Science and the Market for Technology*

Ashish Arora† Sharon Belenzon‡ Jungkyu Suh §

November 13, 2019

Abstract

Well functioning Markets for Technology (MFT) allow inventors to sell their inventions to others that may exploit them better. Indirectly, MFT also enhance welfare by supporting a division of labor between upstream inventors and downstream commercializers. In this paper we explore the relationship between science and MFT and argue that inventions based in science should be more tradable. Science increases the efficiency of inventive activity, and increases the value of the resulting inventions. In addition, science reduces the transfer cost of knowledge. Conceptualizing inventions in scientific terms improves communication between buyers and sellers, reduces search costs for buyers, and enhances buyers' ability to evaluate and integrate the invention. Using large scale data, we establish a positive relationship between science and MFT. We show that this relationship is stronger for novel inventions (inventions that are more different from existing inventions), consistent with science reducing search and integration costs. To isolate the effect of science on the demand for technology acquisition, we exploit a quasi-natural experiment of the arrival of Soviet scientists to American cities. Holding fixed invention quality, an increase in the ex-post scientific understanding of an invention leads to a 22% higher likelihood that it is traded.

*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.

†Duke University, Fuqua School of Business and NBER (ashish.arora@duke.edu)

‡Duke University, Fuqua School of Business and NBER (sharon.belenzon@duke.edu)

§Duke University, Fuqua School of Business (jungkyu.suh@duke.edu)

1 Introduction

The past several decades have witnessed two distinct transformations in the innovation ecosystem: the rise of a Market for Technology and an explosion in scientific research. Data from the USPTO Patent Assignment Dataset (Graham et al., 2018) shows that patent reassignments have risen ten-fold from around 2,000 to over 20,000 cases between 1980 and 2016. U.S. corporations have reported a steady increase in royalty receipts and payments for industrial processes abroad, from \$1.5 billion and \$.4 billion respectively in 1987 to \$12.8 billion and \$4.5 billion in 2017.[1] University licensing revenues have increased ten fold over an even shorter period, from \$218 million in 1991 to \$2.5 billion in 2015 (AUTM, 2015). A similar upward trend is observed in scientific research. 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).[2] 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.[3] Moreover, the reliance of invention on science has increased as well, 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). In this paper, we relate these two phenomena to explore the effect of science on technology trade.

A well-functioning Market for Technology (MFT) enhances welfare by allowing inventors to sell or license inventions to those who can commercialize them more efficiently. However, realizing these gains from trade must contend with technology transfer costs. These costs include the direct cost of transferring knowledge, including relevant know-how, across firm

[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]"Hard science" includes Science and Engineering, excluding Social Sciences, Education, Humanities and Arts.

[3]On 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.)

boundaries. They also include the search costs that potential buyers and sellers incur, as well as contracting costs and the potential inefficiencies arising from incomplete contracts. The application of science to invention reduces these transfer costs and enhances the efficiency of MFT.

Conceptualizing inventions in scientific terms enhances communication between buyers and sellers, reduces search costs for buyers, and enhances buyers' ability to evaluate and integrate the invention. Science generalizes and abstracts phenomena into universal categories. This suggests that inventions based in science will be easier to search and understand by potential buyers (Arora and Gambardella, 1994). Scientific understanding aims beyond empirical relationships. Its objective is to uncover the mechanisms behind the phenomenon. Buyers of scientific inventions know more precisely the constituent parts of an invention and what modulates its behavior. This makes it easier to integrate such inventions into existing technology platforms and predict better their optimal use.

The development of solid-state electronics demonstrates the effect of science on technology buyers. When AT&T's Walter Brattain and John Bardeen patented their point-contact transistor in 1948, they also submitted a paper to the journal *Physical Review* formalizing their discovery.[4] William Shockley similarly followed up his junction transistor invention by writing the definitive textbook for the subject: *Electrons and Holes in Semiconductors* (Riordan and Hoddeson, 1997). The transistor was formulated in universal categories and its mechanisms were understood in scientific terms. This partly accounts for the success that AT&T had in licensing the technology to GE, IBM, Raytheon, Texas Instruments and Tokyo Tsushin Kogyo (Sony) merely four years after the initial invention (Ebert, 2008). The context-independence of knowledge of transistors also explains how major follow-on innovations could emerge from outside of AT&T. Jack Kilby at TI (Texas Instruments), for instance, acquired his knowledge about transistors from a 1952 transistor seminar at Bell Labs but still came up with a radically divergent solution to Bell when solving the problem

[4]This is in sharp contrast to Edison's discovery of the Thermionic effect in 1883, which was not used in signal manipulation until after Owen Richardson's scientific formalization in 1901.

of increasing circuit complexity — the Integrated Circuit. Even though Bell scientists were the original inventors of the transistor, their idea for a follow-on development far less efficient that the integrated circuit design proposed by Kilby (using Germanium) and, a little later, Noyce (using silicon).[5] This example also shows that the original inventor does not have a monopoly on good follow-own development ideas; thus, a market for technology can create gains from trade (Gertner, 2013).

In this paper, we investigate how science increases the rate of trade in technology by increasing gains from trade and reducing the cost of transferring technology from seller to buyer. We proceed in two steps. First, using large scale data on U.S. patents and scientific publications between 1980 and 2016, we establish a positive relationship between science and MFT. We measure the "science basedness" of inventions by whether a patent refers to a scientific article (Marx, 2019). This dataset matches front page Non Patent Literature (NPL) citations in a U.S. patent to peer-reviewed scientific publications from Microsoft Academic Graph (MAG). We use patent reassignments from the USPTO Patent Assignment Database (PAD) to measure transactions in MFT.

Our baseline finding is that a patent which cites a scientific publication has a 23% higher probability of being traded compared to a patent that does not cite science. Since citing science may be positively correlated with measures of invention quality, we control for forward patent citations. We also confirm that this positive relationship between science and trade hold for patents with high stock market value and those sharing prior art in the U.S., European, and Japanese patent jurisdiction (triadic patents). Furthermore, consistent with the view that science-based inventions should have lower search and integration costs, we find that the science-MFT relationship is stronger for novel inventions, that is, inventions that are more different from what is already known. Other things being equal, potential buyers may find it harder to search for and integrate inventions that are less familiar. Thus, transfer costs are more likely to hinder trade in novel patents and should be more responsive

[5] Bell Labs' Jack Morton proposed single transistors that would perform many tasks, rather than Kilby's solution to produce many transistors in one chip.

to science. Our estimates indicate that the effect of science on patent reassignment is around four times larger for novel invention compared to those for "not novel" inventions.

Our second contribution is identifying the causal effect of science on MFT by exploiting an exogenous source of variation. We ask how an *ex-post* increase in the scientific understanding of an invention changes the rate of trade, holding invention quality fixed. We leverage the collapse of the Soviet Union, which spurred the migration of high-caliber scientists into the United States, as a source of exogenous variation in the scientific understanding of inventions. Soviet scientists may have interpreted and explained inventions that are close to their subject matter expertise, reducing search and integration costs for potential buyers. Importantly, this analysis allows us to control for many sources of unobserved differences such as invention novelty and locus of invention (who performs upstream research), as these are dropped in the difference-in-difference analysis. Our findings suggest that patents similar to Soviet science experience a 22% increase in trade probabilities after Soviet scientist migration, compared to a control set of patents that are unrelated to Soviet science. Moreover, this effect becomes 84% larger for patents that are novel, compared to those that are not novel.

In summary, we offer a large-scale empirical investigation of the relationship between science and MFT. Inventions based in science are more likely to be traded. In part, this is because science lowers transfer costs in the market for technology. Science lowers search costs for buyers, increases the ability of buyers to understand and use the invention, and reduces contracting costs. Our findings imply that enhancing scientific understanding can increase social welfare over and above the role of science in generating fundamental inventions, by supporting a market for technology, which allocates ownership rights to the most efficient user of existing inventions, and indirectly, by supporting a division of innovative labor.

The rest of the paper is organized as follows: Section 2 presents the theoretical arguments; Section 3 describes the data and presents the non-parametric evidence; Section 4 includes the empirical analysis and Section 5 concludes.

2 How science affects patent trade

A market for technology arises when the value of the invention to a buyer is greater than its value to its original inventor (the potential seller). That is, an invention will be sold when it can be better exploited by someone other than the inventor, and the cost of transferring knowledge are lower than the gains from trade. Science reduces transfer costs and increases gains from trade.[6]

2.1 Gains from trade and its determinants

The gains from trade for patent is equal to the value a buyer derives from a patent net of the value the inventor (seller) derives from it. There are two primary components that alter these gains.

The first is comparative advantage: buyers with more resources and capabilities to commercialize the invention will increase gains from trade. For instance, a pharmaceutical patent requires complementary experience in navigating FDA regulations and a salesforce catering to physicians. Therefore, larger incumbents such as Eli Lilly would have derived more value from the use of the Swanson-Boyer recombinant DNA patent than a nascent biotechnology company such as Genentech. This explains the patent licensing arrangement in which Lilly produced Humulin by licensing Genentech's patent. It is also consistent with "small business" patent assignees (defined by the USPTO as those with less than 500 employees) exhibiting higher rates of patent trade (Figueroa and Serrano, 2019). Therefore, we predict that the size of the inventor (seller) is negatively correlated with patent trade.

If comparative advantage drives gains from trade, then it follows that markets with a more diverse set of potential buyers are more likely to exhibit trade, compared to markets with homogeneous buyers. Heterogeneity make it more likely, for a given market size, that some potential buyer will have a higher valuation for the invention than the original inventor.

[6] Appendix A formally models these mechanisms.

Hence, we predict that markets with more heterogeneous buyers will exhibit more patent trades.

Inventions may need to be integrated into pre-existing production processes for successful commercialization. Patents that embody unfamiliar technology are less likely to appeal to potential buyers. Indeed, a wide literature on the diffusion of novel technologies is inspired by the widespread empirical pattern that adoption is often slow due to lack of complementary investments and poor modularity of novel inventions to existing technological regimes (David, 1990; Bresnahan et al., 1996; Gross, 2018). Therefore, we predict that novel inventions are, on average, less likely to be traded.

The second mechanism through which gains from trade is affected is via rent dissipation: if buyers and sellers can guard against the erosion of their respective market power, gains from trade will increase (Gans et al., 2002). For instance, if the potential buyer operates in a different product market than the seller, then the transfer of a patent does not dissipate any rents for both parties. However, gains from trade between competing firms would be lower, as the transfer of an invention directly cannibalizes into a seller's existing market power (that is, the invention is worth more to the seller & therefore gains from trade are lower). BP chemicals, for instance, licensed its acetic acid technologies only to markets where it had no market access. However, the firm licensed aggressively in polyethylene, where it had a very low market share, competing with Union Carbide for licensing revenues in this segment (Arora and Fosfuri, 2000). We do not directly explore rent-dissipation, but do find some evidence consistent with its existence in section 5.

2.2 Science and transfer costs

Gains from trade must contend with the existence of transfer costs before a patent trade can occur. Transfer costs for an invention consist of integration costs, uncertain property rights, and expropriation hazards. Science can increase the likelihood of trade by decreasing these costs.

2.2.1 Lowering integration costs

Science can affect buyers by lowering expected integration costs. An extensive literature on technology adoption emphasizes the costs of integrating new technologies into existing systems. Complementary investments are typically highlighted as critical for new technology adoption in numerous cases, such as factory electrification (David, 1990) and client/server computing (Bresnahan and Trajtenberg, 1995). Science breaks down an invention into its prime components, hence making it more modular. Modular systems are "loosely coupled" such that "a change in the design of one component [does not] require compensating design changes in other components" (Sanchez and Mahoney, 1996). The scientific decomposition of an invention into its distinct functional components should allow buyers to anticipate better how each part will interact with existing components, hence lowering the costs of integration.

The rise of fabless firms in the semiconductor industry illustrate this point. The modern semiconductor industry, since its founding by Fairchild Semiconductors in 1957, had largely internalized the design and manufacture of integrated circuits within a firm, providing little room for division of labor between firms. This started to change with Application Specific Integrated Circuit (ASIC) companies such as VLSI Technology and LSI Logic, which designed circuits for systems companies trying to devise new chips for downstream applications (CD players, computers etc.). The mid-1980s saw the advent of "fabless" companies such as Xilinx and Chips and Technologies that only designed chips, and "foundry" companies such as Taiwan Semiconductor Manufacturing Company (TSMC) that only manufactured them. Most recently, the rise of IP licensing firms such as ARM that only license processor cell libraries (without designing chips) have become prominent players, decomposing the value chain even further.

This development of MFT in semiconductors has relied upon sustained increases in the scientific understanding of the surface chemistry of semiconductors, which had been refined throughout the development of the transistor in Bell Labs in the 1950s. Follow-on inventions of science-based tools further lowered transaction costs by letting buyers and sellers better understand what technical component was being traded. For instance, the EDA industry, which provided standardized CAD-like software for chip designers, facilitated communication between designers and foundries.

2.2.2 Clarifying property rights

A fundamental problem of selling knowledge is that the bargaining process requires inevitably disclosing the good itself (Arrow, 1962). Patents help, but do not fully solve the problem. First, the seller needs to convince the potential buyer of the value of the invention, which requires disclosing the invention to the buyer. A scientific decomposition of the information allows for only the relevant parts to be disclosed, without giving away the "secret sauce" (Anton and Yao, 1994). More to the point, patents often do not demarcate the scope of the invention clearly and effectively. Fuzzy patents are not very useful in protecting the invention, and even more so, if the invention has to be traded. Indeed, Arora and Fosfuri (2000) point to the clear, explicit description of "formulae, reaction pathways and operating conditions" represented via Markush structures as one of the reasons why chemical patents tend to work better in protecting property rights (Levin et al., 1987). The semiconductor industry also provides support for this effect. Before the entry of specialized EDA firms, companies such as HP internalized chip designs in their in-house CAD groups, partly because sharing chip layouts with foundries could disclose critical parts of their design. With the publication of "Introduction of VLSI Systems" by Caltech's Carver Mead and Xerox PARC's Lynn Conway in 1980, however, computer scientists began to write more standardized software (Nenni and McLellan, 2013). This shifted the locus of problem-solving to the downstream client, and thereby avoided the inadvertent transfer of "tacit," "sticky" information important to the firm (Von Hippel, 1994).

Conversely, the absence of a scientific clarification of property rights (who owns what, how, and why) has led to intractable patent disputes in new technologies. For instance, the Wrights brothers initiated a patent war against Glenn Curtiss for his use of ailerons in 1908, merely five years after their first flight at Kittihawk. The difficulty of determining infringement likely stemmed from the primitive state of the science of aerodynamics at the time. The resulting stalemate was only ended when the the U.S. government forced a "patent pool" solution during World War I (Merges and Nelson, 1990). Apple's patent battles with Microsoft over the use of the graphical user interface (GUI) in 1988 also started five years after the introduction of its Lisa computer (Hiltzik, 1999). Xerox then joined the fray as it sued Apple in 1989 over copying its GUI research in the Alto computer from the Palo Alto Research Center (PARC). The protracted legal battles regarding design patents surrounding the smartphone between Apple and Samsung, also attest to the difficulty of adjudicating infringement when a new technology is not understood scientifically. In other words, science clarifies the scope of the invention, leading to "crisp" patent boundaries, thereby reducing contracting costs for the invention.

2.2.3 Reducing share of tacit knowledge

Knowledge relevant for exploiting an invention is valuable to buyers but often costly to transfer (Von Hippel, 1994). The cost of transferring such complementary knowledge depends on whether the knowledge can be usefully codified in general categories, or, whether the knowledge is context dependent and difficult to articulate. Context dependent knowledge is often empirically derived, based on experience, and has a large tacit component (Von Hippel, 1994; Polanyi, 2015; Kogut and Zander, 1992; Arora, 1995). Contracts on tacit knowledge are more likely to be incomplete, as the transfer of tacit knowledge requires the active cooperation of the seller, but such cooperation is best induced through longer term relationships rather than arms-length contracts (Arora, 1995). Incomplete contracts increase the risk of hold-up and bargaining breakdowns (Galasso and Schankerman, 2014; Merges and Nelson, 1990). As noted, a scientific understanding of inventions also implies that the complementary knowledge is easier to codify, cheaper to transfer, and less vulnerable to contracting failures.

A classic example of the importance of tacit knowledge for trade is the repeated failure by American industrialists to replicate cellulosic fiber production based on the Bevan, Cross and Topham Patents from Great Britain in the early 1900s (Hounshell, 1988, p.5). Inspired by Counte de Chardonnet's success in generating "artificial silk" from wood pulp in the 1850s, Britain's Charles Cross and Edward Bevan succeeded in refining the product by inventing the viscose process. Rayon production grew from 2 million pounds in 1900 to 15 million pounds in 1910. However, even after buying the American rights to these viscose patents, firms such as the Cellulose Products Company and the General Artificial Silk Company failed to replicate the European processes and were dissolved in the first decade of the 1900s. It was only after the British manufacturer Courtaulds had made a Foreign Direct Investment in the American Viscose Company, and American producers having observed each manufacturing process and importing the machinery, that domestic cellulosic fiber production could begin in earnest. Science can reduce such information loss by increasing the share of explicitly codified knowledge and reducing the share of uncodifiable tacit knowledge. Advances in polymer chemistry, for instance, made it possible for Du Pont to source polyethylene and polyester, in stark contrast to its difficulty with replicating celullosic fibers (Mueller, 1962).

2.3 Science and the gains from trade

In addition to reducing transfer costs, science can also affect gains from trade. We hypothesized that inventors with less commercialization capability are more likely to sell, since the value of an invention is lower to them than it is to larger incumbents with complementary assets. However, potential buyers may harbor doubts about the value of an invention, because patent documents need not disclose all the tacit knowledge required to realize the claims of the patent. These doubts may be exacerbated for smaller firms, universities, and individual inventors, who do not have the downstream assets to prove otherwise. Reliability may therefore explain why research papers by corporations, for instance, tend to be cited more often in patents than those by universities on the same discoveries (Bikard, 2018). Scientific results, however, may bolster the reliability of an invention, and further increase likelihood of sale for small sellers.

Science may increase the potential uses of an invention, making it more attractive to cooperate with other producers (by selling technology) rather than competing with them (by commercializing products). 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). However, new uses were discovered for LEDs beyond indicator lights to general lighting (made possible after Shunji Nakamura's discovery of blue LEDs) and screen displays (diffused after the adoption of Organic LEDs in mobile phones). This enabled specialized firms such as Universal Display Corporation (UDC) to avoid entering the downstream market and sell their intellectual property on dopants to incumbents.[7]

Science can also increase gains from trade by clarifying and generalizing the underlying mechanisms of a novel or obscure invention. This will enable the patent to be relevant to a wider set of buyers. The Bessemer process, for instance, diffused rapidly after its metallurgical properties were sufficiently understood. In a similar vein, connecting more "distant" buyers and sellers to match with each other reduces rent dissipation, thereby increasing gains from trade.

[7]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.

2.4 Science and the entry into invention

So far, we have discussed the determinants of patent trade. However, it is also possible that higher prospects of trade also raise the incentives to invest in science and enter into invention for potential inventors. Prior works have shown that a broader MFT leads to more upstream research (Bresnahan and Gambardella, 1998). This is because a downstream industry with a large number of distinct buyers can encourage entry of upstream specialists, such as technology suppliers, who can spread their fixed cost investments over a larger number of buyers. The petrochemical sector offers evidence to support this dynamic. For instance, the U.S. petroleum industry was fragmented, with many municipalities operating their own refineries. Moreover, oil from different fields had varying amounts of impurities. This created an opening for firms such as UOP, which developed and offered the Dubbs process to refine gasoline to independent refiners. The demand from independent refiners led UOP to develop also the Udex process to separate aromatic compounds from virtually any kind of hydrocarbon feedstock (Gambardella, 2002). Arora et al. (2009) test this proposition in the global chemical industry in the 1980s and 1990s. They find that an increase in the number of downstream chemical firms (buyers of chemical plants) increases the number of upstream Specialized Engineering Firms (SEFs) that provide plant design services.

In sum, while science increases gains from trade, the expansion of MFT itself can also increase science-based invention, which will further strengthen MFT and reinforce this dynamic (Young, 1928). Therefore, we present an empirical identification strategy holding fixed the dynamic effect of MFT on the supply of tradeable inventions, as well other sources of inter-patent heterogeneity.

3 Data

We combine data on patents and peer-reviewed scientific publications to examine the relationship between science and technology markets. 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).[8] We also construct combine measures of invention novelty adopted from previous research (Fleming, 2001). 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.

[TABLE 1 ABOUT HERE]

3.1 Science-based inventions: Patent citations to scientific publications

We define science-based inventions as those that make at least one citation to a scientific article (Narin et al., 1997; Arora et al., 2017; Roach and Cohen, 2013; Sampat, 2010). NPL citations often contain material unrelated to "science" ranging anything from foreign patents, magazine articles, computer code, to trade publications. Therefore, simply counting the number of NPL entries for each patent would fail to accurately capture the use of science by an inventor in a patent. We employ a publicly available dataset from Marx (2019), 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 exclude matches with under a confidence score of 9. We identify 723,351 patents that cite a scientific article in MAG at least once between 1980 and 2011.

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

3.2 Invention quality and patent value: Triadic patents and stock market value

A major concern with identifying the effect of science on patent trades is that citations to science are confounded by higher quality of the underlying invention and of the patent protecting it. Citation to science may indicate that the inventor has been professionally trained, is knowledgeable, and has searched thoroughly through the scientific literature to disclose the NPL citation as prior art.[9] We use three methods to measure the quality of a patent and the underlying invention. We also exploit a quasi-natural experiment in section 5 to fix patent characteristics and vary the supply of science for patents.

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. Since prior art citations are used to delineate the novelty of a patent, it is frequently used as one measure of patent quality. Second, we source the list of "triadic patents" from the OECD, which defines them as patents that share at least one prior art across the three large patent jurisdictions - the European, Japanese, and U.S. patent offices (Dernis and Khan, 2004). That the same invention is patented in all three jurisdictions implies that the value to the inventor is high. Third, we use the stock market valuation of patents from Kogan et al. (2017) for firms listed in American stock exchanges. The authors estimate the dollar value of patents based on excess stock returns for U.S. public firms on the date of the patent's issuance date recorded in the USPTO official gazette. Since price fluctuations are specific to the firm's baseline market value, we normalize them by the market capitalization of the focal firm and then classify patents with below-average stock market returns as low value and vice versa.

[9] Of the 723,351 patents that cite at least one scientific article between 1980 and 2011, there are 264,250 (37%) patents whose citations to science are added exclusively by applicants. On the other hand there are around 84,066 (12%) patents whose citations to science are added exclusively by patent examiners.

3.3 Market for Technology: 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).[10] The USPTO records transfers of ownership that occur between patent assignees. While the reporting of transfers is voluntary, it takes precedence as proof of ownership in cases of legal dispute. Therefore, the assignees that acquire patents have an incentive to report transfers. Since the reassignments can happen for reasons unrelated to MFT, we build on Serrano (2010). We first leverage the cleaning that Marco et al. (2015) has performed and select only transfers marked as "assignment of assignors" by the USPTO, excluding mergers, security interests, name changes, and record corrections. We exclude assignments from an inventor based on a number of string distance measures, comparing inventor name to assignor (seller) name.[11] and whose assignment dates coincide with the patent issue date, since these indicate that the assignment is an initial assignment. For the same reason, we exclude assignments where the disambiguated "initial" assignee name from USPTO PATSVIEW is similar to the buyer (assignee) name on PAD. We furthermore follow the updated data cleaning procedure in Figueroa and Serrano (2019) which leverage Thomson SDC Platinum data to censor patent assignments between M&A target and acquiror companies between 1980 and 2015. Assignments occur frequently between business units within the same firm or between subsidiaries of multi-firm conglomerates. For our purposes, these assignments are not transactions between two independent entities. Therefore, we censor out assignments between similar-named entities, such that assignments between "Microsoft Corporation" and "Microsoft Ventures" are not classified as MFT transactions. Finally, we censor assignments transferring more than 25 patents at a time, which are likely to be M&A transactions. These

[10]For simplicity, we consider patent trades in this paper, which, unlike licensing, entail an exclusive transfer of property rights, such that the original inventor cannot generate revenue from the invention after selling.

[11]We use jaro-winkler, jaccard, and a modified levehnstein distance. The last distance metric, available from python's fuzzywuzzy module, handles arbitrary changes in first-last name orderings (i.e. "Jane Doe" vs "Doe, Jane" name combinations will not incur a distance penalty.)

procedures yield a total of 243,977 patents with at least one (re)assignemnt that we classify as an MFT transaction.

3.4 Novelty

We measure patent novelty using the technology combination familiarity measure from Fleming (2001). We count the number of times the same sub-class combination had appeared before the patent publication date. The assumption is that combinations of sub-classes that appear more often should be more familiar to buyers and sellers of patents. 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. As this variable is highly skewed (14.435), we take the natural logarithm of this measure in the regressions below.

3.5 Size of inventor

We hypothesized that the probability of trade is decreasing in the commercialization capability of an inventor. We measure this first by the declared size of patent assignees in USPTO maintenance fee payment records. Firms with less than 500 employees are classified as small entities per section 41 of the U.S. patent act, and benefit from a 50% reduction in filing and maintenance fees. Therefore, based on the payment date of the maintenance fee, it is possible to identify whether a patent was owned by a small or large patentee at the time of a sale. Second, we match the initial patent assignee names to public company names in Compustat. The fuzzy-matching process takes into account misspellings, name variations and acronyms (e.g. INTERNATL BUSINESS MACH CORP is matched to IBM) for 4,000 U.S. headquartered Compustat firms with positive R&D spending between 1980 and 2015. The matching procedure also covers subsidiaries that we identify by ownership data from BvD Orbis from 2002-2015 and M&A records in SDC platinum for years before 2002. Full detail on the matching process can be found in Arora et al. (2017).

3.6 Market characteristics: Buyer 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 buyers, the more unequal the distribution of valuations of inventions. We first extract the assignee names that are disambiguated in the HBS inventor dataset. 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. By constructing this measure, we implicitly assume that the assignees approximate the potential buyers in a technology market, and that 4-digit IPC classes are appropriate delineators of technology markets.

3.7 Market characteristics: Seller characteristics

We proxy the capability of the marginal seller by the average size of patent holders and the total number of unique sellers. If the marginal seller's capability is lower, average size of patent holders would be lower, while there would be more sellers of patents. 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 the string distance measures used in section 3.3. Assignee name pairs that are sufficiently similar to each other are then treated as a single name. 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.

4 Correlational Evidence

4.1 Science and the gains from patent trade

[FIGURE 1 ABOUT HERE]

Figure 1 presents the key relationships in the raw data. It contrasts the reassignment probability of patents that use (cite) science and those that do not, splitting the sample by novelty, size (maintenance fee payment & Compustat ownership), and buyer value heterogeneity. Our results are summarized in this plot: First, science-based patents are more likely to be traded than non-science based patents (with the exception of Compustat patents). Second, the gain in reassignment probability from using science is stronger for patents that are novel, belong to smaller entities & non-Compustat firms, and are in markets with heterogeneous buyers.

The reassignment share gap between science-using and non-using patents is 1.6% for novel patents, and 0.9% for not novel ones. The same gap for small entity patents is 3.6%, while only 1% for large entity patents. Non-C4 patents in IPCs with high C4 ratios show a 1.4% gain in reassignment probability when using science, while those in low-C4 IPCs barely show any difference (.1%) between science-using and non-using patents. We proceed to present econometric evidence examining the relationship between cites to science and trade. We first estimate patent-level specifications:

$$\begin{aligned} Pr(Reassignment_i = 1) = &\beta_0 Cite\_Science_i + \beta_1 Cite\_Science_i \times Familiarity_i \\ &+ \beta_2 Familiarity_i + \gamma Z_i' + \tau_t + \eta_c + \epsilon_i \end{aligned} \tag{1}$$

Specification (1) tests our prediction that novel (unfamiliar, obscure) patents are less likely to be reassigned, and that science mitigates this effect. *Reassignment* receives a value of one for patents that are reassigned at least once during their term and zero for patents that are never reassigned within our time window.[12] *Cite_Science* is a dummy variable that

[12]A patent that is reassigned multiple times gets the same *Reassignment* value of one as a patent that

receives the value of one for patents with at least one NPL citation to Microsoft Academic. *Familiarity* is measured by patent subclass combination familiarity (Fleming, 2001). We expect $\hat{\beta}_1 < 0$ and $\hat{\beta}_2 > 0$.

$$\begin{aligned} Pr(Reassignment_i = 1) = &\beta_0 Cite\_Science_i + \beta_1 Cite\_Science_i \times Seller\_Size_i \\ &+ \beta_2 Seller\_Size_i + \gamma Z_i' + \tau_t + \eta_c + \epsilon_i \end{aligned} \tag{2}$$

Specification (2) tests our prediction that higher seller commercialization capability leads to less trade. Seller commercialization capability is measured by whether the patent is owned by a "small" firm defined by the USPTO, and by whether it is owned by a Compustat company. We expect $\hat{\beta}_1 < 0$ and $\hat{\beta}_2 < 0$ That is, we expect that large inventors are less likely to sell their invention than smaller inventors, but that the use of science reduces the gap by reducing transfer costs. As well, the increase in gains from trade may be more relevant for bigger inventors.

$$\begin{aligned} Pr(Reassignment_i = 1) = &C4_i \times \{\beta_0 Cite\_Science_i \\ &+ \beta_1 Cite\_Science_i \times Buyer\_Concentration_i \\ &+ \beta_2 Buyer\_Concentration_i\} \\ &+ \{\beta_4 Cite\_Science_i \\ &+ \beta_5 Cite\_Science_i \times Buyer\_Concentration_i \\ &+ \beta_6 Buyer\_Concentration_i\} \\ &+ \gamma Z_i' + \tau_t + \eta_c + \epsilon_i \end{aligned} \tag{3}$$

Specification (3) tests the prediction that markets with more heterogeneous buyers will see more trade. We measure buyer concentration by the share of C4 patentee-owned patents in a focal patent's 4-digit IPC. $C4_i = 1$ for patents owned by C4 patentees. Because C4

has been reassigned once.

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 measure the effect of the C4 patentee share separately for patents belonging to C4 and non-C4 patentees. Our coefficient of interest is $\hat{\beta}_6$, which we expect to be positive since it measures the effect of buyer concentration (or inequality) on trade for non-C4 patentees. Furthermore, we expect $\hat{\beta}_5 > 0$, since science moderates this effect.

We hypothesize that science reduces transfer costs, and therefore expect $\hat{\beta}_0 > 0$ for specifications (1) and (2) and $\hat{\beta}_4 > 0$ for (3). Since patent quality may also affect reassignments, we control for forward patent citations in all three specifications.[13] τ_t and η_c denote year and 4-digit IPC dummies, respectively. ϵ_i is an *iid* error term.

Table 2 presents the Linear Probability Model (LPM) estimates. As expected, the coefficient estimate for *Cite_Science* are positive and significant (ranging between ranges between 11 to 16% in magnitude relative to the sample mean, depending on the controls), consistent with the prediction that science reduces transaction costs and thereby increases the probability of trade. Without any controls (but with patent grant year and 4-digit IPC fixed effects), citing science is associated with a 23% increase in reassignment probability.[14]

Columns 1 and 2 test the effect of patent familiarity on reassignment. As expected, the coefficient $\hat{\beta}_1 < 0$ and $\hat{\beta}_2 > 0$ and statistically significant. Patents whose subclass combinations are in the first decile of *Combination Familiarity* scores (in other words, novel patents) are 1% less likely to be traded compared to those in the tenth decile (not novel patents). However, novel patents that are based in science are 22% more likely to be traded than novel patents not based in science, whereas the same difference for not novel patents is 6%: the effect of science on reassignment is close to four times larger in novel patents. Columns 3 through 6 test the effect of seller (inventor) size on reassignment. Patents belonging to "small" firms with under 500 employees are around 6% more likely to be sold

[13]Number of forward patent citations normalized by the number of citations received by all patents in the focal patent's publication year

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

than those owned by "large" firms (column 3). On the other hand, small firm patents that use science are 43% more likely to be sold, compared to small firm patents not based in science (column 4). Patents owned by non-Compustat owners are 25% more likely to be sold relative to Compustat patents (column 5), while those using science are 40% more likely to be sold, compared to non-Compustat patents not based in science (column 6). Buyer concentration results in column 7 show that more concentrated patent markets (those with higher C4 patentee shares) have higher reassignment rates. For non-C4 patentees, being in a concentrated market where 90% of all patentees are C4 patentees leads to a 2% gain in reassignment probability compared to one where only 10% of all patentees are C4 patentees. However, patents that use science in those concentrated markets are 16% more likely to be traded than those that do not use science (column 8).

[TABLE 2 ABOUT HERE]

4.2 Science, MFT, and entry into invention

Consistent with science indirectly increasing the entry into invention by increasing the probability of trade, we expect the commercialization capability of the marginal seller (inventor) to decrease as the use of science in an IPC-year increases. We proxy the commercialization capability of the marginal seller by (i) the share of small sellers and (ii) the number of sellers in a patent market. We test the following specification for 4 digit IPC c at year t:

$$y_{ct} = \beta_0 + \beta_1 \frac{No.\ of\ Science\ Citing\ Patents_{ct}}{Patent\ Stock_{ct}} + \gamma Z' + \eta_c + \tau_t + \epsilon_{ct} \tag{4}$$

where y_{ct} is (i) the share of patents issued to "small" entities by their payment of maintenance fees and (ii) the number of unique sellers from the USPTO (identified through the cleaning procedure in section 3.7). We control for higher technological opportunity by average forward citations because technological advances may encourage the entry of new sellers. We also include IPC (η_c) and year(τ_t) fixed effects to exclude the effect of any year or technology class-specific differences. $\epsilon_{c,t}$ is the *iid* error term. We expect $\hat{\beta}_1 > 0$.

[TABLE 3 ABOUT HERE]

Table 3 presents the estimation results. As expected, $\hat{\beta}_1 > 0$. We find that IPC-years that have a higher share of patents citing science tend to have larger number of patentees (sellers), while the average sizes of patentees are smaller. Our results indicate that a one standard deviation increase in science-citing patent share from the sample mean translates to a 15% gain in the share of small entities (from 28% to 32%). A similar gain is observed for number of sellers: there are 0.068 sellers per patent on average, but IPC-years that cite science a standard deviation more often have 0.078 sellers per patent.

4.3 Science and patent quality

A major threat to the validity of our correlational results is that patents that cite science may systematically differ from non-citing patents. In particular, they may represent higher quality inventions and better crafted patents. Therefore, in addition to controlling for forward patent citations, we split the samples by whether they are evaluated as valuable or not by the stock market, and by whether the assignee has filed a patent for the same invention in three distinct patent jurisdictions (U.S., Europe, and Japan). If the effect of science on patent trade is only measuring the effect of quality differences, then we expect the coefficient for $Cite\_Science$ to be statistically insignificant and weaker for columns 2 and 4 in table 4 (the dependent variable is patent reassignment, as before), which only include patents above average stock market value and triadic patents respectively. We find the opposite: the relationship between science and reassignment for patents with high stock market value and triadic patents is statistically significant and 1.5 to 2.5 times larger in magnitude compared to those with lower stock market value and non-triadic patents. In the next section, we exploit a quasi-natural experiment to fix invention and seller characteristics and exogenously vary the supply of science. This will help us control for unobserved heterogeneity.

[TABLE 4 ABOUT HERE]

5 Towards identifying the causal effect of science on MFT

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. Yet, other mechanisms can generate the positive science-MFT relationship. For example, if science-citing inventions, especially novel ones, are more likely to be produced by specialized institutions that lack commercial capabilities, we would find that the science-citing patents are more likely to be traded. To identify a causla effect of science on patent trade, thus, our analysis has to control for potential unobserved differences in inventions (or inventors) as the use of science varies. In this section, we show that trade rates rise when a scientific understanding of a patent improves *ex post* invention, with the arrival of Soviet scientists in the United States.

5.1 Soviet scientist migration after the Cold War

We employ a shock to the scientific understanding of inventions in the United States emanating from the collapse of the Soviet Union, which resulted in a significant immigration of scientists from the erstwhile Soviet Union to the United States. We hypothesize that the arrival of Soviet scientists improved the ability of potential buyers to find, evaluate, and use inventions related to Soviet science, and perhaps also reduced transfer costs.

Throughout the Cold War, the Soviet Union invested heavily in basic science. Soviet scientists won several Nobel prizes (four in physics, two in chemistry), beat the United States to the first satellite and manned spaceflight, and arrived at basic scientific inventions such as the laser (Phokhorov) independently from Western science. However, knowledge of these results were only incompletely communicated to the West during the Cold War. Soviet scientific publications were written primarily in Russian.[15] While translations of prominent mathematics and physics journals had become more common by the 1980s due to large scale translation efforts by the National Science Foundation, the translation process itself would delay the dissemination of knowledge from between six months to a year, and the price of the translated journals was up to to twenty-eight times the price of their Russian language originals. Moreover, cover-to-cover journal translations would not be always available, depending on the discipline (Hollings, 2016; Garfield, 2006). With the fall of the Soviet Union, state funding for basic scientific institutions in the erstwhile Soviet Union were reduced, and the ensuing economic crisis in the mid 1990s made employment in Russia even more precarious (Mirzabekov, 1993). Coupled with changes in U.S. immigration policy facilitating the relocation of Soviet scientists such as the Soviet Scientist Immigration Act of 1992, conditions in former Soviet republics shortly after the Cold War increased the number of Soviet scientists immigrating into the United States (Ganguli, 2015; Borjas and Doran, 2012).

5.1.1 Identification Strategy

The timing and location choice of immigrating scientists differentially affects the state of scientific understanding in a region in the United States. If there are inventions in this region that are similar to the subject matter that Soviet scientists have worked on, then they will be better described by those scientists after immigration. A better translation of an existing invention, in turn, will make it more tradable. On the other hand, inventions in the same region that are unrelated to Soviet science will be unaffected by the arrival of Soviet scientists.

Treatment ("Soviet") and Control ("non-Soviet") Patents — To determine which patent is close to Soviet science, we calculate the cosine similarity of patent-to-Soviet publication pair using a customized Term Frequency-Inverse Document Frequency (TF-IDF)

[15] Publications funded by the Soviet Academy of Sciences were required to publish in the language.

metric for each word in both documents. Given these pairwise closeness scores, we can classify a patent as "Soviet" or "Non-Soviet" based on how often the patent is ranked close to publications in Soviet journals.[16] We first identify 642,477 publications from 102 prominent Soviet journals listed in the International Science Foundation's (ISF) Individual Emergency Grant program. The ISF was established by George Soros in 1992 with an initial funding of $100 million with the purpose of preventing the collapse of Soviet science. Its grant programs provided financing for leading basic scientists in the former USSR, as far as they could show a proven publication record on prominent Soviet and Russian journals (Ganguli, 2017).

We match this list of journals in the Web of Science dataset and extract information on the title and abstracts of these papers.[17] For each publication, we calculate its distance to all U.S. patents. We then check whether a focal patent is listed in the top 100 closest patents for each publication. We add the number of times the patent has been matched within the top 100 closest set to a Soviet publication and normalize this count measure by the number of times a focal patent has been matched within the top 100 closest patent set to each publication in the Web of Science (regardless of whether they are Soviet publications or not). This normalization mitigates concerns that our measures are confounded by the patent being closer to science in *general*:[18]

Identifying Assumptions — We assume that whether or not an invention is close to Soviet science *before* the arrival of Soviet science is unrelated to factors affecting its potential rate of trade. Treatment is triggered when Soviet scientists arrive in a region. The outcome of interest is the reassignment probability of treatment and control patents in the affected region. If Soviet science helps with reducing search and integration costs for the "Soviet-similar" (treatment group) patents, then these patents should experience an increase in reassignment frequency post-migration relative to the control patents (patents that are unrelated to Soviet publications in the same MSA-IPC-year as the treatment patent).

[16] Appendix C provides a detailed explanation of the textual algorithm we used in the analysis.

[17] Appendix C.2 includes the names of the Soviet journals and number of articles in each journal to which we match our patent sample.

[18] For a detailed explanation on the calculation method, see appendix C below.

Furthermore, if we assume treatment group assignment is not dependent upon the timing of scientist migration, then changes in trade rates post-migration should be driven by changes in the value potential buyers attribute to these inventions.[19]

It is important to emphasize that our empirical investigation holds invention characteristics fixed and examines the change in reassignment probabilities after the arrival of Soviet scientists. This directly addresses the concern that the baseline results we found in section 4 were driven by patent quality, since the change in reassignment probability in this experiment can only be caused by *ex-post* differences in scientific understanding, whereas the underlying invention remains unchanged over time. We are agnostic, however, on whether this effect impinges on knowledge transfer costs or gains from trade, as evidence below suggests that both may be occurring.

5.2 Difference-in-Difference estimation for Soviet shock

We estimate a diff-in-diff model for patent i in year t:

$$\begin{aligned} Pr(Reassignment_{it} = 1) = &\beta_1 Soviet_i \times Post\ Migration_t + \beta_2 Post\ Migration_t \\ &+ \theta_i + \tau_t + \epsilon_{it} \end{aligned} \tag{5}$$

Reassignment receives the value of one on the year the patent has been reassigned (the earliest reassignment date is used for multiple patent reassignments) and zero before this date. Patents that are never reassigned will have zero *Reassignment* values throughout their term (or until 2011, whichever is earlier), while patents reassigned will have zero values until the year before they are reassigned. *Soviet* receives the value of one if the patent document's subject matter is semantically similar to those encountered in the texts of soviet

[19] We find very few citations from U.S. patents before 1992 to the Soviet publications in the ISF list, suggesting that Soviet science is unlikely to have influenced the inventions themselves.

scientific publications and zero if it is not similar.

$$Sovietness\ Normalized_i = \frac{No.\ of\ times\ matched\ top100\ to\ soviet\ articles}{No.\ of\ times\ matched\ top100\ to\ ALL(any)\ articles} \tag{6}$$

If a patent has a $Sovietness\ Normalized_i$ score of above 0.5, it is classified as a Soviet (treatment) patent. Otherwise, it is classified as a non-Soviet (control) patent.[20] In our sample, around 10% of all patents are classified as treatment patents.[21] For each treatment patent, we select a control patent by matching on the treatment patent's 4-digit IPC, MSA (Metropolitan Statistical Area), publication year, science citation, and normalized forward patent citations. This mitigates concerns that $\hat{\beta}_1$ is biased by certain technology classes or regions being inherently closer to Soviet science.

Post Migration is a dummy variable indicating whether a Soviet scientist migrant has arrived in the patent's MSA before a focal year. After migration, the treatment group patents in that MSA are then "treated," while the control group patents remain unaffected. We source immigration data for Soviet scientists from Ganguli (2015), which identifies 809 soviet scientist migrants in 179 MSAs from 1986 to 2002 by tracking the publication records of scientists published in prominent Soviet journals listed by the International Science Foundation (ISF). Migration is recorded if an author for a Soviet journal article whose publication address before 1992 is recorded inside the Soviet Union subsequently publishes under an address within the United States. We include patent fixed effects (θ_i) to fix inventor and invention characteristics. We also include year (τ_t) fixed effects to preclude differences in patent values due to fluctuating market conditions over time. $\epsilon_{i,t}$ is the *iid* error term.

Our coefficient of interest is β_1, which is the differential change in reassignment probabilities for patents close to Soviet publications compared to the change for the control patents far from Soviet publications (Average Treatment Effect on Treated). If Soviet scientists re-

[20] We perform robustness checks using alternative similarity score cutoffs (e.g. by defining treatment for patents in the top decile), but the results are not sensitive to these changes.

[21] This measures a dimension of the use of science in patents distinct from citations-based measures, because it measures how frequently the same, rare concept has been used between inventions and scientific publications.

duce transaction costs or increase gains from trade with "Soviet-science", we should expect $\hat{\beta}_1 > 0$. We find in column 1 of table 5 that $\hat{\beta}_1$ is positive and statistically significant. Treatment patents that are similar to Soviet science experience a 22% larger increase in reassignment probability (relative to the sample mean) after the arrival of Soviet scientists, compared to a control patents.

[TABLE 5 ABOUT HERE]

The trade advantage conferred by better scientific understanding should be larger for more novel patents. We therefore expect the effect on patent trade that we have isolated through the arrival of Soviet scientists to be greater for novel patents. In columns 2 and 3, we split our sample of Soviet and non-Soviet patents by the combination familiarity scores described in section 4 ("Novel" columns refer to patents above the median familiarity scores, while "Not Novel" columns include samples in the below median familiarity scores.). Column 2 and 3 show that the effect of Soviet science on patent reassignment is around 84% larger for novel patents, compared to those that are not.

Columns 4 through 7 split the sample by our two measures of size (maintenance fee, and Compustat ownership status) and finds results opposite to section 4.1. For instance, the difference-in-difference coefficient estimate for "Not Small" and "Compustat" patents are larger than those for "Small" and "Non-Compustat" patents. This implies that the arrival of Soviet scientists affects not only the reduction of transaction costs, but also the reduction of rent dissipation for incumbents. If the Soviet scientists allow firms to explore more distant buyers (i.e., potential buyers that do not compete with the firm in the product market), then incumbents with existing market power would benefit more from the reduction in rent dissipation than smaller firms with no market power. Indeed, the results for buyer heterogeneity suggests that Soviet scientists also increase gains from trade, separately from the reduction of transaction costs: patents that belonged to more concentrated patent markets (column 9) tend to have a strong improvement in reassignment rates post-Soviet migration, whereas those in patent markets with relatively equal buyers (column 8) do not show similar effects.

It is therefore possible that Soviet scientists increased the value of patents to potential buyers by elucidating the science underlying an existing invention.

The validity of our results depends on our treatment group assignment being unbiased in terms of other patent characteristics. For instance, if patents similar to Soviet scientific articles were systematically of higher quality, we may expect them to be traded more frequently than patents of lower quality. However, the two panels in of appendix tables E1 and E2 show that patents in the treatment group (similar to Soviet publications) have slightly lower five-year forward patent citations (3.967) compared to those in the control group (4.619). An even more direct comparison of forward patent citations pre-trends between treatment and control patents is presented in appendix figure E1, which shows the number of citations received each year for a ten-year window before and after the first Soviet scientist arrives in the focal patent's MSA. A violation of the parallel trends assumption would have shown citations received by each group to follow a divergent path before the migration event at year zero. While citations received increase for both groups of patents in the pre-migration period, migration itself does not seem to succeed such a divergence. If our Soviet similarity assignments had been biased (for instance, "Soviet patents" happen to be "higher quality" patents), then Soviet-similar patents should show more forward patent citations received over the years. This is opposite to what we find in appendix figure E1. Finally, assuming that the use of science in general tends to increase a patent's tradability (as we argue throughout this paper), our Soviet similarity assignments would be biased if Soviet-similar patents were more likely to cite science. However, appendix tables E1 and E2 show that treatment group patents cite science less often (2.9%) than control group patents (4%). Therefore, to the extent there is a difference in ex-ante characteristics between Soviet-similar and dissimilar patents, it is more likely that they introduce a conservative bias to our estimates.

6 Conclusion

This paper aims at advancing our understanding of how science affects the rate and direction of innovation. Science, by strengthening the market for technology, can enhance social welfare by moving inventions to those that are best able to commercialize them, and by supporting a division of labor between upstream inventors and downstream commercializers. Science generalizes phenomena into universal categories and unravels the mechanisms that underpin phenomena, which enhances communication between buyers and sellers, reduces search costs for buyers, and enables buyers to evaluate and integrate inventions. We provide evidence consistent with the idea that scientific codification of inventions makes them more tradable. We further show that for a given invention, a deeper scientific understanding of the invention increases the likelihood of the invention being traded, likely because potential buyers are better able to understand and use the invention.

Our main contribution is to establish that science based inventions are more likely to be traded. Patents that reference a scientific article are 11-16% more likely to be traded than patents that do not reference science. This relationship is especially strong for inventions that are novel (different from existing knowledge), which is consistent with the view that science clarifies concepts embodied in unfamiliar inventions. Patents that are novel are less likely to be traded than those that are not, but novel patents that cite science are 22% more likely to be traded than novel patents that do not cite science. This relationship is not due to unobserved differences in quality of invention or type of inventor. Leveraging the arrival of Soviet scientists to American cities caused by an exogenous political event (the end of the Cold War), we find evidence that *ex-post* increases in scientific understanding of existing inventions makes them more tradable. All patents are more likely to trade after the arrival of high-caliber Soviet scientists in their cities, but patents that are semantically similar to science done in the Soviet Union grow 22% more in trade probabilities compared to those that are not similar.

In summary, we offer a large-scale empirical investigation of the relationship between science and MFT. Inventions based in science are more likely to be traded. In part, this is because science lowers search costs for buyers and increases their ability to understand and use the invention. Our findings imply that enhancing scientific understanding can increase social welfare over and above its role in generating fundamental inventions. Science can also increase social value by supporting a market for technology, which allocates ownership rights to the most efficient user of existing inventions, and indirectly, by supporting a division of innovative labor.

References

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., and Sheer, L. (2017). Back to Basics: Why do Firms Invest in Research? Technical report, National Bureau of Economic Research.

Arora, A. and Fosfuri, A. (2000). The market for technology in the chemical industry: causes and consequences. Revue dconomie 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.

Arora, A., Vogt, W. B., and Yoon, J. W. (2009). Is the division of labor limited by the extent of the market?: evidence from the chemical industry. Industrial and Corporate Change, 18(5):785–806.

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.

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

Bikard, M. (2018). Made in Academia: The Effect of Institutional Origin on Inventors Attention to Science. Organization Science, 29(5):818–836.

Borjas, G. J. and Doran, K. B. (2012). The collapse of the Soviet Union and the productivity of American mathematicians. The Quarterly Journal of Economics, 127(3):1143–1203.

Bresnahan, T. and Gambardella, A. (1998). The division of inventive labor and the extent of the market. In Helpman, E., editor, General purpose technologies and economic growth, pages 253–281. MIT press.

Bresnahan, T., Greenstein, S., Brownstone, D., and Flamm, K. (1996). Technical progress and co-invention in computing and in the uses of computers. Brookings Papers on Economic Activity. Microeconomics, 1996:1–83.

Bresnahan, T. F. and Trajtenberg, M. (1995). General purpose technologies Engines of growth? Journal of econometrics, 65(1):83–108.

David, P. A. (1990). The dynamo and the computer: an historical perspective on the modern productivity paradox. The American Economic Review, 80(2):355–361.

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

Ebert, L. B. (2008). What the Story of the Invention of the Transistor Teaches Us About 21st Century Patent Practice, 8 J. Marshall Rev. Intell. Prop. L. 80 (2008). The John Marshall Review of Intellectual Property Law, 8(1):4.

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.

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

Gambardella, A. (2002). Successes and failures in the markets for technology. Oxford Review of Economic Policy, 18(1):52–62.

Ganguli, I. (2015). Immigration and Ideas: What Did Russian Scientists Bring to the United States? Journal of Labor Economics, 33(S1):S257–S288.

Ganguli, I. (2017). Saving Soviet science: The impact of grants when government R&D funding disappears. American Economic Journal: Applied Economics, 9(2):165–201.

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.

Garfield, E. (2006). The history and meaning of the journal impact factor. Jama, 295(1):90–93.

Gertner, J. (2013). The idea factory: Bell Labs and the great age of American innovation. Penguin.

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.

Gross, D. P. (2018). Scale versus scope in the diffusion of new technology: evidence from the farm tractor. The RAND Journal of Economics, 49(2):427–452.

Hiltzik, M. A. (1999). Dealers of lightning: Xerox PARC and the dawn of the computer age. HarperCollins Publishers.

Hollings, C. D. (2016). Scientific communication across the Iron Curtain. Springer.

Hounshell, D. A. (1988). The Making of the Synthetic Fiber Industry in the United States.

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.

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.

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

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.

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

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

Mirzabekov, A. D. (1993). Visit Russia, help science. Nature, 363(6430):576–576.

Mueller, W. F. (1962). The origins of the basic inventions underlying du pont's major product and process innovations, 1920 to 1950. In The rate and direction of inventive activity: Economic and social factors, pages 323–358. Princeton University Press.

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

Nenni, D. and McLellan, P. (2013). FABLESS: the transformation of the semiconductor industry. sl: SemiWiki. com. SemiWiki.com.

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

Chicago Press.
Riordan, M. and Hoddeson, L. (1997). Crystal fire: the invention of the transistor and the birth of the information age. WW Norton & Company.
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.
Sanchez, R. and Mahoney, J. T. (1996). Modularity, flexibility, and knowledge management in product and organization design. Strategic management journal, 17(S2):63–76.
Serrano, C. J. (2010). The dynamics of the transfer and renewal of patents. The RAND Journal of Economics, 41(4):686–708.
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.
Von Hippel, E. (1994). Sticky information and the locus of problem solving: implications for innovation. Management science, 40(4):429–439.
Young, A. A. (1928). Increasing Returns and Economic Progress. The Economic Journal, 38(152):527–42.

Table 1: Summary Statistics for Main Variables

	Obs.	Mean	S.D.	p10	p50	p90
Patent Publication Year	3,884,291	1998.886	8.725	1986	2000	2010
Reassignment	3,883,777	0.063	0.243	0	0	0
Cite Science	3,883,777	0.186	0.389	0	0	1
Forward Patent Citations (Normalized)	3,884,291	0.924	1.625	0	1	2
5-year Forward Patent Citations	3,884,291	6.093	10.837	0	3	14
Stock Market Value of Patent	1,218,301	12.614	38.122	0	4	27
Combination Familiarity	3,882,310	76.799	386.215	0	1	155
Small Entity	3,689,667	0.226	0.418	0	0	1
Compustat Patent	3,884,291	0.254	0.435	0	0	1
Triadic Patent	3,884,291	0.304	0.460	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. *Forward Patent Citations (Normalized)* counts the number of prior art citations the patent has received until 2015 and normalizes this quantity by the average number of forward patent citations received by patents published in the focal patent's publication year. *5-year Forward patent citations* counts the number of prior art citations the patent has received within five years of its publication. *Stock Market Value of Patent* is based on the cumulative abnormal returns in the firm's market value at the issuance event of the patent per Kogan et al. (2017). *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. *Triadic Patent* is equal to one if the patent shares a prior art in the USPTO, EPO, and JPO, and zero otherwise.

Table 2: Science and the Gains from Patent Trade

	Novelty		Size				Buyer Concentration	
	(1) Reassignment	(2) Reassignment	(3) Reassignment	(4) Reassignment	(5) Reassignment	(6) Reassignment	(7) Reassignment	(8) Reassignment
Cite Science	1.104**	1.390**	1.086**	0.656**	1.210**	2.084**		
	(0.038)	(0.045)	(0.039)	(0.040)	(0.038)	(0.049)		
ln(Combination Familiarity+1)	0.012*	0.046**						
	(0.006)	(0.007)						
Forward Patent Citations (Normalized)	0.874**	0.874**	0.897**	0.897**	0.894**	0.894**	0.882**	0.881**
	(0.020)	(0.020)	(0.021)	(0.021)	(0.020)	(0.020)	(0.021)	(0.021)
Cite Science × ln(Combination Familiarity+1)		-0.200**						
		(0.016)						
Small Entity			0.350**	-0.150**				
			(0.038)	(0.041)				
Cite Science × Small Entity				2.698**				
				(0.100)				
Compustat Patent					-1.558**	-0.969**		
					(0.030)	(0.033)		
Cite Science × Compustat Patent						-2.530**		
						(0.068)		
C4 Patentee=0 × Cite Science							1.231**	0.997**
							(0.043)	(0.062)
C4 Patentee=1 × Cite Science							-0.009	-0.055
							(0.072)	(0.129)
C4 Patentee=0 × C4 Share							0.778**	0.487*
							(0.189)	(0.194)
C4 Patentee=1 × C4 Share							3.421**	3.365**
							(0.225)	(0.231)
C4 Patentee=0							0.000	0.000
							(.)	(.)
C4 Patentee=1							-4.874**	-4.894**
							(0.055)	(0.056)
C4 Patentee=0 × Cite Science × C4 Share								2.436**
								(0.481)
C4 Patentee=1 × Cite Science × C4 Share								0.324
								(0.599)
Avg of DV	6.281	6.281	6.265	6.265	6.282	6.282	6.456	6.456
IPCs	632	632	632	632	632	632	632	632
Years	32	32	31	31	32	32	31	31
R^2	0.013	0.013	0.013	0.014	0.014	0.014	0.012	0.012
N	3,882,309	3,882,309	3,689,236	3,689,236	3,883,776	3,883,776	3,658,880	3,658,880

Notes: The dependent variable is equal to one if the patent has ever been reassigned in the USPTO PAD dataset and zero otherwise. *Cite Science* indicates whether or not a front page NPL citation has been made by the focal patent to a peer-reviewed scientific article in Microsoft Academic. *Forward patent citations* counts the number of forward patent citations normalized by the number of citations received by all patents in the focal patent's publication year. *Familiarity* is equal to $ln(1 + Combination\_Familiarity)$, where $Combination\_Familiarity$ is calculated by the number of times the same subclass combination has appeared before the publication date of a patent (detail in Fleming (2001)). The coefficient estimates are multiplied by 100 for ease of reporting. All specifications include fixed effects for 4-digit IPCs and patent publication years. Standard errors are robust to arbitrary heteroscedasticity.

Table 3: Science, MFT, and entry into invention

	(1) Small Entity Share	(2) No. of Sellers
Avg Cites to Science	0.260**	0.062**
	(0.034)	(0.009)
Avg Forward Patent Citations	-0.004	0.021**
	(0.009)	(0.002)
log(Patent Stock + 1)	-0.016**	-0.004*
	(0.005)	(0.002)
Avg of DV	0.281	0.068
IPCs	334	337
Years	31	32
R2	0.918	0.571
N	6,913	7,173

Notes: Unit of observation is at the 4 digit IPC-year level. *Avg Cites to Science* is the count of patents in a 4 digit IPC-year that have made a citation to a peer-reviewed scientific article in MAG from their front page NPL citation section, normalized by patent stock. *Avg Forward patent citations* counts the normalized forward patent citations in each 4 digit IPC-year and normalizes by patent stock. Small entity share is the number of small entity (¡500 employee) patents divided by patent stock. *Number of sellers* equals the number of unique patent sellers that have been identified for each 4 digit IPC-year, normalized by patent stock. All columns include fixed effects for 4-digit IPC and patent publication years. Standard errors are clustered at the 4-digit IPC level.

Table 4: Relationship between Science and Patent Reassignment, by Quality

	Stock Market Value		Triadic	
	(1)	(2)	(3)	(4)
	below avg	above avg	No	Yes
Cite Science	0.246	0.333*	0.625**	1.468**
	(0.154)	(0.165)	(0.101)	(0.150)
Forward Patent Citations (Normalized)	0.458**	0.318**	0.891**	0.809**
	(0.101)	(0.041)	(0.074)	(0.088)
Avg of DV	2.966	7.048	5.888	7.182
IPCs	600	609	631	629
Years	31	31	32	32
Firms	380	3,149		
R^2	0.038	0.187	0.013	0.015
N	374,798	374,418	2,702,037	1,181,739

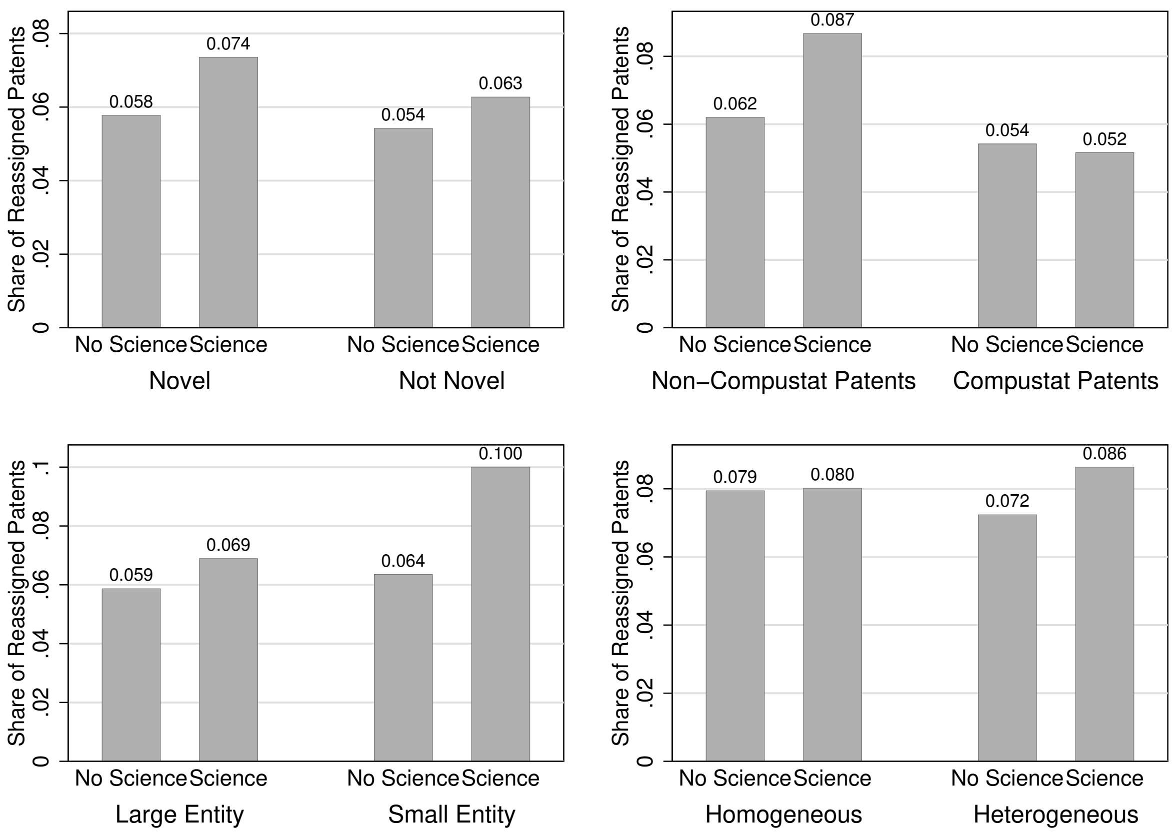
Notes: The dependent variable is equal to one if the patent has ever been reassigned in the USPTO PAD dataset and zero otherwise. Stock market value for patents are calculated by normalizing the Kogan et al. (2017) patent value measure by the market capitalization of the Compustat firm which owns the patent. Columns 1 and 2 include patent publication year, 4-digit IPC, and Compustat firm fixed effects, with standard errors clustered at the firm level. Columns 3 and 4 include patent publication year and 4-digit IPC fixed effects, with standard errors clustered at the 4-digit IPC level. All coefficient estimates are multiplied by 100 for ease of reporting.

Table 5: DIFFERENCE IN DIFFERENCES WITH SOVIET SHOCK

	All	Novelty		Size				Buyer Heterogeneity	
	(1) All	(2) Not Novel	(3) Novel	(4) Small	(5) Not Small	(6) Not Compustat	(7) Compustat	(8) Homogeneous	(9) Heterogeneous
Post_Migration × Soviet_Patent	0.098**	0.073*	0.134**	-0.084	0.181**	0.008	0.141**	0.042	0.128*
	(0.024)	(0.032)	(0.040)	(0.046)	(0.038)	(0.030)	(0.034)	(0.051)	(0.052)
Post_Migration	0.221**	0.218**	0.228**	0.242**	0.303**	0.312**	0.144**	0.664**	0.377**
	(0.035)	(0.039)	(0.039)	(0.050)	(0.054)	(0.042)	(0.050)	(0.050)	(0.056)
Avg of DV	0.441	0.448	0.431	0.416	0.456	0.564	0.308	0.577	0.326
Years	32	32	32	31	31	32	32	32	32
Patents	120,738	71,350	49,388	26,335	65,978	63,478	57,260	30,300	30,168
R^2	0.192	0.194	0.190	0.203	0.196	0.198	0.180	0.185	0.191
Observations	2,239,681	1,316,664	923,017	489,921	1,208,789	1,166,307	1,073,374	558,611	561,829

Notes: Unit of observation is at the patent-year level. Dependent variable is equal to one on the year the focal patent has been reassigned and zero before. *Post Migration* equals one after the first year in which a Soviet scientist arrives at the focal patent's MSA. *Soviet Similar* equals one if the normalized semantic similarity scores to Soviet scientific publications is above 0.5 for a patent. All coefficient estimates are multiplied by 100 for ease of reporting. All standard errors are clustered at the MSA level.

Figure 1: Science and the Determinants of Patent Trade



Notes: The bars plot the share of reassigned patent in each group, split by those that cite science and those that do not. The 1st quadrant splits the sample by whether patent owner names are matched to Compustat firm names. The 2nd quadrant splits the sample by whether the patent is in the 1st decile of combination familiarity scores (Novel) or in the 10th decile (Not Novel). The 3rd quadrant splits the sample by whether patent owner is classified as a "small" entity per section 41 of the U.S. patent act. Patents owned by non-"small' owners are classified as "large". The 4th quadrant first takes patents owned by non-C4 owners and then splits this sample by whether the patent is in a 4-digit IPC with the highest (4th quartile) share of four-patentee concentration ratios (heterogeneous) or the lowest (1st quartile) ratios (homogeneous).

A A model of trade in patents

We present a simple model to formalize our comparative statics and clarify the multiple channels through which the use of science can condition the probability of a patent being traded.

Let there be I inventors. Some of these could be individuals, universities, or firms. To start with, assume each inventor is endowed with an invention (i.e., we begin with inventions being exogenously assigned.) The inventor can commercialize the invention herself and earn payoff y_i where

$$y_i = q_i(x_i + \epsilon_i) \tag{7}$$

where q_i is the quality of the i^{th} invention. Inventions based on science may have higher quality. For the i^{th} invention, there are N_i firms that may buy it to commercialize it themselves i.e., there are N_i possible innovators. Each innovator's baseline payoff is zero. If it commercializes $i's$ invention it can earn a payoff y_{ik} and zero otherwise.

$$y_{ik} = q_i(x_{ik} + \epsilon_{ik}) - \tau_i \tag{8}$$

In (8), τ_i represents transaction and transfer costs (henceforth transfer costs), and x_{ik} represents the ability to extract value from invention i of the i^{th} potential buyer. Note that a key assumption is that transfer costs are independent of the quality of the invention.

Transfer costs have several components.

- Contracting cost: Contracts involve lawyers and meetings, which have direct and indirect costs
- Adaptation Costs: Inventions often need accompanying know-how, some of which may require training and service.
- Imperfect contracting: Effective commercialization of an invention may require the cooperation of the inventor. Such cooperation may not be forthcoming, even if contractually agreed upon, in adequate measure.

There are two primary components to x_{ik}

- Comparative advantage: how well the resources and capabilities the potential buyer posses are suited to the commercialize the invention.
- Rent dissipation: The extent to which the potential buyer can enhance its market power in the downstream product market (or guard against the erosion of its existing market power if another firm commercializes the invention).

We assume efficient bargaining. Therefore, if trade with at least one of the N_i potential buyers offers a surplus that is at least as great as the transfer costs, the invention will be traded. This is an important assumption that sweeps away considerations of asymmetric information. It also means that we can focus directly on what we observe, namely whether the invention is traded or not, without having to discuss how the net surplus (gains from trade minus transaction cost) are divided between the buyer and seller.

A.1 probability of trade

The invention is not traded if

$$
\begin{aligned}
& y_i \geq \max_k^N {}_i\{y_{ik}\} & (9)\\
\iff\ & x_i + \epsilon_i - \max_k^{N_i}\{x_{ik} + \epsilon_{ik}\} \geq -\frac{\tau_i}{q_i} & (10)
\end{aligned}
$$

We assume that ϵ_i and ϵ_{ik} are distributed *iid* as Type I extreme value distribution (Gumbell distribution). This will yield the familiar logit expressions.

The probability of the invention not being traded, P = Pr(no trade)

$$
\begin{aligned}
P(\text{no trade}) &= \frac{\exp(x_i + \frac{\tau_i}{q_i})}{\exp(x_i + \frac{\tau_i}{q_i}) + \sum_k^{N_i} \exp(x_{ik})} \\
&= \frac{1}{1 + \sum_k^{N_i} \exp(x_{ik} - x_i - \frac{\tau_i}{q_i})}
\end{aligned} \tag{11}
$$

A.1.1 special case

Consider first the case where all buyers are similar except for the idiosyncractic valuation. That is, suppose $x_{ik} = \bar{x}_i \forall k$ where $\bar{x}_i$ is the value that all innovators can derive from the invention. Then the probability of No Trade can be rewritten as

$$
P(\text{No Trade}) = \frac{1}{\exp(-x_i - \frac{\tau_i}{q_i} + \bar{x}_i)N_i + 1} = \frac{\frac{1}{N_i}}{\exp(-c + \bar{x}_i) + \frac{1}{N_i}} \tag{12}
$$

Note that the probability of trade, is simply $1 - P(\text{No Trade})$. Equation 12 shows that science increases the probability of trade through three channels

1. The number of potential innovators,N_i
2. The difference in comparative advantage between inventors and potential innovators plus any rent-dissipation effects, $x_i - \bar{x}_i$
3. The quality of the invention, and the transfer costs, $\frac{\tau_i}{q_i}$.

A.2 buyer concentration

Consider the no trade condition 11. We wish to understand how this is affected by an increase in the heterogeneity of buyers (holding the number of buyers and their average x_{ik} constant).[22]

[22] From Rothschild and Stiglitz 1970 (Rothschild, M. and Stiglitz, J.E., 1970. Increasing risk: I. A definition. Journal of Economic theory, 2(3), pp.225-243.) we know that if $g(x)$ is a convex function of x, then a second order stochastic shift in the distribution of x will imply that $Eg(x)$ will increase.

Here we prove the special case where we compare 12 with 11. That is, 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 $\mathrm{E}_k[x_{ik}] = \bar{x}_i$.

Note that 11 can be written as

$$
\begin{aligned}
P(\text{no trade}) &= \frac{1}{1 + \sum_k^{N_i} \exp(-x_i + x_{ik} - \frac{\tau_i}{q_i})} \\
&= \frac{\frac{1}{N_i}}{\frac{1}{N_i} + \frac{1}{N_i}\sum_k^{N_i} \exp(x_{ik} - c_i)} \qquad \approx \frac{1}{1 + N_i \mathrm{E}_k\left[\exp(x_{ik} - c_i)\right]} \\
&\text{where } c_i = x_i + \frac{\tau_i}{q_i}
\end{aligned} \tag{13}
$$

Comparing 13 and 12, we see that

$$
\begin{aligned}
&\frac{1}{1 + N_i \mathrm{E}_k\left[\exp(x_{ik} - c_i)\right]} < \frac{1}{1 + N_i(\exp(\mathrm{E}_k[x_{ik} - c_i]))} \\
&\text{because } \mathrm{E}_k\left[\exp(x_{ik} - c_i)\right] > \exp(\mathrm{E}_k[x_{ik} - c_i])
\end{aligned} \tag{14}
$$

where the last inequality follows from applying Jensen's Inequality and recognizing that the exponential is a convex function, and expectations are being taken over the commercialization capabilities of the buyers, x_{ik}, for a given inventor.

This shows that if there is more inequality among buyers in terms of commercialization capability x_{ik}, the probability of No Trade falls.

A.3 Role of science

Conceptualizing inventions in terms of science reduce τ_i and perhaps increase q_i. The combined result is to increase the effective gains from trade.

A.3.1 Novelty of invention

More novel inventions may not be comprehensible to buyers. Thus, novelty might reduce N_i. Let the level of novelty be denoted by z and the level of science used in invention be denoted by s. From 12, we can see that

$$
\begin{aligned}
\frac{\partial P(\text{no trade})}{\partial z} &= -\frac{1}{(\exp(-x_i - \frac{\tau_i}{q_i} + \bar{x}_i)N_i + 1)^2} \frac{\partial N_i}{\partial z} \geq 0 \\
\frac{\partial P(\text{no trade})}{\partial s} &= \frac{\exp(-x_i - \frac{\tau_i}{q_i} + \bar{x}_i)N_i}{q_i(\exp(-x_i - \frac{\tau_i}{q_i} + \bar{x}_i)N_i + 1)^2} \frac{\partial \tau_i}{\partial s} \leq 0 \\
\frac{\partial^2 P(\text{no trade})}{\partial s \partial z} &= -\frac{2\exp(-x_i - \frac{\tau_i}{q_i} + \bar{x}_i)N_i}{q_i(\exp(-x_i - \frac{\tau_i}{q_i} + \bar{x}_i)N_i + 1)^3} \frac{\partial \tau_i}{\partial s} \frac{\partial N_i}{\partial z} \leq 0
\end{aligned} \tag{15}
$$

As we see in 15, novel inventions are less likely to be traded. However, this effect is smaller for science based inventions. Put differently, the diminution in the probability of trade due to greater novelty is less marked for science based inventions.

A.3.2 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. For simplicity of notation, but without any loss of generality, we show the formal derivations for the case with uniform buyers.

$$\frac{\partial P(\text{no trade})}{\partial x_i} = \frac{\exp(-x_i - \frac{\tau_i}{q_i} + \bar{x}_i)N_i}{(\exp(-x_i - \frac{\tau_i}{q_i} + \bar{x}_i)N_i + 1)^2} \geq 0 \tag{16}$$

inventor capabilities and science As we see in equation 16, larger firms are less likely to sell (x_i is higher for larger firms). To sign the interaction between inventor size and science, note that 16 can be rewritten as

$$\frac{\partial P(\text{no trade})}{\partial x_i} = \frac{a}{(a+1)^2} \qquad \text{where}$$
$$a = \exp(-x_i - \frac{\tau_i}{q_i} + \bar{x}_i)N_i$$

This expression increases with a for $a < 1$:

$$\frac{d\frac{a}{(a+1)^2}}{da} = \frac{1-a}{(a+1)^3}$$

Note that a increases with science. Formally

$$\frac{\partial a}{\partial s} = -\frac{\partial \tau}{\partial s}\frac{1}{q_i}\exp(-x_i - \frac{\tau_i}{q_i} + \bar{x}_i)N_i = -\frac{\partial \tau}{\partial s}\frac{1}{q_i}a \geq 0 \tag{17}$$

The effect of size on the probability of trade is conditioned by science. Formally,

$$\begin{aligned} \frac{\partial^2 P(\text{no trade})}{\partial x_i \partial s} = \frac{1-a}{(a+1)^3}\frac{\partial a}{\partial s} &\leq 0 \text{ if } a > 1 \\ &\geq 0 \text{ if } a < 1 \end{aligned} \tag{18}$$

Finally, note that the probability of no trade, $P(\text{no trade }) = \frac{1}{1+a} \approx 0.9$ in the sample. Therefore, $a \approx \frac{1}{0.9} - 1 \leq 1$. We therefore expect that the probability of no trade will increase with size, and science will moderate the increase. Put differently, the probability of trade will fall with size and the use of science in invention will exacerbate the fall.

A.3.3 buyer concentration and science

Consider the probability of no trade with heterogeneous buyers, as in 11

$$\begin{aligned}\frac{\partial P(\text{no trade})}{\partial s} &= \frac{\partial \tau}{\partial s}\left(\frac{\sum\limits_{k}^{N_i} \exp(x_{ik}-c_i)}{(\sum\limits_{k}^{N_i} \exp(x_{ik}-c_i)+1)^2}\right) \\ &= \frac{\partial \tau}{\partial s}\left(\frac{N_i\, E_k[\exp(x_{ik}-c_i)]}{(N_i\, E_k[\exp(x_{ik}-c_i)+1)]^2}\right) \le 0\end{aligned} \tag{19}$$

note that 21 can be rewritten as

$$\frac{\partial P(\text{no trade})}{\partial s} = \frac{\partial \tau}{\partial s}\frac{b}{(b+1)^2} \le 0 \qquad \text{where}$$
$$b = N_i\, E[\exp(x_{ik}-c_i)]$$

The corresponding derivative for homogenous buyers from 12 is

$$\begin{aligned}\frac{\partial P(\text{no trade})}{\partial s} &= \frac{\partial \tau}{\partial s}\left(\frac{N_i \exp(E[x_{ik}-c_i])}{(N_i \exp(E[x_{ik}-c_i])+1)^2}\right) \\ &= \frac{\partial \tau}{\partial s}\frac{a}{(a+1)^2} \le 0\end{aligned} \tag{20}$$

where $a = N_i \exp(E[x_{ik}-c_i])$

The probability of no trade decreases with b for $b > 1$, but increases with b for $b < 1$. Formally, we have that if $b < 1$

$$\frac{\partial \frac{b}{(b+1)^2}}{\partial b} = \frac{1-b}{(b+1)^3} \ge 0$$

What is the value of b? Note that the probability of no trade is $\frac{1}{1+N_i b}$. In the sample, the probability of no trade is greater than 90%. This implies that $b = \frac{1}{9N_i} < 1$. A similar calculation shows $a < 1$

Therefore, we assume that $b < 1, a < 1$. Then because $b > a$, the probability of no trade increases with science more slowly when buyers are heterogeneous. Put differently, the probability of trade increases faster with science when buyers are heterogeneous.

Formally, we have that

$$\begin{aligned}&\frac{\partial P(\text{no trade})}{\partial s}|_{heterog} - \frac{\partial P(\text{no trade})}{\partial s}|_{homog} = -\frac{\partial \tau}{\partial s}\left(\frac{a}{(1+a)^2} - \frac{b}{(1+b)^2}\right) \ge 0 \\ &\implies \frac{\partial P(\text{trade})}{\partial s}|_{heterog} \le \frac{\partial P(\text{trade})}{\partial s}|_{homog}\end{aligned} \tag{21}$$

A.4 specialization and division of innovative labor

We began by assuming that inventions were exogenously assigned. However, the prospect of being able to trade an invention, thereby increasing expected returns from inventing, would make investing in invention more attractive. For simplicity, suppose the inventor captures the entire surplus from invention. The expected payoff from invention, Π is

$$\Pi = c_0 \ln \left(\exp(x_i) + N_i(\exp(\bar{x}_i - \frac{\tau_i}{q}) \right) - R \tag{22}$$

where R is the investment required to produce an invention, and c_0 is a constant that depends on the variability in the idiosyncratic valuations ϵ_{ik}. For simplicity we assume that $\tau_i = \tau(s)$ so that transaction costs depend only on the amount of science used. In particular, they do not depend on the identity of the inventor. By a similar logic, $\bar{x}_i$ and N_i also do not vary with i, although they may vary with industry conditions and the use of science.

A.4.1 Entry into invention

Suppose that potential inventors differ only in the value they could capture by commercializing it themselves. It follows that each potential inventor in given industry and technology class, with a given amount of science, expects the same transaction costs, quality, and number of potential buyers. The marginal inventor is indifferent between investing in invention or not investing. That is the marginal inventor is characterized by x^*[23] such that

$$\Pi(x^*) = c_0 \ln \left(\exp(x*) + N(\exp(\bar{x} - \frac{\tau}{q}) \right) - R = 0 \tag{23}$$

Intuitively, and as 23 confirms, that an increase in the number of potential innovators, or a decrease in transaction cost, the marginal inventor would have a lower x_i. That is, more inventors would invest, creating more invention. Importantly, the effect is due to an increase in the probability of trade. That is, as Adam Smith noted long ago, enhanced possibility for trade encourage entry. In this instance, a reduction in the transaction cost in the market for technology encourages entry into invention, especially for smaller inventors. Similarly, an increase in the productivity of research (captured by a reduction in R or an increase in q) would directly increase the probability of trade (via 12). They would also *indirectly* increase the probability of trade by reducing the x_i associated with the marginal inventor.

Formally, we have that

$$\begin{aligned}
\frac{d\Pi(x^*)}{ds} &= \frac{\partial \Pi(x^*)}{\partial s} + \frac{\partial \Pi(x^*)}{\partial x^*}\frac{\partial x^*}{\partial s} = 0 \\
&\implies c_0 \left(-\frac{\partial \tau}{\partial s} P(\text{trade}) + (1 - P(\text{trade}))\frac{\partial x^*}{\partial s} \right) = 0 \\
&\implies \frac{\partial x^*}{\partial s} \leq 0
\end{aligned} \tag{24}$$

The result in 24 follows upon noting that the expected payoff of an inventor, $\Pi(x)$,

[23] We drop the inventor subscript i because we are focusing on the marginal inventor

increases with the commercialization capability of the inventor, x, and with science s.

A.4.2 science and the equilibrium in the market for technology

It therefore follows that probability of No Trade will decrease with science, s, directly, as in 12, but also indirectly, because inventors with a lower ability to commercialize their inventions, will enter the market. Formally,

$$\frac{dP(\text{no trade})}{ds} = \frac{a}{(1+a)^2}\left(\frac{\partial \tau}{\partial s} + \frac{\partial x*}{\partial s}\right) \leq 0 \tag{25}$$

B 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.

C Calculating proximity scores between patents and scientific publications

Step 1: Bag of words

As a first step to calculating the proximity between a patent and a publication, we use the bag-of-words approach to extract all words from the claims text of all USPTO patent documents and title and abstracts of Web of Science peer-reviewed scientific articles. For each patent and article, we create a vector of all word stems. Each word stem is weighted by the inverse of its frequency in the complete patent corpus. The inverse frequency index is

$$I_i = N_i \times (1 - \frac{p_i}{P}) \tag{26}$$

N_i is the number of times ith word stem in the word stem vector appears throughout the claims section of the USPTO patents. p_i is the number of patent documents that contain the ith word stem, and P is the number of patents issued by USPTO. Each item in the index represents the weight assigned to extracted word stems according to their specificity across all USPTO patent documents.[24]

Step 2: Distance between words

Similar ideas might be described using different text. Thus, a major challenge is how to compute the "technical distance" between two words, that is how to calculate the likelihood that two different words describe the same technical concept. To address this challenge, we develop a dictionary that aims to measure the probability that two distinct words refer to the same technical concept. For this purpose, we identify words used in patent documents deemed to be technically similar by human experts, the patent examiners themselves. We use prior-art patents referenced by examiners in rejecting patent applications for a lack of novelty or obviousness to compute a measure of technical distance between two given words.

To create the technical distance between two words we follow these steps. First, we extract from the USPTO's Public PAIR (Patent Application Information Retrieval) system a random sample of about 150,000 non-final rejection letters. We include only non-final rejection letters with rejections pertaining to novelty and non-obviousness as outlined in 35 U.S.C. 102 and 35 U.S.C. 103 of the USPTO's Manuals for Patent Examining Procedure. The letters are available as images and thus are converted into a text format. For each patent with a non-final rejection, we extract the text of the original patent application associated with that rejection as well as the text of the prior-art patents cited as the reason for the rejection ("rejection prior-art patents"). In cases where multiple rejections are associated with the same application, we extract the relevant (modified) application claims for each rejection.

[24] An important part of the word stemming process is mapping acronyms and technical concepts. For example, the acronym RAM refers to Random Access Memory. Thus, in our textual comparison algorithm, when the sequence of words Random Access Memory appears, we collapse it into RAM. Acronyms appear in capital letters on patent documents. We retain all words with at least two capital letters and manually search for their meaning. To mitigate cases where multiple meanings exist for a given acronym, we preform the acronym-meaning match at the four-digit IPC level. (Chemical compounds also appear in capital letters, but we leave them unchanged.)

Second, we extract all relevant word stems from the claims section of the focal patent application and corresponding prior-art patents listed by the patent examiner as the basis for a rejection.[25] At the end of this step, we have relevant word stems extracted from the rejected applications and prior-art patents listed on non-final rejection letters. Next, we calculate the proximity between each pair of the word stems based on their co-occurrence. To account for the baseline tendency of two word stems to co-occur across two documents, for each rejected application and rejection prior-art patent pair, we construct a control pair by linking the rejected application with a control patent that was not cited as a reason for the rejection but is in the same 4-digit IPC (International Patent Classification) and has the same application year as the rejection prior-art patent. Proximity between a pair of word stems is calculated as the ratio of the number of times the pair appears in the rejected application and rejection prior-art patent to the number of times it appears in the rejected application and the control prior-art patent. More precisely, proximity between two word stems is calculated as:

$$Proximity_{w1,w2} = \frac{(A \bigcup R)_{w1,w2}}{(A \bigcup C)_{w1,w2}} \tag{27}$$

$(A \bigcup R)_{w1,w2}$ is the number of times the words $w1$ and $w2$ co-occur within the focal application A and rejection prior-art patent R. $(A \bigcup C)_{w1,w2}$ is the number of times the words $w1$ and $w2$ co-occur in the focal application A and control patent C. Because the same word stem pair $w1$ and $w2$ can co-occur in more than one application and rejection prior-art patent pair, we average the proximity scores between $w1$ and $w2$ across all application and rejection prior-art patent pairs, denoted by $\bar{P}_{W1_i,W2_i}$.

Step 3: Textual overlap between documents

The final step of our algorithm is to construct a similarity score between a pair of patent and publication based on their words and the "technical distance" between these words from Step 2. To derive the textual proximity between a patent and a scientific article, we create a vector of words for each document with their corresponding weights (i.e. inverse frequency) as described in step 1. We then calculate the cosine proximity score between the two word vectors $W1$ and $W2$, each vector consisting of n elements, while taking into account the average word pair proximity, $\bar{P}_{W1_i,W2_i}$ calculated in step 2:

$$PS_{W1,W2} = \frac{\sum_{i=1}^{i=n} W1_i \times W2_i \times \bar{P}_{w1_i,w2_i}}{\sqrt{\sum_{i=1}^{i=n} W1_i^2}\sqrt{\sum_{i=1}^{i=n} W2_i^2}} \tag{28}$$

We normalize the proximity score $PS_{W1,W2}$ to be between 0 and 1 by dividing it by $\max(PS_{W1_i,W2_i})$. A score of one indicates the highest similarity, and a score of zero indicates the lowest similarity between two documents.

Step 4: Aggregating the similarity measures

Given the pairwise similarity scores between publications and patents, we rank patents

[25] We use original applications rather than the final patent documents because claims can change through patent examination process and thus using the original applications allows us to compare the relevant set of claims between the applications and rejection prior-art patents.

from the most similar to the least similar for each Web of Science publication. We cut off the top 100 such patents per publication and then count the number of these "top 100-matches" per patent ("WoS Similarity"). We then identify Soviet-era journals from the ISF's eligible publications list from Gangui (2015) (reproduced in appendix C.2 below)) and count the number of times patents are ranked top 100 similar to the articles in these journals ("Soviet Similarity"). This allows us to calculate a measure of how similar a patent i is to a Soviet publication normalized by how similar it is to scientific publications in general:

$$Normalized\ Soviet\ Similarity_i = \frac{Soviet\ Similarity_i}{WoS\ Similarity_i}$$

We define a patent to be in the "treatment" group if $Normalized\ Soviet\ Similarity_i > 0.5$. In unreported regressions, we change the threshold value from 0.5 to 0.9, and are able to replicate all results in table 5.

C.1 Comparing Soviet Patents to Non Soviet Patents

We also compare the relative frequencies with which "Soviet" patents occur at the 3 digit IPC level and do not find a systematic concentration in certain technical areas. Table C1 counts the number of Soviet and non-Soviet patents in each 3 digit IPC, downward sorts them and shows the five most frequent and five least frequent IPC classes. If Soviet similarity would vary by technology class, then we would expect to see the rankings to show no overlap between the two — we find the opposite: two out of the top and bottom five IPCs are common in both Soviet and non-Soviet tables.

Table C1: Number of Patents per 3-digit IPCs, by Soviet Similarity

Similar to Soviet Science			
Rank	3 Digit IPC	No. of Patents	Description
1	H01	54516	Basic electric elements
2	B60	22217	Vehicles in general
3	H04	21166	Electric communication technique
4	G06	18293	Computing; calculating; counting
5	G11	17412	Information storage
...	...	...	...
117	C40	19	Combinatorial technology
118	C05	8	Fertilisers; manufacture thereof
119	C14	8	Skins; hides; pelts; leather
120	B82	5	Nanotechnology
121	C13	3	Sugar industry
Not Similar to Soviet Science			
Rank	3 Digit IPC	No. of Patents	Description
1	H01	307002	Basic electric elements
2	G06	278471	Computing; calculating; counting
3	A61	266123	Medical or veterinary science; hygiene
4	H04	231120	Electric communication technique
5	G01	227601	Measuring; testing
...	...	...	...
117	C13	460	Sugar industry
118	D07	431	Ropes; cables other than electric
119	C14	427	Skins; hides; pelts; leather
120	B82	370	Nanotechnology
121	G12	363	Instrument details

Notes: This table ranks the number of patents per 3 digit IPC class by whether they are close to a Soviet scientific publication.

C.2 Soviet Journal List (International Science Foundation)

Journal Name	Number of Articles
ACOUSTICAL PHYSICS	3,031

ANTIBIOTIKI I KHIMIOTERAPIYA	797
ARKHIV PATOLOGII	3,016
ASTRONOMICHESKII ZHURNAL	3,910
BIOFIZIKA	7,130
BIOKHIMIYA	1,601
BIOLOGICHESKIE MEMBRANY	2,784
BIOLOGIYA MORYA-MARINE BIOLOGY	1,293
BIOORGANICHESKAYA KHIMIYA	4,725
COMBUSTION EXPLOSION AND SHOCK WAVES	6,241
DIFFERENTIAL EQUATIONS	7,529
DOKLADY AKADEMII NAUK	10,801
DOKLADY AKADEMII NAUK SSSR	93,913
EURASIAN SOIL SCIENCE	4,477
FARMAKOLOGIYA I TOKSIKOLOGIYA	3,699
FIZIKA METALLOV I METALLOVEDENIE	10,246
FIZIKA NIZKIKH TEMPERATUR	3,495
FIZIKA TVERDOGO TELA	19,093
FIZIOLOGICHESKII ZHURNAL	1,543
FUNCTIONAL ANALYSIS AND ITS APPLICATIONS	2,156
GENETIKA	7,089
GEOKHIMIYA	5,309
GEOMAGNETIZM I AERONOMIYA	5,424
GEOTECTONICS	1,326
HIGH TEMPERATURE	9,048
INORGANIC MATERIALS	16,135
IZVESTIYA AKADEMII NAUK FIZIKA ATMOSFERY I OKEANA	952
IZVESTIYA AKADEMII NAUK SSSR SERIYA BIOLOGICHESKAYA	2,621
IZVESTIYA AKADEMII NAUK SSSR SERIYA FIZICHESKAYA	11,877
IZVESTIYA AKADEMII NAUK SSSR SERIYA GEOLOGICHESKAYA	1,970
IZVESTIYA SIBIRSKOGO OTDELENIYA AKADEMII NAUK SSSR SERIYA KHIMICHESKIKH NAUK	1,333
IZVESTIYA VYSSHIKH UCHEBNYKH ZAVEDENII FIZIKA	9,973
IZVESTIYA VYSSHIKH UCHEBNYKH ZAVEDENII KHIMIYA I KHIMICHESKAYA TEKHNOLOGIYA	5,147
IZVESTIYA VYSSHIKH UCHEBNYKH ZAVEDENII RADIOFIZIKA	2,672
JETP LETTERS	14,958
JOURNAL OF ANALYTICAL CHEMISTRY OF THE USSR	6,866
JOURNAL OF EVOLUTIONARY BIOCHEMISTRY AND PHYSIOLOGY	2,276
JOURNAL OF MICROBIOLOGY EPIDEMIOLOGY AND IMMUNOBIOLOGY USSR	431
KARDIOLOGIYA	9,861
KHIMICHESKAYA FIZIKA	2,485
KHIMIKO-FARMATSEVTICHESKII ZHURNAL	7,516
KINETICS AND CATALYSIS	8,095
KOLLOIDNYI ZHURNAL	697
KOORDINATSIONNAYA KHIMIYA	2,377
KOSMICHESKAYA BIOLOGIYA I AVIAKOSMICHESKAYA MEDITSINA	2,421

KRISTALLOGRAFIYA	6,531
KVANTOVAYA ELEKTRONIKA	9,749
KYBERNETIKA	2,179
MATHEMATICAL NOTES	8,376
MEASUREMENT TECHNIQUES	9,433
MIKOLOGIYA I FITOPATOLOGIYA	2,739
MIKROBIOLOGIYA	733
MOLEKULYARNAYA BIOLOGIYA	293
NEUROPHYSIOLOGY	2,290
OKEANOLOGIYA	4,137
OPTIKA I SPEKTROSKOPIYA	15,537
PARAZITOLOGIYA	1,469
PETROLEUM CHEMISTRY	3,025
PISMA V ZHURNAL TEKHNICHESKOI FIZIKI	5,682
PRIBORY I TEKHNIKA EKSPERIMENTA	2,669
PRIKLADNAYA MATEMATIKA I MEKHANIKA	582
RADIOCHEMISTRY	926
RADIOTEKHNIKA I ELEKTRONIKA	10,540
RUSSIAN JOURNAL OF INORGANIC CHEMISTRY	5,792
RUSSIAN JOURNAL OF PLANT PHYSIOLOGY	2,804
RUSSIAN MATHEMATICAL SURVEYS	4,004
SBORNIK MATHEMATICS	1,728
SEMICONDUCTORS	6,876
SIBERIAN MATHEMATICAL JOURNAL	5,034
TEORETICHESKAYA I EKSPERIMENTALNAYA KHIMIYA	1,297
TERAPEVTICHESKII ARKHIV	14,261
THEORETICAL AND MATHEMATICAL PHYSICS	6,590
THEORY OF PROBABILITY AND ITS APPLICATIONS	3,612
UKRAINSKII BIOKHIMICHESKII ZHURNAL	2,914
UKRAINSKII FIZICHESKII ZHURNAL	4,436
USPEKHI FIZICHESKIKH NAUK	4,148
USPEKHI KHIMII	3,221
VESTNIK AKADEMII MEDITSINSKIKH NAUK SSSR	2,218
VESTNIK AKADEMII NAUK SSSR	3,989
VESTNIK MOSKOVSKOGO UNIVERSITETA SERIYA 1 MATEMATIKA MEKHANIKA	3,182
VESTNIK MOSKOVSKOGO UNIVERSITETA SERIYA 2 KHIMIYA	4,187
VESTNIK MOSKOVSKOGO UNIVERSITETA SERIYA 3 FIZIKA ASTRONOMIYA	2,656
VOPROSY MEDITSINSKOI KHIMII	4,291
VOPROSY ONKOLOGII	4,585
VOPROSY VIRUSOLOGII	4,206
VYSOKOMOLEKULYARNYE SOEDINENIYA SERIYA A	7,905
VYSOKOMOLEKULYARNYE SOEDINENIYA SERIYA B	5,229
ZHURNAL EKSPERIMENTALNOI I TEORETICHESKOI FIZIKI	11,724
ZHURNAL FIZICHESKOI KHIMII	26,686
ZHURNAL OBSHCHEI BIOLOGII	2,923
ZHURNAL OBSHCHEI KHIMII	30,563

ZHURNAL ORGANICHESKOI KHIMII	16,450
ZHURNAL TEKHNICHESKOI FIZIKI	13,372
ZHURNAL VYSSHEI NERVNOI DEYATELNOSTI IMENI I P PAVLOVA	6,265
ZOOLOGICHESKY ZHURNAL	8,100

Notes: The list comprises prominent Soviet and Russian journals in which eligible ISF applicants needed to show publication records for (Ganguli, 2017). The number of articles is calculated by the authors using Web of Science data.

D Assigning geolocational data to patents

We use the address of an inventor to determine the location of a patent. If there are more than two inventors, we take the address that appears the most frequent as the location of the patent (majority vote). In case of a tie, we randomly take one of the addresses as the patent's location. We source the inventor addresses from the HBS inventor dataset, which contains city and state information for inventors for all U.S. patents from 1975 to 2010. Since cities may often be too small and states too large to gauge the impact of immigration on the cost of knowledge transfer for buyers and sellers of patents, we classify each city that appears in the HBS dataset into Metropolitan Statistical Areas (MSA) and Primary Metropolitan Statistical Areas (PMSA) current as of 1990. Specifically, we download the historical delineation files for the 1990 Decennial Census from the Census TIGER database.[26] We also download delineation files for U.S. cities from the ESRI USA data available from Baruch's Geoportal website.[27] We then use the "spatial join" feature in ArcGIS Pro in order to determine which cities lie within our 179 MSAs and PMSAs. Cities that are not classified within an MSA or PMSA are classified by their state e.g., "Not in MSA/PMSA California." Patents are then classified into one of the MSA/PMSAs based on the majority rule described above.

[26]https://www.census.gov/geo/maps-data/data/tiger-line.html

[27]https://www.baruch.cuny.edu/confluence/display/geoportal/ESRI+USA+Data

E Comparison of Patents by Content Similarity to Soviet Science

Table E1: Summary Statistics for Patents Close to Soviet Science

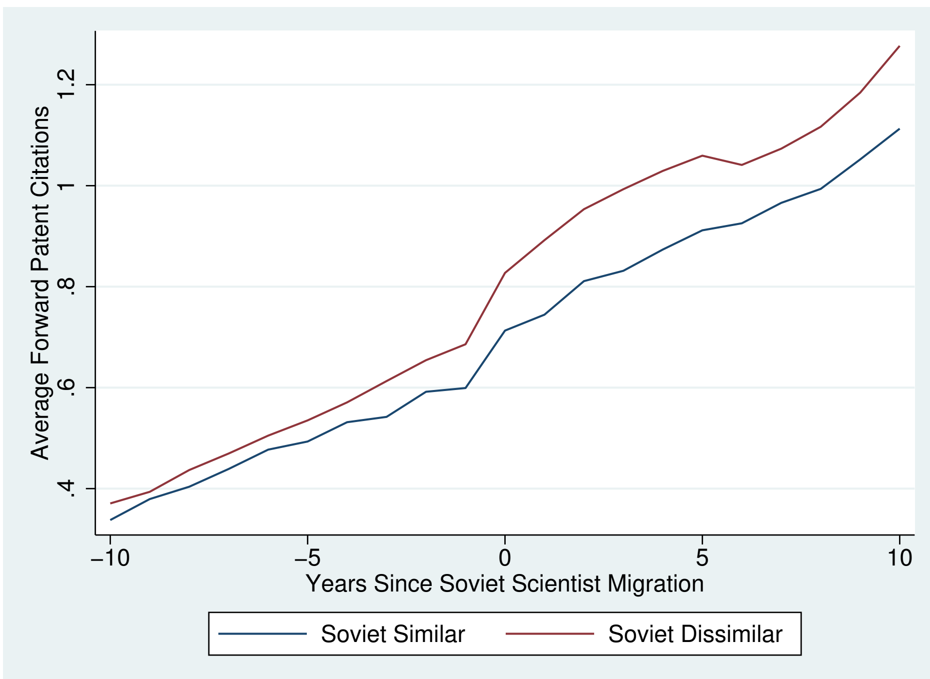
	count	mean	sd	p10	p50	p90
Patent Publication Year	25911	1986.863	4.710	1981	1987	1993
Reassignment	25911	0.095	0.293	0	0	0
Cite Science	25911	0.029	0.167	0	0	0
Forward Patent Citations (Normalized)	25911	0.977	1.179	0	1	2
5-year Forward Patent Citations	19424	3.967	4.557	1	3	8
Stock Market Value of Patent	9976	7.761	10.927	1	4	18
Combination Familiarity	25900	66.057	141.791	0	3	209
Small Entity	20028	0.376	0.484	0	0	1
Compustat Patent	25911	0.373	0.484	0	0	1
Triadic Patent	25911	0.167	0.373	0	0	1

Table E2: Summary Statistics for Patents NOT Close to Soviet Science

	count	mean	sd	p10	p50	p90
Patent Publication Year	94827	1986.960	4.913	1981	1987	1993
Reassignment	94827	0.080	0.272	0	0	0
Cite Science	94827	0.040	0.195	0	0	0
Forward Patent Citations (Normalized)	94827	1.085	1.320	0	1	2
5-year Forward Patent Citations	72987	4.619	5.603	1	3	10
Stock Market Value of Patent	48144	9.108	16.324	1	5	20
Combination Familiarity	94794	50.750	129.224	0	1	153
Small Entity	72285	0.260	0.439	0	0	1
Compustat Patent	94827	0.502	0.500	0	1	1
Triadic Patent	94827	0.217	0.412	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 MAG, and zero otherwise. *Forward Patent Citations (Normalized)* counts the number of prior art citations the patent has received until 2015 and normalizes this quantity by the average number of forward patent citations received by patents published in the focal patent's publication year. *5-year Forward patent citations* counts the number of prior art citations the patent has received within five years of its publication. *Stock Market Value of Patent* is based on the cumulative abnormal returns in the firm's market value at the issuance event of the patent per Kogan et al. (2017). *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. *Triadic Patent* is equal to one if the patnet shares a prior art in the USPTO, EPO, and JPO, and zero otherwise. Treatment Group is defined as patents with a normalized similarity score to Soviet journals above 0.5 (details on classification method in section 5).

Figure E1: Forward Patent Citations by Treatment Group



Notes: This graph plots forward patent citations per patent received by patents in the treatment and control group for 10 years before and after the first Soviet scientist migrant arrived in the focal patent's MSA. The timing and location of migrants are based on data from Ganguli (2015)