted Funding Shocks)

	(1)	(2)	(3)
	OLS	1st Stage IV	2nd Stage IV
Dependent Variable:	ln(Share of Reassigned Patents)	ln(Avg Cites to Science)	ln(Share of Reassigned Patents)
ln(Avg Cites to Science)	1.040**		0.650**
	(0.056)		(0.085)
Δ ln(Gov. R&D Contracting) (Predicted)		0.306**	
		(0.005)	
ln(Gov. R&D Contracting (Pre, $1Bn))	-0.476**	0.122**	-0.481**
	(0.086)	(0.016)	(0.081)
ln(Gov. non-R&D Contracting (Pre, $1Bn))	0.314**	0.008*	0.295**
	(0.031)	(0.004)	(0.032)
ln(Number of Patents)	-0.042**	-0.003	-0.027**
	(0.007)	(0.006)	(0.006)
Share of Small Assignees	0.396**	-0.276**	0.305**
	(0.052)	(0.004)	(0.054)
Avg of DV	2.059	0.109	2.059
SD of Science	0.128		0.128
Cragg-Donald F-Stat		321.527	
Year Fixed Effects	Yes		Yes
R^2	0.119		
N	1,928	1,928	1,928

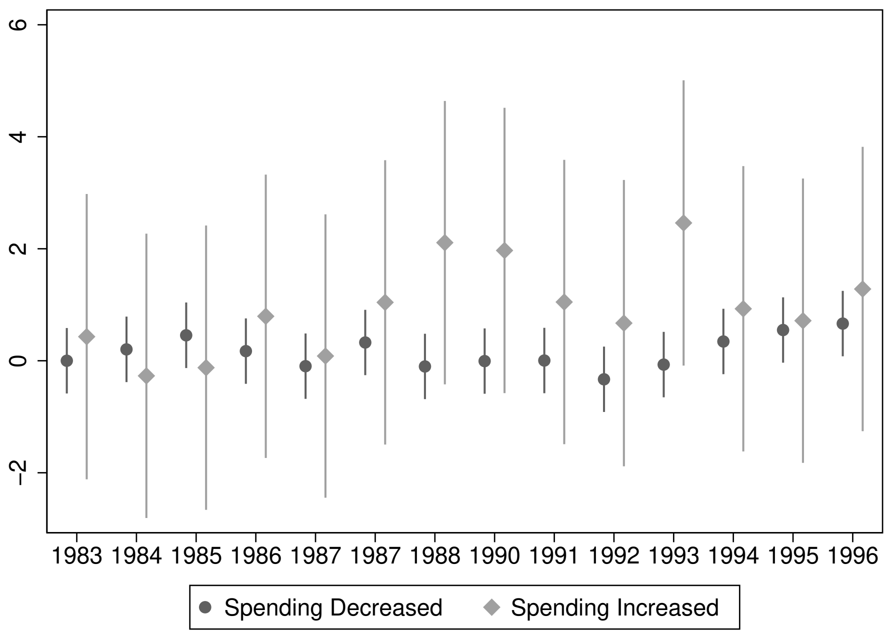
Notes: Analysis is at the 4-digit IPC-patent publication year level. Sample period is 1992 and 2000 inclusively for all columns. $\Delta log(Gov.\ R\&D\ Contracting)(Predicted)$ is calculated as the logged difference in predicted values from an OLS specification in equation A8. Other variable definitions are identical to table 9. All specifications include patent publication year fixed effects. Standard errors are clustered at the year level.

A potential threat to the validity of table 9 in section 5.4 is that our results are biased by secular patterns in MFT preceding our sample period between 1992 and 2000. For instance, it is possible

[45] The 2001 questionnaire for the FFS states that "A performer [of R&D] is either an intramural group or organization carrying out an operational function or an extramural organization or person receiving support or providing services under a contract **or grant**." (brackets and emphasis added by authors)

[46] The funding data is aggregated at the PSC-year level and crosswalked to the 4-digit IPC year level. The weights are based on the relevance of each 4-digit IPC to a PSC using patenting data of vendor names matched to the DISCERN database.

Figure A4: Share of Patents Traded vs Shock



Notes: This plots the coefficients from regressing share of patents traded within 5 years of grant at the 4 digit IPC-year level against patent publication years, splitting the sample by whether the focal IPC was negatively (dark circle marker) or positively (light diamond marker) shocked by the end of the Cold War. The base group is 1982.

that areas where federal R&D spending increased were also areas where patent market regulations were selectively relaxed or random scientific discoveries were concentrated in. If so, then we should expect there to be a considerable difference in the level of MFT activity for IPCs whose federal R&D spending increased and those whose spending decreased. We regress at the 4-digit IPC-year level the share of traded patents against year dummies after splitting the sample where average Federal R&D spending increased and those where spending decreased. Figure A4 plots the coefficients of these OLS specifications (where 1982 is the base year) and finds that there is no statistically significant difference between the groups positively and negatively affected by federal spending shocks.

A.3.1 R&D contracting and science

We find in column 1 of table 9 that a standard deviation larger government contract R&D shock around the end of the Cold War leads to a 7% increase in papers in the relevant Web of Science field.

Given that corporate share in total scientific publications is generally low compared to academia,[47] the existence of the 7% effect may be suspect if one assumes that government contract R&D procurement is principally carried out by the corporate sector. Nevertheless, we find that between 1986 and 1992 (when the funding shocks are calculated), U.S. public firms account for only around 22% ($6.7 Bn) of all contract R&D (which averages $32 Bn) in our procurement data. Moreover, around a third (114/357) of R&D contracting firms never publish in science, while only around 40% operate a corporate lab for 1988 and 1991.[48] This suggests that a sizable share of the contract R&D work is sub-contracted to public

[47]For example, the NSF's 2018 Science and Engineering Indicators produces well over 70% of all peer-reviewed scientific papers between 2003 and 2016 (`https://www.nsf.gov/statistics/2018/nsb20181/report/sections/academic-research-and-development/outputs-of-s-e-research-publications`)

[48]We link our procurement data to a comprehensive directory of all corporate industrial laboratories in the United States from Png (2019). The dataset is available at `https://scholarbank.nus.edu.sg/handle/10635/150104` and contains the number of professionals, doctorates, and technicians reported by American firms for 1981, 1983, 1985, 1988, 1991, 1994,

entities that may in turn publish follow-on research.

This conjecture is consistent with prior research that argues that American firms started to "externalize" their R&D operations "through such mechanisms as consortia, collaboration with US universities and federal laboratories" since the beginning of the 1980s (Mowery, 1998, p.646). On the other hand, universities began to rely more heavily on industry support: between 1960 and 1995, industry contribution to university research tripled to 7%, while more than 1050 research institutes "seeking to support research on issues of direct interest to industry" were being run by 1992 (Mowery, 1998, p.648).[49]

1997. The dataset also links lab names to Compustat GVKEYs, which allows us to link the data to the procurement data.

[49] 57% of these were established during the 1980s.

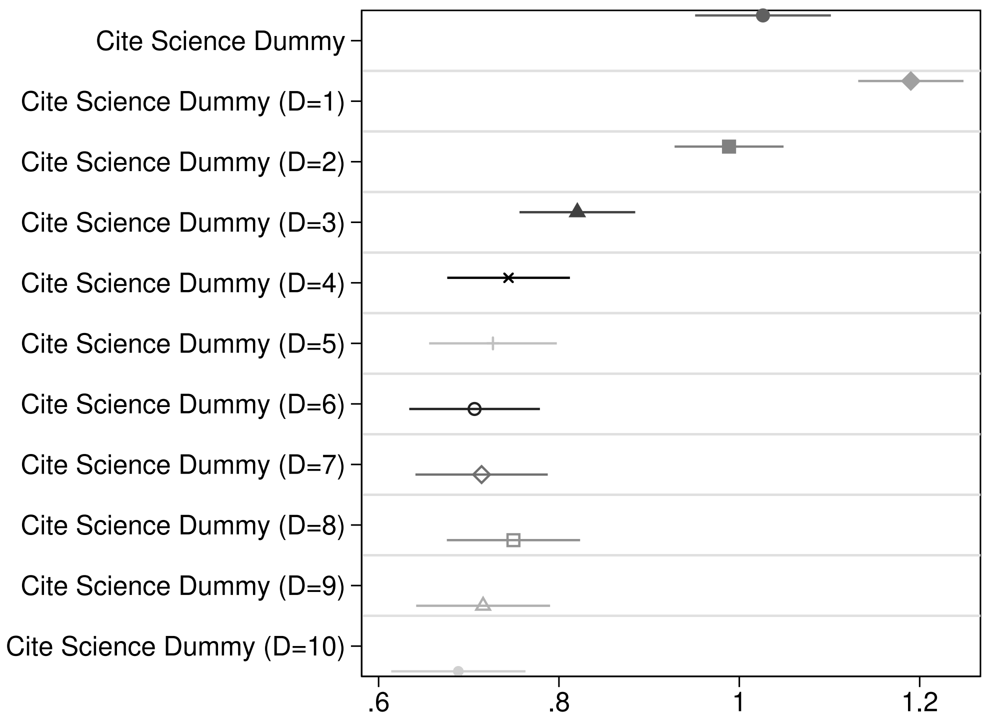
Appendix B Auxiliary Results

Table B1: Reliance on Science and MFT, by Degrees of Connection to Citation Frontier)

	Dependent Variable: Reassignment=1										
	(1) D=0	(2) D=1	(3) D=2	(4) D=3	(5) D=4	(6) D=5	(7) D=6	(8) D=7	(9) D=8	(10) D=9	(11) D=10
Cite Science Dummy	1.026**										
	(0.038)										
Cite Science Dummy (D=1)		1.190**									
		(0.030)									
Cite Science Dummy (D=2)			0.989**								
			(0.031)								
Cite Science Dummy (D=3)				0.820**							
				(0.033)							
Cite Science Dummy (D=4)					0.744**						
					(0.035)						
Cite Science Dummy (D=5)						0.727**					
						(0.036)					
Cite Science Dummy (D=6)							0.706**				
							(0.037)				
Cite Science Dummy (D=7)								0.714**			
								(0.037)			
Cite Science Dummy (D=8)									0.750**		
									(0.038)		
Cite Science Dummy (D=9)										0.716**	
										(0.038)	
Cite Science Dummy (D=10)											0.688**
											(0.038)
Triadic Patent Dummy	0.914**	0.912**	0.932**	0.946**	0.954**	0.957**	0.960**	0.962**	0.964**	0.967**	0.969**
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Number of Claims	0.082**	0.081**	0.082**	0.083**	0.083**	0.083**	0.084**	0.084**	0.083**	0.083**	0.083**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Length of First Claim	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Avg of DV	6.283	6.283	6.283	6.283	6.283	6.283	6.283	6.283	6.283	6.283	6.283
Science Citing Patents	695,164	1,708,136	2,304,626	2,568,996	2,635,670	2,603,063	2,526,127	2,428,737	2,322,123	2,208,138	2,090,316
4-digit IPC Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
N	3,882,632	3,882,632	3,882,632	3,882,632	3,882,632	3,882,632	3,882,632	3,882,632	3,882,632	3,882,632	3,882,632

Notes: *Reassignment* is a binary variable equal to 100 if the patent has ever been reassigned in the USPTO PAD dataset. *Cite Science* is equal to one if there has been a citation to Microsoft Academic Graph (MAG), and zero otherwise. *Cite Science Dummy (D=n* is equal to one if there has been a citation to MAG n times removed from the citation frontier. Other variable definitions are identical to those in table 3

Figure B1: Reliance on Science and MFT, by Degrees of Connection to Citation Frontier



Notes: This figure plots the coefficient estimates for the Cite Science Dummies in Table B1, in increasing order of distance from the citation frontier.

Table B2: Reliance on Science and MFT, Forward Patent Citation Controls)

	DV: Reassignment=1		
	(1) Continuous	(2) Quintiles	(3) Deciles
Cite Science Dummy	0.780**	0.825**	0.802**
	(0.039)	(0.038)	(0.038)
5-year Forward Patent Cites	0.104**		
	(0.002)		
Triadic Patent Dummy	0.810**	0.820**	0.809**
	(0.029)	(0.029)	(0.029)
Number of Claims	0.069**	0.070**	0.069**
	(0.001)	(0.001)	(0.001)
Length of First Claim	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)
Forward Patent Citations (1st Quintile)		0.000	
		(.)	
Forward Patent Citations (2nd Quintile)		0.302**	
		(0.035)	
Forward Patent Citations (3rd Quintile)		0.673**	
		(0.036)	
Forward Patent Citations (4th Quintile)		1.293**	
		(0.037)	
Forward Patent Citations (5th Quintile)		2.593**	
		(0.040)	
Forward Patent Citations (1st Decile)			0.000
			(.)
Forward Patent Citations (2nd Decile)			-0.098
			(0.052)
Forward Patent Citations (3rd Decile)			0.175**
			(0.046)
Forward Patent Citations (4th Decile)			0.385**
			(0.046)
Forward Patent Citations (5th Decile)			0.546**
			(0.047)
Forward Patent Citations (6th Decile)			0.757**
			(0.048)
Forward Patent Citations (7th Decile)			1.131**
			(0.049)
Forward Patent Citations (8th Decile)			1.414**
			(0.050)
Forward Patent Citations (9th Decile)			1.900**
			(0.051)
Forward Patent Citations (10th Decile)			3.262**
			(0.055)
Avg of DV	6.265	6.265	6.265
4-digit IPC Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
R^2	0.013	0.013	0.013
N	3,882,632	3,882,632	3,882,632

Notes: Unit of analysis is at the patent level. “5-year Forward Patent Cites” counts the number of forward patent citations a focal patent receives. Column 2 and 3 respectively include dummies for quintiles and deciles of “5-year Forward Patent Cites” by patent grant year and 4-digit IPC. Other variable definitions are identical to those in table 3

Table B3: Science and MFT, by characteristic of cited science (2nd degree connected)

	DV: Reassignment = 1			
	(1) Recent	(2) Specialized	(3) Novel	(4) All
Avg(Lag to Cited Science (D=2))	-0.004			-0.004
	(0.004)			(0.004)
1-Normalized Field Counts (D=2)		0.292*		0.403**
		(0.140)		(0.144)
log(Avg(MAG Combination Familiarity (D=2)) + 1)			-0.020	-0.024*
			(0.011)	(0.011)
Triadic Patent Dummy	1.932**	1.927**	1.925**	1.923**
	(0.068)	(0.068)	(0.068)	(0.068)
Number of Claims	0.063**	0.062**	0.062**	0.062**
	(0.002)	(0.002)	(0.002)	(0.002)
Length of First Claim	-0.000*	-0.000*	-0.000*	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Avg of DV	7.466	7.449	7.449	7.449
IPC Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
R^2	0.015	0.015	0.015	0.015
N	685,836	685,836	679,822	679,822

Notes:Sample is limited to patents that cite at least one scientific article in MAG. *Lag to Cited Science (D=2)* is defined as the difference in the grant year of a patent and the publication year of a scientific paper cited by a scientific paper in its front page NPL citation list ("D=2" cited papers). *Avg(Lag to Cited Science)* averages this value for each sciece-citing patent. *MAG Combination Familiarity (D=2)* calculates the number of times the same WOS Field combination has been cited by a paper since 1790 with an exponential time decay rate of 18% (Fleming, 2001). *Avg(MAG Combination Familiarity (D=2))* averages this value for each patent that cites science. *Normalized Field Counts (D=2)* equals the number of unique WOS Fields found in scientific papers cited by a scientific paper in the focal patent's front page NPL citation list, divided by the number of these "D=2" cited papers.

Table B4: Post Cold War Federal R&D Shifts and MFT (Bootstrapped Standard Errors)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	1st Stage IV	2nd Stage IV	1st Stage IV	2nd Stage IV	1st Stage IV	2nd Stage IV
Dependent Variable:	ln(Share of Reassigned Patents)	ln(Avg Cites to Science)	ln(Share of Reassigned Patents)	ln(Avg Cites to Science)	ln(Share of Reassigned Patents	ln(Avg Cites to Science)	ln(Share of Reassigned Patents
ln(Avg Cites to Science)	1.040**		0.905**		0.666**		0.650**
	(0.051)		(0.093)		(0.201)		(0.197)
Number of Papers (Predicted, 1000s)		0.075**					
		(0.003)					
Δ ln(Gov. R&D Contracting)				0.189**			
				(0.009)			
Δ ln(Gov. R&D Contracting) (Predicted)						0.316**	
						(0.016)	
ln(Gov. R&D Contracting (Pre, $1Bn))	-0.476**	-0.005	-0.478**	0.125**	-0.481**	0.094**	-0.481**
	(0.078)	(0.029)	(0.091)	(0.030)	(0.092)	(0.029)	(0.088)
ln(Gov. non-R&D Contracting (Pre, $1Bn))	0.314**	-0.062**	0.307**	-0.037**	0.296**	0.034**	0.295**
	(0.028)	(0.012)	(0.043)	(0.012)	(0.043)	(0.012)	(0.042)
ln(Number of Patents)	-0.042**	-0.002	-0.037*	0.022**	-0.027	-0.010*	-0.027
	(0.007)	(0.004)	(0.016)	(0.004)	(0.017)	(0.004)	(0.016)
Share of Small Assignees	0.396**	-0.244**	0.364**	-0.242**	0.308**	-0.263**	0.305**
	(0.047)	(0.011)	(0.053)	(0.012)	(0.068)	(0.012)	(0.064)
Avg of DV	2.059	0.109	2.059	0.109	2.059	0.109	2.059
SD of Science	0.128		0.128		0.128		0.128
Cragg-Donald F-Stat		700.301		358.905		320.356	
Year Fixed Effects	Yes		Yes		Yes		Yes
R^2	0.119						
N	1,928	1,928	1,928	1,928	1,928	1,928	1,928

Notes: This table replicates the results in table 9 and A3 with bootstrapped standard errors. Analysis is at the 4-digit IPC-patent publication year level. All specifications include patent publication year fixed effects. Standard errors are from 1000 bootstrapped samples.