Volume Title: The Economics of Artificial Intelligence: An Agenda

Volume Authors/Editors: Ajay Agrawal, Joshua Gans, and Avi Goldfarb, editors

Volume Publisher: University of Chicago Press

Volume ISBNs: 978-0-226-61333-8 (cloth); 978-0-226-61347-5 (electronic)

Volume URL: http://www.nber.org/books/agra-1

Conference Date: September 13–14, 2017

Publication Date: May 2019

Chapter Title: Artificial Intelligence and International Trade

Chapter Author(s): Avi Goldfarb, Daniel Trefler

Chapter URL: http://www.nber.org/chapters/c14012

Chapter pages in book: (p. 463 – 492)

19

Artificial Intelligence and International Trade

Avi Goldfarb and Daniel Trefler

The last 200 years have produced a remarkable list of major innovations, not the least of which is artificial intelligence (AI). Like other major innovations, AI will likely raise average incomes and improve well-being, but it may also disrupt labor markets, raise inequality, and drive noninclusive growth. Yet, even to the extent that progress has been made in understanding the impact of AI, we remain largely uninformed about its international dimensions. This is to our great loss. A number of countries are currently negotiating international agreements that will constrain the ability of sovereign governments to regulate AI, such as the North American Trade Agreement (NAFTA) and the Trans-Pacific Partnership (TPP)-11. Likewise, governments around the world are freely spending public funds on new AI clusters designed to shift international comparative advantage toward their favored regions, including the Vector Institute in Toronto and the Tsinghua-Baidu deep-learning lab around Beijing. The international dimensions of AI innovations and policies have not always been well thought out. This work begins the conversation.

China has become the focal point for much of the international discussion. The US narrative has it that Chinese protection has reduced the ability

Avi Goldfarb holds the Rotman Chair in Artificial Intelligence and Healthcare and is professor of marketing at the Rotman School of Management, University of Toronto, and a research associate of the National Bureau of Economic Research. Daniel Trefler holds the J. Douglas and Ruth Grant Canada Research Chair in Competitiveness and Prosperity at the University of Toronto and is a research associate of the National Bureau of Economic Research.

The authors thank Dave Donaldson and Hal Varian for their thoughtful feedback. Trefler acknowledges the support of the "Institutions, Organizations and Growth" Program of the Canadian Institute for Advanced Research (CIFAR). For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see http://www.nber.org/chapters/c14012.ack.

of dynamic US firms such as Google and Amazon to penetrate Chinese markets. This protection has allowed China to develop significant commercial AI capabilities, as evidenced by companies such as Baidu (a search engine like Google), Alibaba (an e-commerce web portal like Amazon), and Tencent (the developer of WeChat, which can be seen as combining the functions of Skype, Facebook, and Apple Pay). While no Chinese AI-intensive company has household recognition outside of China, everyone agrees that this will not last. Further, a host of behind-the-border regulatory asymmetries will help Chinese firms to penetrate Canadian and US markets.

Even the Pentagon is worried. Chinese guided-missile systems are sufficiently sophisticated that they may disrupt how we think of modern warfare; large and expensive military assets such as aircraft carriers are becoming overly vulnerable to smart weapons.[1] This may do more than transform the massive defense industry; these AI developments may radically shift the global balance of power.

As international economists, we are used to hype and are typically dismissive of it. Despite AI's short life—Agrawal, Gans, and Goldfarb (2018) date its commercial birth to 2012—AI's rapid insinuation into our daily economic and social activities forces us to evaluate the international implications of AI and propose best-policy responses. Current policy responses often rest on a US narrative of a zero-sum game in which either the United States or China will win.[2] Is this the right premise for examining AI impacts and for developing AI policies? Further, calls for immediate action by prominent experts such as Bill Gates, Stephen Hawking, and Elon Musk will likely encourage governments to loosen their pocketbooks, but will government subsidies be effective in promoting broad-based prosperity or will subsidies become yet another form of ineffective corporate welfare? What specific policies are likely to tip the balance away from ineffective corporate handouts?

Using comparative advantage theory, trade economists have thought long and hard about the right mix of policies for successfully promoting industry. Many of our theories imply a laissez-faire free-trade approach. However, since the early 1980s our theories have shown that certain types of government interventions may be successful, for example, Krugman (1980), Grossman and Helpman (1991), and the more informal theories of Porter (1990). These theories emphasize the role of scale and the role of knowledge creation and diffusion. Unfortunately, the precise policy prescriptions produced by these theories are very sensitive to the form of scale and the form

1. *New York Times*, Feb. 3, 2017. See also *Preparing for the Future of Artificial Intelligence*, Office of the President, Oct., 2016.

2. For example, https://www.economist.com/news/business/21725018-its-deep-pool-data-may-let-it-lead-artificial-intelligence-china-may-match-or-beat-america and http://www.reuters.com/article/us-usa-china-artificialintelligence/u-s-weighs-restricting-chinese-investment-in-artificial-intelligence-idUSKBN1942OX?il=0.

of knowledge creation/diffusion. And competition can play an important role too, for example, in Aghion et al. (2001, 2005) and Lim, Trefler and Yu (2017).

We therefore start in section 19.2 by identifying the key features of AI technology in regard to scale and knowledge. To date there are no models that feature the particular scale and knowledge characteristics that are empirically relevant for AI. In section 19.3 we use these features (a) to offer some suggestions for what an appropriate model might look like, and (b) to draw implications for policy. This leads to high-level thinking about policy. For example, it provides a foundation for assessing recent proposals put forward by AI researcher Geoff Hinton and others on the potential benefit of public investments in AI.[3] However, these models are not sufficiently fine-grained to directly capture existing regulatory issues that "go behind the border" such as privacy policy, data localization, technology standards, and industrial regulation. In section 19.4 we therefore review the many behind-the-border policies that already impact AI and discuss their implications for comparative advantage and the design of trade agreements. We begin with a factual overview of the international dimensions of AI.

19.1 From Hype to Policy

Statistics about where AI is being done internationally and how it is diffusing can be tracked in a number of ways, for example, the number of basic research articles, patents and patent citations produced in a region; the number of start-ups established in a region; or the market capitalization of publicly traded AI-based companies in a region. We look at two of these indicators: basic research and market capitalization. For the former, we collected time-series data on the institutional affiliation of all authors of papers presented at a major AI research conference, namely, the Association for the Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence. In table 19.1, we compare the 2012 and 2017 conferences. In 2012, 41 percent of authors were at US institutions, but by 2017 this was down to 34 percent. The two other largest declines were recorded by Canada and Israel. While these countries all increased their absolute number of participants, in relative terms they all lost ground to China, which leapt from 10 percent in 2012 to 23 percent in 2017.

We have not examined patent numbers, but suggestive work by Fujii and Managi (2017) points to weaker international diffusion of AI: US technology giants such as IBM and Microsoft remain far and away the world's dominant patent applicants.

Another indication of the economic future of AI comes from the largest

3. "Artificial Intelligence is the Future, and Canada Can Seize It" by Jordan Jacobs, Tomi Poutanen, Richard Zemel, Geoffrey Hinton, and Ed Clark. *Globe and Mail*, Jan. 7, 2017.

Table 19.1 Participants at a major AI conference

Country	2012 (%)	2017 (%)	Change (%)
United States	41	34	–6
China	10	23	13
United Kingdom	5	5	0
Singapore	2	4	2
Japan	3	4	1
Australia	6	3	–2
Canada	5	3	–3
India	1	2	1
Hong Kong	3	2	–1
Germany	4	2	–1
France	4	2	–2
Israel	4	2	–3
Italy	2	2	–1
Other	10	10	0

Notes: Participation rates at the Association for the Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence. For example, of the papers presented at the 2017 conference, 34 percent of authors had a US affiliation.

public companies in the world by market capitalization. Table 19.2 lists the twelve largest companies worldwide. What is striking about the table is the number of companies that might subjectively be described as "AI intensive." Seven of the twelve companies are heavily engaged in AI (such as Alphabet/Google), three are in finance (where the use of AI is growing rapidly), and one has a substantial pharmaceutical presence (where AI is likely to soon be reducing development costs). What makes table 19.2 relevant for international trade is the fact that two of the largest companies worldwide are now Chinese AI-intensive firms (Tencent and Alibaba). It is truly remarkable that two high-tech companies based out of China—private companies, not state-owned enterprises—are among the largest companies in the world. While we had to move beyond the round number of ten to make this point, it is striking nonetheless. It points to the major global shake-up that is coming.

Some would conclude from tables 19.1 and 19.2 that almost all of the world's largest companies will soon be competing directly against Chinese companies when—not if—these Chinese companies go global. In 2000, Robin Li signaled his agreement by moving to China to establish Baidu. The flood of US-trained talent returning to China has continued. This year, former Microsoft executive Qi Lu joined Baidu as chief operating officer (COO). In describing China, Lu writes, "We have an opportunity to lead in the future of AI."[4] Not everyone agrees. Some have argued that China's AI-intensive companies will not be globally competitive until they compete head-on in China with global leaders such as Google. This flies in the face of

4. *The Economist*, July 15, 2017.

Table 19.2 World's largest public companies and AI exposure

Company	Market value ($)	AI exposure
1. Apple	754	High
2. Alphabet	579	High
3. Microsoft	509	High
4. Amazon	423	High
5. Berkshire Hathaway	411	Rising
6. Facebook	411	High
7. ExxonMobil	340	Low
8. Johnson & Johnson	338	Rising
9. JPMorgan Chase	314	Rising
10. Wells Fargo	279	Rising
11. Tencent Holdings	272	High
12. Alibaba	269	High

Notes: Market capitalization of the largest public companies as of March 31, 2017, from PWC (2017). "AI exposure" is our subjective assessment of the role of AI in company performance.

a long history of Chinese export successes in other fields. Indeed, Sutton and Trefler (2016) describe both theoretically and empirically how developing countries such as China initially enter new markets at a low level of quality, but over time develop the capabilities to deliver high-quality, internationally competitive goods and services.

Many experts are weighing in on how to counter the "Chinese threat" and, more generally, how to enrich local economies through cluster policies that support sustained competitive advantage in AI-based market segments. Geoff Hinton and collaborators have convinced Canadian governments to develop a major AI institute that would "graduate the most machine-learning PhDs and master's students globally" and "become the engine for an AI supercluster that drives the economy of Toronto, Ontario, and Canada."[5] Hinton also emphasizes the importance of access to data. "Why? Because for a machine to 'think' intelligently, it must be trained with lots of data."

While there are potential benefits from Hinton's initiative, it raises two important points that loom large in our thinking. First, economists who specialize in clusters are deeply skeptical about the efficacy of cluster policies (e.g., Duranton 2011). Such policies have failed more often than not, and the theoretical justification for cluster policies is highly sensitive to assumptions about knowledge diffusion. For example, will Hinton's PhDs stay in Canada and will the knowledge they generate be commercialized in Canada? Second, a host of behind-the-border regulations on privacy, data localization, technology standards, and industrial policy will affect the ability of Canadian firms to access data relative to their competitors in larger markets such as the

5. *Globe and Mail*, Jan. 7, 2017.

United States, Europe, and China. What is the current state of these domestic data regulations, how do they effect trade patterns, do they serve a public interest, are they being used as disguised protection to generate comparative advantage, and should they be covered by international trade agreements (as some would have been in the TPP e-commerce chapter)?

The following sections help answer these questions and move us toward better policies for promoting AI and preventing both corporate welfare and welfare-reducing disguised protection.

19.2 The Technological Backdrop: Scale, Scope, Firm Size, and Knowledge Diffusion

The Oxford English Dictionary defines AI as "the theory and development of computer systems able to perform tasks normally requiring human intelligence." This has meant different things at different times. In the 1960s and 1970s, computer scientists approached this using rules, if-then statements, and symbolic logic. It worked well for factory robots and for playing chess. By the 1980s, it became clear that symbolic logic could not deal with the complexities of nonartificial settings, and AI research slowed substantially. Various approaches continued to be supported in a small number of locations, including by the Canadian Institute for Advanced Studies (CIFAR).

The recent resurgence in AI research is driven by one such approach: the insight that computers can "learn" from example. This approach is often called "machine learning" and is a field of computational statistics. The algorithm that has received the most attention is back propagation in neural networks, most notably through "deep learning," but there is a large suite of relevant technologies including deep learning, reinforcement learning, and so forth. Because the current excitement about AI is driven by machine learning, we focus on this particular set of algorithms here.

For our purposes, we need to zero in on those aspects of AI technology that are central to thinking about the economics of AI. We identify four aspects: economies of scale associated with data, economies of scale associated with an AI research team, economies of scope in the use of the team for multiple applications, and knowledge externalities.

19.2.1 Economies of Scale from Data

Statistical predictions improve with the quantity and quality of data. Recall from statistics 101 that the quality of prediction increases with N (or, more precisely with root N). All else being equal, this means that companies that have more observations will generate more accurate predictions. It is in this sense that economies of scale matter. Still, because predictions increase in root N, then, while scale matters, there are decreasing returns to scale in terms of the accuracy of prediction.

It is subtler than this, however. Google and Microsoft both operate search engines. Google has claimed their search engine has higher market share because it has better quality.[6] Microsoft has claimed the higher quality is a direct consequence of scale. By having more data, Google can predict what people want in their search results more accurately. Google responds that Microsoft has billions of search results. While Google has more data, surely the law of large numbers applies before one billion results. And so, more data does not give a meaningful advantage. Microsoft's response is the essence of where economies of scale bind. While they have billions of searches, many search queries are extremely rare. Microsoft may only see two or three, and so Google can predict those rare queries much better. If people choose search engines based on quality differences in rare searches, then Google's better data will lead to a substantial increase in market share. Having a larger share gives Google more data, which in turn improves quality and supports an even larger share.

The source of economies of scale here is therefore in the form of direct network externalities. More customers generate more data, which in turn generates more customers. This is different from the literature on two-sided markets and indirect network externalities. The network externalities resemble the phone network, rather than externalities between buyers and sellers on a marketplace like Ebay. This is significant in a trade context because the trade literature has emphasized two-sided matching, for example, in Rauch (1999) and McLaren (2000). This is also different from all of the trade and market structure literature, which emphasize economies of scale that are driven by fixed costs, so trade theory does not currently have models that are applicable to the AI technology environment.

The direct network externalities environment leads to a core aspect of competition in AI: competition for data. The companies that have the best data make better predictions. This creates a positive feedback loop so that they can collect even more data. In other words, the importance of data leads to strong economies of scale.

19.2.2 Economies of Scale from the Overhead of Developing AI Capabilities

Another source of economies of scale in AI involves the fixed cost of building an AI capability within a firm. The main cost is in personnel. Much of the software is open source, and in many cases hardware can be purchased as a utility through cloud services. The uses of AI need to be big enough to justify the substantial cost of building a team of AI specialists. World leaders in AI command very high pay, often in the millions or tens of millions.

6. There is a chicken and egg problem, whether good algorithms drive market share or whether market share drives hiring that leads to better algorithms. For one point of view, see https://www.cnet.com/news/googles-varian-search-scale-is-bogus/.

Top academic researchers have been hired to join Google (Hinton), Apple (Salakhutdinov), Facebook (LeCunn), and Uber (Urtasun). So far, there has been a meaningful difference between employing the elite researchers and others in terms of the capabilities of the AI being developed.

19.3.3 Economies of Scope

Perhaps more than economies of scale, the fixed cost of building an AI capacity generates economies of scope. It is only worth having an AI team within a company if there are a variety of applications for them to work on. Many of the currently leading AI firms are multiproduct firms. For example, Google parent Alphabet runs a search engine (Google), an online video service (YouTube), a mobile device operating system (Android), an autonomous vehicle division (Waymo), and a variety of other businesses. In most cases, the economies of scope happen on the supply side through AI talent, better hardware, and better software.

Another important source of economies of scope is the sharing of data across applications. For example, the data from Google's search engine might be valuable in helping determine the effectiveness of YouTube advertising, or its mapping services might be needed for developing autonomous vehicles. The sharing of data is a key source of international friction on disguised protection behind the border. Differences in privacy policies mean that it is easier to share data across applications in some countries compared to others. For example, when Ebay owned PayPal, it faced different restrictions for using the PayPal data in Canada compared to the United States. We will return to this subject later.

This contrasts with the main emphasis in the trade literature on economies of scope, which emphasizes the demand side. Economies of scope in AI do not seem to be about demand externalities in brand perception or in sales channels. Instead, they appear to be driven by economies of scope in innovation. A wider variety of potential applications generates greater incentives to invest in an AI research team, and it generates more benefits to each particular AI project due to the potential to share data across applications.

19.3.4 Knowledge Externalities

There is a tension in discussing knowledge diffusion in the AI sphere. On the one hand, the spectacular scientific advances are often taught at universities and published in peer-reviewed journals, providing businesses and government personnel with quick and easy access to frontier research. Further, there is the migration of personnel across regions and countries as the above examples of Robin Li and Qi Lu show. This suggests that knowledge externalities are global in scope.

On the other hand, AI expertise has also tended to agglomerate in several narrowly defined regions globally. As with other information technologies, much of the expertise is in Silicon Valley. Berlin, Seattle, London, Boston, Shanghai, and to some extent Toronto and Montreal can all claim to be hubs of AI innovation. This suggests that AI involves a lot of tacit knowledge that is not easily codified and transferred to others.

In fact, the traditional discussion of knowledge externalities takes on a more nuanced hue in the context of AI. Can these researchers communicate long distance? Do they have to be together? How important are agglomeration forces in AI? As of 2017, AI expertise remains surprisingly rooted in the locations of the universities that invented the technologies. Google's DeepMind is in London because that is where the lead researcher lived. Then the first expansion of DeepMind outside the United Kingdom was to Edmonton, Alberta, because Richard Sutton, a key inventor of reinforcement learning, lives in Edmonton. Uber opened an AI office in Toronto because it wanted to hire Raquel Urtasun, a University of Toronto professor.

Generally, there are a small number of main AI research departments: Stanford, Carnegie Mellon University, the University of Toronto, and several others. Their location is often surprisingly disconnected from headquarters, and so companies open offices where the talent is rather than forcing the talent to move to where the company is.

As we shall see, the exact nature of knowledge externalities is terribly important for understanding whether cluster and other policies are likely to succeed. The nature of these externalities also has some unexpected implications such as the implications of noncompete clauses (Saxenian 1994) and the asymmetries in access to knowledge created by asymmetries in who can speak English versus who can speak Chinese versus who can speak both.

19.3 Trade Theory and the Case for Industrial and Strategic Trade Policies

There are many voices in the industrialized world arguing for industrial policies and strategic trade policies to promote rising living standards. Many of these voices point to the achievements of China as an example of what is possible. Much of what is claimed for China, and what was once claimed for Japan, is of dubious merit. China has redirected vast resources from the rural poor and urban savers toward state-owned enterprises that have massively underperformed. Those firms continue to be major players in the economy and a major drag on economic growth (Brandt and Zhu 2000). It is thus significant that China's greatest commercial successes in AI have come from private companies. So if we are to make the case for industrial and strategic trade policies, we cannot blithely appeal to Chinese state-directed successes. Rather, we must understand the characteristics of industries that increase the likelihood that government policy interventions will be successful.

To this end, we start with a vanilla-specific factors model of international trade (Mussa 1974; Mayer 1974) in which the case for departures from free trade is weak. We then add on additional elements and examine which of these is important for policy success. The first conclusion is that scale and knowledge externalities are critical. The second is that these two elements alone are not enough: their precise form also matters.

19.3.1 Scientists, Heterogeneous Scientists, and Superstar Scientists

Many factors enter into the location decisions of AI firms including access to local talent, local financing/management, and local markets. In this section, we focus on the role of university-related talent. Among the participants of this conference are three head researchers at top AI companies: Geoffrey Hinton (University of Toronto and Google), Russ Salakhutdinov (Carnegie Mellon University and Apple), and Yann LeCun (New York University and Facebook). Each joined his company while retaining his academic position, and each continues to live near his university rather than near corporate headquarters. These three examples are not exceptional, as indicated by the above examples of DeepMind and Richard Sutton, and Raquel Urtasun and Uber.

Scientists. We begin with the simplest model of trade that allows for two types of employees, scientists, and production workers. There are two industries, search engines and clothing. Production workers are employed in both industries and move between them so that their wages are equalized across industries. Scientists are "specific" to the search engine industry in that they are very good at AI algorithms and useless at sewing. We also assume that scientists and workers cannot migrate internationally. Then it is immediately obvious that the more scientists a country has, the larger will be both the size and service exports of the search engine industry.

We start with this benchmark model because, in this setting, without scale or externalities there is no scope for market failure and hence *there is no simple case for any trade policy other than free trade*. For example, consider a policy of restricting imports of search engine services, as China has done with Google. This restriction helps Chinese scientists but can hurt Chinese production workers and consumers (Ruffin and Jones 1977).

There are several departures from this benchmark model that lead to welfare-enhancing export subsidies and other departures from free trade. As we shall see, the two most important are economies of scale and knowledge creation. However, we start instead with profits because profits are at the core of arguments supporting strategic trade policies (Krugman 1986). Since there are no profits in the specific factors model, we introduce profits by introducing scientists of heterogeneous quality.

Heterogeneous Scientists. Consider an industry in which firms provide a search engine and generate advertising revenue. There is a continuum of scientists distinguished by their "quality" q. A firm is distinguished by the quality of its chief scientist and hence firms are also indexed by q. A higher-quality scientist produces a better search engine. A firm engages in activity a that increases advertising revenues $r(a)$ where $r_a > 0$. Let $p(q)$ be the proportion of consumers who choose firm q's search engine. It is natural to assume that $p_q > 0$ that is, a better scientist produces a more desirable search engine. The firm's profit before payments to the scientist is $\pi(a,q) = p(q)r(a) - c(a)$ where $c(a)$ is the cost of the firm's ad-generating activity. In this model the firm is essentially the scientist, but we can delink the two by assuming that the scientist is paid with stock options and so receives a fraction $(1 - \mu)$ of the profits. It is straightforward to show that profit $\pi(a,q)$ is supermodular in (a,q). This implies positive assortative matching; firms with better scientists engage in more ad-generating activity. This means that firms with better scientists will also have more users ($p_q > 0$), more revenues $[\partial r(a(q),q)/\partial q > 0]$, and higher profits $[\partial\pi(a(q),q)/\partial q > 0]$. Putting these together, better scientists anchor bigger and more profitable firms.[7]

To place this model into an international-trade setting, we assume that there are multiple countries, a second constant-returns-to-scale industry (clothing), and no international migration of scientists or workers. Because there are profits in the search engine industry, policies that expand that industry generate higher profits. This is the foundation of *strategic trade policy*. In its simplest form, if there are supernormal profits then tariffs and other trade policies can be used to shift profits away from the foreign country and to the domestic country.

Strategic trade policy was first developed by Brander and Spencer (1981) and variants of it have appeared in many of the models discussed below. Unfortunately, the case for strategic trade policy is not as clear as it might seem. Its biggest logical problem is the assumption of positive profits: if there is free entry, then entry will continue until profits are driven to zero.[8] This means that any government policy that encourages entry of firms or training of scientists will be offset by inefficient entry of firms or scientists. Put simply, strategic trade policies only work if there are profits, but with free entry there are no profits (see Eaton and Grossman 1986). The conclusion we draw from this is that the model needs enriching before it can be used to justify trade policy.

Before enriching the model, we note that there are two other compelling

7. The first-order condition for advertising activities is $\mu\pi_a = \mu(pr_a - c_a) = 0$. We assume that the second-order condition is satisfied: $\mu\pi_{aa} < 0$. Supermodularity is given by $\partial^2\mu\pi(a,q)/\partial a\partial q = p_q r_a > 0$. The result that advertising activity levels $a(q)$ are increasing in q comes from differentiating the first-order condition: $\mu p_q r_a + \mu\pi_{aa}a_q = 0$ or $a_q = -p_q r_a/\pi_{aa} > 0$. The result that average revenues $p(q)r(a)$ are increasing in q follows from $\partial p(q)r(a(q))/\partial q = p_q r + pr_a a_q > 0$. The result that profits $\pi(a(q),q)$ are increasing in q follows from $\partial\mu\pi(a,q)/\partial q = \mu\pi_a a_q + \mu p_q r(a) = \mu p_q r(a) > 0$ where we have used the first-order condition ($\pi_a = 0$).

8. Free entry implies that ex ante profits are zero. Of course, ex post profits (operating profits of survivors) are always positive; otherwise, survivors would exit.

reasons for being skeptical about the efficacy of strategic trade policy. First, such policies set up political economy incentives for firms to capture the regulatory process used to determine the amount and form of government handouts. Second, the logic of strategic trade policy fails if there is retaliation on the part of the foreign government. Retaliation generates a trade war in which both countries lose. Artificial intelligence meets all the conditions that Busch (2001) identifies as likely to lead to a trade war. We now turn to enriching our model.

Superstar Scientists.[9] Strategic trade policies are more compelling in settings where scale and/or knowledge creation and diffusion are prevalent. To this end we follow section 19.2 in assuming that there are economies of scale in data. This will cause the market to be dominated by a small number of search engine firms; that is, it will turn our model into something that looks like a superstar model. To be more precise, it is a little different from standard superstar models that make assumptions on the demand side (Rosen 1981). The superstar assumptions here are on the supply side.

Modifying our model slightly, we introduce scale in data by assuming that the share of consumers choosing a search engine ($p(q)$) is increasing at an increasing rate ($p_{qq} > 0$);[10] $p_{qq} > 0$ implies that profits and scientist earnings increase at an increasing rate, that is, they are convex in q.[11] This, in turn, implies that the distribution of firm size becomes highly skewed toward large firms. It also implies that the shareholders of large firms will make spectacular earnings, that is, the 1 percent will pull away from the rest of society.

In this setting we expect that a small number of large firms will capture most of the world market for search engines. Further, these firms will be hugely profitable. We have in mind a situation like that found empirically in the search engine market. The top five leaders are (billions of monthly visitors in parentheses): Google (1.8), Bing (0.5), Yahoo (0.5), Baidu (0.5), and Ask (0.3).[12] If the Chinese government subsidizes Baidu or excludes Google from China, then Baidu captures a larger share of the market. This generates higher profits and higher earnings for shareholders within China, making China better off both absolutely and relatively to the United States. Depending on the details of the model, the United States may or may not be absolutely worse off.

This example is very similar to the mid-1980s discussions about commercial jet production. At a time when it was understood that there was room for only two players in the industry (Boeing and McDonnell Douglas were the leaders), the European Union (EU) heavily subsidized Airbus and ultimately

9. To our knowledge there are no superstar-and-trade models beyond Manasse and Turrini (2001), which deals with trade and wage inequality.

10. This is an ad hoc assumption, but to the extent that it has the flavor of scale economies, we will see less ad hoc variants in the models reviewed below.

11. From a previous footnote, $\partial \pi(a(q), q)/\partial q = p_q r(a)$. Hence $\partial^2 \pi(a(q), q)/\partial q^2 = p_{qq} r + p_q r_a a_q > 0$.

12. Source: http://www.ebizmba.com/articles/search-engines, July, 2017.

forced McDonnell Douglas to exit. These EU subsidies were enormous, but may nevertheless have been valuable for EU taxpayers.[13]

Our superstars model provides a more compelling case for government intervention because scale in data acts as a natural barrier to entry that prevents the free-entry condition from offsetting the impacts of government policies. Thus, the government can beneficially subsidize the education of AI scientists and/or subsidize the entry of firms, for example, by offering tax breaks, subsidies, expertise, incubators, and so forth. This establishes that scale economies and the supernormal profits they sometimes imply strengthen the case for strategic trade policy.

There is, however, one more assumption we have made that is essential to the argument for strategic trade policy, namely, that there are no international knowledge spillovers. In the extreme, if all the knowledge created, for example, by Canadian scientists, moved freely to the United States or China, then a Canadian subsidy would help the world, but would not differentially help Canada. This establishes the critical role of knowledge diffusion (in addition to scale) for thinking about government policies that promote AI.

Empirics. What do we know about superstar effects empirically? Nothing from the trade literature. We know that superstars matter for the rate and direction of innovation in academic research. We know that universities have played a key role in developing AI expertise and that a small number of university-affiliated chief scientists have played a key role in developing new technologies. We also have some evidence of a knowledge externality. Azoulay, Graff Zivin, and Wang (2010) show that the death of a superstar scientist in a field slows progress in the research area of the superstar. The field suffers as scientists associated with the deceased superstar produce less research. While Azoulay, Graff Zivin, and Wang do not consider AI, their work points to the existence of knowledge spillovers that are local rather than global.

Inequality. This discussion has not had much to say about inequality. In our superstars model, industrial policy and strategic trade policies are successful precisely because they promote large and highly profitable firms. We know that these firms account for an increasing share of total economic activity and that they are likely major contributors both to falling labor shares (Autor et al. 2017) and to rising top-end inequality. Thus, the policies being supported by our model do not lead to broad-based prosperity. This cannot be ignored.

Extensions. While the above model of AI science superstars is useful, it

13. The subsidies have continued unabated for over four decades. In 2016, the World Trade Organization (WTO) found that WTO-noncompliant EU subsidies were $10 billion. This does not include the WTO-compliant subsidies. Likewise, the WTO found comparable numbers for WTO-noncompliant US subsidies of Boeing. See Busch (2001) for a history. This raises the possibility that subsidies that are intended to get a firm "on its feet" become permanent, which is yet another reason to be skeptical about strategic trade policies.

has a number of other problems. It is beyond the scope of this chapter to resolve these problems through additional modeling. Instead, we highlight each problem and review the related international trade and growth literatures in order to provide insights into how the model might be improved and what the implications of these improvements are for thinking about trade and trade policy. The problems we cover are the following.

1. The scale assumption $p_{qq} > 0$ is ad hoc. In subsection B below, we consider scale returns that are external to the firm and show that the form of the scale returns matters for policy.
2. In our model, there is no knowledge creation within firms and no knowledge diffusion across firms and borders. In subsection C below, we review endogenous growth models and show that the form of knowledge diffusion, whether it is local or global, matters for policy.
3. Our model ignores the geography of the industry and so does not speak to economic geography and "supercluster" policies. We review the economic geography literature in subsection D below.
4. In section E below we discuss the implications for supercluster policies.

19.3.2 Increasing Returns to Scale External to the Firm—A Basic Trade Model

We start with a simple trade model featuring economies of scale whose geographic scope is variable, that is, regional, national, or international. The model captures the core insights of richer models developed by Ethier (1982), Markusen (1981), and Helpman (1984), along with more recent developments by Grossman and Rossi-Hansberg (2010, 2012).

Firm i produces a homogeneous good using a production function

$$q_i = Q^{\alpha} F(L_i, K_i),$$

where L_i is employment of labor, K_i is employment of capital, F displays constant returns to scale, Q is industry output ($Q = \Sigma_i q_i$), and $0 < \alpha < 1$; Q^{α} is like a Solow residual in that it controls productivity. The idea is that a firm's productivity depends on the output of all firms.[14] If Q is *world* output of the industry, then productivity Q^{α} is common to all firms internationally and scale has no implications for comparative advantage. On the other hand if Q is *national* output of the industry, then the country with the larger output Q will have higher productivity Q^{α} and hence will capture the entire world market.

Artificial intelligence as an industry has a technology that lies somewhere between national returns to scale (Q is national output) and international returns to scale (Q is international output). With national returns

14. Each firm ignores the impact of its output decision on Q so that returns to scale can be treated as external to the firm.

to scale, a government policy such as tariffs or production subsidies that increases domestic output will increase national welfare because the policy raises average productivity at home and also drive exports. Whether it helps or hurts the foreign country depends on a number of factors such as the strength of the scale returns (the size of a) and the size of the countries (Helpman 1984). Most important, the domestic benefits of industrial and trade policies depend on the geographic extent of scale, that is, how much of it is national versus international.

Whether scale operates at the national or international level is not easy to assess and has not been attempted for AI. For the DRAM market in the 1980s, Irwin and Klenow (1994) show that external economies of scale were entirely international rather than national. Other evidence that AI economies are international is the fact that AI algorithms have been disseminated internationally via scientific journals and teaching, and research and development (R&D)-based AI knowledge has diffused internationally via imitation and reverse engineering. On the other hand, the colocation of AI researchers in Silicon Valley and a handful of other technology hubs is suggestive of national and even subnational returns to scale. Azoulay, Graff Zivin, and Wang (2010) also suggests the existence of subnational returns to scale. Clearly, more research is needed on the extent of national versus international returns to scale in AI.

19.3.3 Knowledge Creation and Diffusion: Endogenous Growth

In the previous section, scale was external to the firm and, relatedly, firms did no research. We now introduce firm-level research. Conveniently, some of the key implications of firm-level innovation are similar to those from the previous section, namely, that trade policy depends in large part on the extent to which knowledge spillovers are national or international. To see this, we review the main endogenous growth models that feature international trade. These are Grossman and Helpman (1989, 1990, 1991), Rivera-Batiz and Romer (1991), and Aghion and Howitt (2009, ch. 15). In these models, firms conduct costly R&D and there is an externality that affects these costs. The dominant model in the trade literature features quality ladders (Grossman and Helpman 1991) featuring vertical (quality) differentiation. The highest-quality firm takes the entire market and earns profits.[15]

Innovation improves the quality of the frontier firm by a constant proportion λ. At date $t > 0$, let $n(t)$ be the number of quality improvements during the time interval $(0,t)$ so that the frontier quality is $\lambda^{n(t)}$. Firms invest an amount r in R&D and this generates an endogenous probability $p(r)$ of becoming the quality leader (with quality $\lambda^{n(t)+1}$).

A key feature of the R&D process is an externality: innovators stand

15. Ex post profits are needed in order to justify R&D expenses. However, these models have a free-entry condition that drives ex ante profits to zero.

on the shoulders of giants in the sense that they improve on the frontier level of quality. Had they improved on their own quality, there would be no externality. A two-sector, two-country quality ladder model appears in Grossman and Helpman (1991). Grossman and Helpman assume that there is a standard constant-returns-to-scale sector and a quality sector.[16]

Another popular approach is Romer's (1990) expanding-varieties model. Final goods producers combine varieties of intermediates using a constant elasticity of substitution (CES) production function so that there is love of variety. At any date t there is a measure $N(t)$ of varieties. The marginal returns to new varieties are positive, but diminishing. The key "building on the shoulders of giants" externality is that the cost of developing a new variety is inversely proportional to the measure of varieties. As a result, innovation costs fall over time, generating endogenous growth. A one-sector, two-country extension appears in Rivera-Batiz and Romer (1991). A two-sector, two-country extension appears in Grossman and Helpman (1991).

This brief review leads to a number of observations. As in the previous section, the benefit of trade policy depends on whether the externality operates at the national or international levels; Q of the previous section is replaced here by either $\lambda^{n(t)}$ or $N(t)$. Hence, if each firm builds on the *international* frontier $\lambda^{n(t)}$ or the *international* number of varieties $N(t)$, then there are no implications for comparative advantage; however, if each firm builds on its national $\lambda^{n(t)}$ or national $N(t)$ then the frontier country will develop an increasingly strong comparative advantage in the quality or expanding-varieties sector. With national-level externalities one country will capture the lion's share of the quality/varieties sector. Further, a country can capture this sector by using R&D and trade policies.

Endogenous growth models provide important insights into the details of R&D and trade policies. Research and development policies directly target the knowledge externality and so are preferred to (second-best) trade policies. One R&D policy avenue is to promote *knowledge diffusion*. This can be done through subsidies to nonprofit organizations targeting local within-industry interactions and industry-university collaborations. A second R&D policy avenue is to promote *knowledge creation* through R&D subsidies that are available to all firms, universities, and students. There is a tension between these two avenues; knowledge diffusion can discourage knowledge creation since knowledge diffusion to competitors reduces the returns to innovation. However, the tension is sometimes constructive: Silicon Valley emerged from the shadows of Massachusetts' Route 128 partly because of an "open-source attitude" (Saxenian 1994) and Califor-

16. Placing endogenous growth into a two-sector model so as to facilitate a discussion of comparative advantage is not easy because the sector with improving quality slowly takes over the entire economy unless other price or nonprice "congestion" forces prevent this.

nian restrictions on noncompete clauses (Marx and Fleming 2012). It is less likely that diffusion of knowledge to foreign countries will be as beneficial domestically.

This class of models discourages policies that target individual firms or that "pick winners." To understand why industry leaders should *not* be advantaged by policy, note that counterintuitively, industry leaders will be the least innovative firms due to the "market-stealing" effect. If an entrant innovates, it steals the market from the leader. If a leader innovates, it cannibalizes itself. Leaders therefore have *less* of an incentive to innovate. Aghion et al. (2001, 2005) address this counterintuitive result by developing a model in which leaders innovate in order to escape the competition. Aghion et al. (2017) and Lim, Trefler, and Yu (2017) are currently developing international trade models featuring escape the competition.

In the context of AI, none of the above endogenous growth models is ideal, leading us to conjecture about what an appropriate model might look like. The advantage of endogenous growth models is that they emphasize knowledge creation and diffusion. Thinking more deeply about AI development and commercialization, it is useful to distinguish two aspects of what is done in the AI research departments of large firms. First, they improve AI algorithms, which have the flavor of quality ladders. (Recall that quality can be something that is perceived by consumers *or*, as is relevant here, something that reduces marginal costs.) Second, AI research departments develop new applications of existing AI; for example, Google uses AI for its search engine, autonomous vehicles, YouTube recommendations, advertising network, energy use in data centers, and so forth. This suggests an expanding-varieties model, but one that operates *within* the firm. We are unaware of any endogenous growth models that have both these features. Grossman and Helpman (1991) have the first and Klette and Kortum (2004) have the second. Combining them in one model is not trivial and analytic results would likely have to be replaced with calibration.

19.3.4 New Economic Geography and Agglomeration

The discussion in the previous section points to the possibility that knowledge spillovers are subnational, and this leads naturally to a theory of regional clusters such as Silicon Valley. New economic geography or NEG (Krugman 1980) does not typically consider knowledge spillovers, but it does consider other local externalities that drive regional clusters. Three mechanisms have been particularly prominent: (a) demand-side "home-market effects," (b) upstream-downstream linkages, and (c) labor-market pooling. All of these theories feature two key elements: costs of trading across regions (e.g., tariffs) and increasing returns to scale at the firm level (which can be thought of as the fixed costs of developing a new product). We explain the role of these two elements in the context of home-market effects.

Consider a model with CES monopolistic competition and two regions ($j = 1, 2$). There are varieties of machines and the larger the set of machines to choose from, the more productive are the producers. Let N_j be the measure of machine varieties available in region j. Then with CES production functions, productivity is proportional to N_j.[17] The fundamental factor pushing for agglomeration is the strength of this love-of-variety/productivity externality. (This is related to the externality in Romer's expanding varieties model, which is also proportional to N_j.) As in previous models, the externality operates at the local level rather than at the international level. This externality encourages firms to colocate or agglomerate since the agglomeration of firms drives up N_j and productivity. The fundamental factor pushing against this agglomeration is trade costs: a firm can avoid trade costs by locating close to consumers rather than close to other producers. The main insight of this model is that in equilibrium a disproportionate share of the world's firms will locate in a single region, and this region will thus have higher productivity. As a result, this region will be richer. Notice that firms are choosing to set up where the competition is greatest and where wages and property values are the highest.

The above model of agglomeration has been extended in countless ways (e.g., Krugman and Venables 1995; Fajgelbaum, Grossman, and Helpman 2011; Duranton and Puga 2001) and it is easy to think of applications where the force for agglomeration is not the variety of machines, but the variety of knowledge held by firms. If this knowledge is tacit (meaning it cannot be codified and transmitted in a document), then knowledge spillovers are only transmitted locally via face-to-face interactions. In this case, knowledge externalities lead firms to agglomerate. The result is regions like Silicon Valley.

19.3.5 Cluster Policies

Cluster policies have long been the politician's best friend, yet economists remain highly critical of them. In surveying the evidence for the success of these policies, Uyarra and Ramlogan (2012) write "There is no clear and unambiguous evidence that over the long term clusters are able to generate strong and sustainable impacts in terms of innovation, productivity or employment." One of the world leaders in the economics of clusters, Gilles Duranton, titled his 2011 survey "'California Dreamin': The Feeble Case for Cluster Policies." Yet clusters remain fashionable.

In light of what we have described, the first question is: When are cluster policies likely to succeed? The answer is that they are most likely to succeed when there is clear evidence of scale economies and of knowledge creation together with local knowledge diffusion. Artificial intelligence displays these

17. More precisely, productivity is proportional to $N^{1/(\sigma-1)}$ where $\sigma > 1$ is the elasticity of substitution between varieties.

characteristics, though the extent of international knowledge diffusion cannot be ignored.

The second question is: What policies are likely to work? To answer this question we turn to the insights of Ajay Agrawal, Director of Rotman's Creative Destruction Lab (CDL), and Michael Porter, the business guru of cluster policies. We start with Agrawal. Agrawal identifies two problems with developing AI in the Canadian context. First, there is a shortage of people with the skills to scale up companies. Agrawal calls these people 1000Xers. Second, the cost of information about a start-up's quality is so high that capital markets cannot identify the best and the brightest start-ups. Agrawal's CDL addresses both of these problems by linking start-ups with serial entrepreneurs who can identify a good start-up, tap into 1000Xers for growth, and pass on valuable information about start-up quality to investors globally.

Another approach to the question of what policies are likely to work utilizes Porter's (1990) diamond, which emphasizes four features of clusters: (*a*) factor conditions such as universities and an abundant supply of AI scientists, (*b*) home-market-demand externalities for AI, (*c*) externalities flowing from suppliers of specialized intermediate inputs into AI such as financial services, and (*d*) a competitive environment. Items *b*–*d* involve effects that have already been described in our discussion of knowledge spillovers and lie at the heart of local agglomeration. Item *a* is a more conventional economic factor, that is, drive down the price of the key input by subsidizing its supply. Yet Porter's research shows that many clusters are driven primarily by *a*. That is to say, the single most important policy in practice is simple: follow Hinton's advice in training a large number of AI scientists locally.

Our models also suggest two difficulties with Hinton's advice that must be shored up. First, there is international rather than national knowledge diffusion due to the fact that, for example, Canadian-trained scientists are likely to leave Canada for Silicon Valley, China, and other AI hotspots. This suggests value in programs like those used successfully in Singapore that require student loans to be repaid if the student does not work in Singapore for a minimum number of years.

Second, scale in data is a huge problem for a small country like Canada. To understand appropriate solutions for this, we now turn to the details of national regulatory environments that affect data and the use of AI.

19.4 Behind-the-Border Trade Barriers: The Domestic Regulatory Environment

Given these models, we next turn to the specific regulatory issues that are likely to impact trade policy. Many of the core trade issues around AI involve access to data. Data is a key input into AI, and there are a number of government policies that affect data access and data flows. To the extent these regulations vary across countries, they can advantage some countries' AI industries. The models above suggest that this advantage can have consequences if there are economies of scale, local externalities, and/or rents.

We highlight five policies in particular. The first three involve data: domestic privacy policy, data localization rules, and access to government data. The others are development of the regulation of AI application industries (such as autonomous vehicles) and protection of source code. Privacy policy, data localization, and source code access have already become significant trade issues. For example, the TPP addresses all three of these, as do the US Trade Representative's NAFTA renegotiation objectives. The US position is that strong Canadian and Mexican privacy rules, localization requirements, and access to foreign source code are all impediments to US exports of AI-related goods. In other words, the emphasis on trade policy in these areas is that regulation could be disguised protection that helps domestic firms and hurts foreign firms. In the discussion below, we explore the extent to which this starting assumption is appropriate.

Privacy Regulation. Privacy regulation involves policies that restrict the collection and use of data. Such regulation differs across locations. Privacy policy has the power to limit or expand the ability of firms to use AI effectively. Restrictions on the use of data mean restrictions on the ability to use AI given the data available; however, restrictions on the use of data may also increase the supply of data available if it leads consumers to trust firms that collect the data. Although the theory is ambiguous, thus far, the empirical evidence favors the former effect on balance. Stricter privacy regulations reduce the ability of firms and nonprofits to collect and use data and therefore leads to less innovative use of data (Goldfarb and Tucker 2012). Thus, firms in some countries may benefit from favorable privacy policy.

We believe the most useful analogies for privacy policy in trade relate to labor and environmental regulations. Such regulations also differ across countries for a variety of reasons. They could reflect differences in preferences across countries, or could be perceived as normal goods that wealthier countries are willing to pay for but poorer countries are not (Grossman and Krueger 1995). There is room for reasonable disagreement on how data might be collected or used. Some countries will restrict the information used in prediction while others will not. For example, for insurance, the data that can be used varies by state, with different states providing a variety of restrictions on the use of race, religion, gender, and sexual orientation in insurance.[18] Even with such restrictions, if other variables provide surrogates for such categories, it is possible that firms may be forced to abandon AI methods entirely for more transparent prediction technologies. In terms of

18. http://repository.law.umich.edu/cgi/viewcontent.cgi?article=1163&context=law_econ_current.

privacy policy, we think it is useful to take as given that there are differences across countries in their preferences for policies that restrict the collection and use of data.

Given these differences in preferences, what are the implications for trade? Suppose that the optimal privacy policy for growing an AI industry involves relatively few restrictions on data. Artificial intelligence requires data, and so the fewer government restrictions on data collection, the more rapidly the industry grows.[19] To the extent that young firms tend to grow by focusing on the domestic market, this will advantage the growth of AI firms in some countries relative to others. Thus, lax privacy policies may help domestic industry relative to countries with strict policies just as lax labor and environmental regulation may help the domestic industry.

This suggests the potential of a "race to the bottom" in privacy policy. Evidence for such races has been found in enforcement of labor policies (e.g., Davies and Vadlamannati 2013) and in environmental policies (e.g., Beron, Murdoch, and Vijverberg 2003; Fredriksson and Milliment 2002). There is evidence that privacy regulation does disadvantage jurisdictions with respect to their advertising-supported software industries. In particular, Goldfarb and Tucker (2011) examined a change in European privacy regulation (implemented in 2004) that made it more difficult for European internet firms to collect data about their online customers. This regulatory change was particularly likely to reduce the effectiveness of advertising on websites that relied on customer-tracking data. Using a consistent measure of the effectiveness of thousands of online advertising campaigns, the results showed that European online advertising became about 65 percent less effective after the regulation took effect, compared to before the regulation and compared to advertising in other jurisdictions, mainly the United States. In other words, privacy regulation seemed to reduce the ability of companies to use data effectively. In a different context, Miller and Tucker (2011) show that state-level privacy restrictions can reduce the quality of health care. While this evidence does not pertain to AI, just like AI, online advertising and health care use data as a key input. In other words, the same forces will likely be at play for privacy regulation that restricts the ability of AI to operate.

Under strategic trade models, such races to the bottom are likely to matter if there are rents to be gained from AI. Under endogenous growth models with local spillovers and various agglomeration models, this could create an equilibrium in which the AI industry moves to the country with the most lax policies. Currently, privacy policies are much stricter in Europe than in the

19. Importantly, this is not a statement about the optimal privacy policy from the point of view of a firm. If consumers have a preference for privacy, the private sector can provide it even in the absence of regulation. For a richer debate on this point, see Goldfarb and Tucker (2012) and Acquisti, Taylor, and Wagman (2016).

United States or China.[20] Furthermore, there are a number of differences in such policies between the United States and China. This may give the United States and China an advantage over Europe in this industry.

If stricter privacy policy is likely to hamstring domestic firms in favor of foreign ones, we would expect policy to emphasize avoiding such a race to the bottom; however, recent trade negotiations have instead focused on privacy regulation as disguised protection. For example, this argument is at odds with the current US trade negotiation objectives, which want to weaken Canadian privacy laws. Based on the existing evidence from other data-driven industries, we believe this will help the Canadian industry relative to the US industry in the long run, even if it benefits American companies that already do business in Canada in the short run. In addition, TPP's chapter 14 on Electronic Commerce contains provisions that attempt to limit disguised protection, but contains almost no language that encourages harmonization in privacy policies beyond a request in Article 14.8.5 to "endeavor to exchange information on any such [personal information protection] mechanisms . . . and explore ways to extend these or other suitable arrangements to promote compatibility between them." The words "endeavor" and "explore" are what are known in the trade policy literature as "aspirational" language and generally have no force. The CETA agreement is even more vague with respect to electronic commerce generally. The electronic commerce section, chapter 16, says little but "recognize the importance of" electronic commerce regulation and interoperability and that "the Parties agree to maintain a dialogue on issues raised by electronic commerce."[21]

It is important to note that this is not a statement about company strategy. The market may discipline and provide consumer protection with respect to privacy. Apple, in particular, has emphasized the protection of the personal information of its customers as it has rolled out AI initiatives, and it is an open question whether this strategy will pay off in terms of consumer loyalty and access to better quality, if limited, data.

We also want to emphasize that we do not have a position on the optimal amount of privacy as enforced by regulation. In fact, we think this is a difficult question for economists to answer. Given that the empirical evidence suggests that privacy regulation, on balance and as implemented thus far, seems to reduce innovation, the determination of the optimal amount of privacy should not focus on maximizing innovation (through, as the TPP

20. Canada sits somewhere in the middle. Europe is strict on both data collection and its uses. Canada's core restrictions involve use for a purpose different from the collection context. The United States emphasizes contracts, and so as long as the privacy policy is clear, companies can collect and use data as they wish (at least outside of certain regulated industries like health and finance).

21. https://ustr.gov/sites/default/files/TPP-Final-Text-Electronic-Commerce.pdf, http://www.international.gc.ca/trade-commerce/trade-agreements-accords-commerciaux/agr-acc/ceta-aecg/text-texte/16.aspx?lang=eng.

emphasizes in article 14.8.1, "the contribution that this [privacy protection] makes to enhancing consumer confidence in electronic commerce"). Instead, it is a balance of the ethical value of (or even right to) privacy and the innovativeness and growth of the domestic AI industry.

To reiterate, privacy regulation is different from many other regulations because privacy (perhaps disproportionately) hamstrings domestic firms. Therefore, trade negotiations should not start with the assumption that privacy regulation is disguised protection. Instead, discussions should start with the public policy goal of the "social benefits of protecting the personal information of users of electronic commerce" that is also mentioned in article 14.8.1 of the TPP. Then, if needed, discussions can move to any particular situation in which a privacy regulation might really be disguised protection. As we hope is clear from the above discussion, domestic privacy regulations that restrict how firms can collect and use data are unlikely to be disguised protection. We next turn to two other regulations that might use privacy as an excuse to favor, rather than hamstring, domestic firms.

Data Localization. Data localization rules involve restrictions on the ability of firms to transmit data on domestic users to a foreign country. Such restrictions are often justified by privacy motivations. Countries may want data to stay domestic for privacy and (related) national security reasons. In particular, the argument for data localization emphasizes that governments want the data of their citizens to be protected by the laws of the domestic country. Foreign national security agencies should not have access to data that occurs within a country, and foreign companies should be bound by the laws of the country where the data were collected. The argument against such localization (at least in public) is technical: such localization imposes a significant cost on foreign companies wanting to do business. They need to establish a presence in every country, and they need to determine a system that ensures that the data is not routed internationally (something that is technically costly, particularly for integrated communications networks such as within Europe or within North America). US-based companies have lobbied against such requirements.[22]

On the technical side, consider two parties, A and B, who reside in the same country. Internet traffic between A and B cannot be confined within national borders without specific technical guidance (and some cost to quality) because the internet may route data indirectly. In addition, data on a transaction between A and B may be stored on a server located in a different country. Furthermore, if A and B reside in different countries, then the data on that transaction will likely be stored in both countries.[23]

Data localization is an issue for AI because AI requires data. And it often involves merging different data sources together. The quality of aggregate

22. https://publicpolicy.googleblog.com/2015/02/the-impacts-of-data-localization-on.html.
23. Dobson, Tory, and Trefler (2017).

predictions from AI will be lower if the scale of data is limited to within a country. In other words, localization is a way to restrict the possible scale of any country in AI, but at the cost of lower quality overall.

Put differently, data localization is a privacy policy that could favor domestic firms. Unlike the consumer protection privacy policies highlighted above, it can favor domestic over foreign firms because the foreign-firm AI experts may not have access to the data. The TPP recognizes this and explicitly restricts it in Article 14.11.3a, which states that the cross-border transfer of information should not be restricted in a manner that would constitute "a disguised restriction on trade."[24]

Privileged Access to Government Data. Another potential restriction on trade that might be justified by privacy concerns involves access to government data. Governments collect a great deal of data. Such data might be valuable to training AIs and improving their predictions. Such data include tax and banking data, education data, and health data. For example, as the only legal provider of most health care services in Ontario, the Ontario government has unusually rich data on the health needs, decisions, and outcomes of 14 million people. If domestic firms are given privileged access to that data, it would create an indirect subsidy to the domestic AI industry.

We think the most useful analogy in the current trade literature is the perennial softwood lumber trade dispute between Canada and the United States. In the softwood lumber case, most timber in Canada is on government-owned land, while in the United States, most timber is on privately owned land. The US complaints allege that Canadian timber is priced too low, and is therefore a government subsidy to the Canadian lumber industry. While there have been various agreements over the years, the disagreement has not been fully resolved. The superficial issue is what a fair price should be for access to government resources. The real issue is whether legitimate regulatory differences can be argued to convey unfair advantage and therefore constitute a trade-illegal subsidy.

Government data can be seen similarly. Links between the state and the corporation vary by country, and this might help some corporations more than others. What is a fair price for access to the data? Importantly, governments may not want to give foreign firms access to such data for the same privacy and national security issues that underlie motivations for data local-

24. Related to the issue of data localization is the question of who owns data collected on domestic individuals by foreign individuals or firms. For example, consider an American company that uses Peruvians' cell phones to gather data on agriculture and climate. Who owns the rights to that data? Are the Americans allowed to profit from that data? Are contracts between the individual actors enough, or is there a need for international laws or norms? The data might not be collected if not for the private companies, but the companies use the data in their own interest rather than in the public interest or in the interest of the Peruvians who provided the data. The recent attempts at a joint venture between Monsanto and John Deere, along with the US Department of Justice antitrust concerns that scuttled the deal, highlight how tangible this issue is.

ization. Thus, seemingly reasonable differences across countries in their data access policies can end up favoring the domestic industry.

Industrial Regulation. Most international agreements have a section on competition policy and industrial regulation. This is because regulation can be a source of unfair comparative advantage or disadvantage. In AI applications, this list is long. In addition to the points around data and privacy highlighted above, many applications of AI involve complementary technologies in which standards might not yet exist and the legal framework might still be evolving.

For example, in autonomous vehicles, a variety of standards will need to be developed around vehicle-to-vehicle communication, traffic signals, and many other aspects of automotive design. Most of these standards will be negotiated by industry players (Simcoe 2012), perhaps with some government input. As in other contexts, national champions can try to get their governments to adopt standards that raise costs for foreign competition. This leads to the possibility of international standards wars. This is particularly true of standards that are likely to involve a great deal of government input. For example, suppose governments require that the AI behind autonomous vehicles be sufficiently transparent that investigators are able to determine what caused a crash. Without international standards, different countries could require information from different sensors, or they could require access to different aspects of the models and data that underlie the technology. For companies, ensuring that their AI is compatible with multiple regulatory regimes in this manner would be expensive. Such domestic regulations could be a way to favor domestic firms. In other words, domestic technology standards around how AI interacts with the legal regime is a potential tool for disguised restriction on trade.

The autonomous vehicle legal framework is evolving, with different countries (and even states within the United States) allowing different degrees of autonomy on their public roads. Drones are another example where, in the United States, the Federal Aviation Administration (FAA) strictly regulates American airspace, while China and some other countries have fewer restrictions. This may have allowed China's commercial drone industry to be more advanced than the industry in the United States.[25] Thus, regulation can also impact the rate of innovation and therefore comparative advantage.

Source Code. To the extent that AI may discriminate, governments may demand information about the algorithms that underlie the AI's predictions under antidiscrimination laws. More generally with respect to software, including AI, governments may demand access to source code for security reasons, for example, to reduce fraud or to protect national security. Thus, using consumer protection or national security as an excuse, governments

25. https://www.forbes.com/sites/sarahsu/2017/04/13/in-china-drone-delivery-promises-to-boost-consumption-especially-in-rural-areas/#47774daf68fe.

could reduce the ability of foreign firms to maintain trade secrets. Furthermore, cyber espionage of such trade secrets may be widespread, but that is beyond the scope of this chapter.[26] Broadly, this issue has been recognized in the TPP negotiations, with Article 14.17 emphasizing that access to source code cannot be required unless that source code underlies critical infrastructure or unless the source code is needed to obey other domestic regulations that are not disguised restrictions on trade.

Other policies that might affect the size of domestic AI industries include intellectual property, antitrust, R&D subsidies, and national security. If AI is the next important strategic industry, then all of the standard questions arise with respect to trade policies in these industries. We do not discuss these in detail because we think the trade-specific issues with respect to these policies are not distinct to AI, but are captured more generally by the discussion of innovation and trade. The main point for these other aspects of domestic policy with respect to AI and trade is that there are economies of scale in AI at the firm level. Furthermore, we expect some of the externalities from the AI industry to remain local.

19.5 AI and International Macroeconomics

Before concluding, it is important to recognize that AI will have implications for international macroeconomics. For example, suppose that China does succeed in building a large AI industry. This will likely increase its trade surplus with the rest of the world, particularly in services. Furthermore, suppose that China manages to control wage inflation through promoting migration from rural to urban areas, and by relaxing the one-child policy. Then, this is likely to put upward pressure on the renminbi (RMB) and downward pressure on the dollar.

This will have implications for US labor markets. At the low end of the market, a weakening dollar might repatriate manufacturing jobs. At the high end of the market, skilled US workers will for the first time be exposed to competition from a low-wage country. In isolation, this would reduce one dimension of domestic US inequality.

If the Chinese market becomes open to US technology giants (and vice versa), both the Melitz (2003) model and the Oberfield (2018) model of trade predict that the giants will grow even larger. In the context in which these companies have already absorbed one-fifth of US value added, and may have contributed to US top-end inequality, the impact of international trade in further growing these impacts may increase top-end inequality.

26. https://obamawhitehouse.archives.gov/sites/default/files/omb/IPEC/admin_strategy_on_mitigating_the_theft_of_u.s._trade_secrets.pdf.

19.6 Conclusion

How will artificial intelligence affect the pattern of trade? How does it make us think differently about trade policy? In this article we have tried to highlight some key points.

First, the nature of the technology suggests that economies of scale and scope will be important. Furthermore, as a knowledge-intensive industry, knowledge externalities are likely to be important. Prior literature on other industries suggests that such externalities are often local, but more evidence is needed. Second, the trade models that are likely to be most useful in understanding the impact of AI are those that account for these points, specifically, scale, knowledge creation, and the geography of knowledge diffusion. These models suggest that whether AI-focused trade policies (or AI-focused investments in clusters) are optimal will depend very much on the presence of scale and the absence of rapid international knowledge diffusion. Third, we discussed whether and how regulation might be used to favor domestic industry. We highlighted that privacy policy that targets consumer protection is unlike many other regulations in that it is likely to hamstring domestic firms, even relative to foreign ones. So, rather than focusing trade discussions on how privacy policy might be used as a disguised restriction on trade, such discussions should emphasize regulatory harmonization so as to avoid a race to the bottom. In contrast, several other policies may be used to favor domestic firms including data localization rules, limited access to government data, industry regulations such as those around the use of drones, and forced access to source code.

Generally, this is an exciting new area for trade research and policy. There is still much to learn before we have a comprehensive understanding of these questions.

References

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman. 2016. "The Economics of Privacy." *Journal of Economic Literature* 54 (2): 442–92.

Aghion, Philippe, Antonin Bergeaud, Matthieu Lequien, and Marc Melitz. 2017. "The Impact of Exports on Innovation: Theory and Evidence." Working paper, Harvard University.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. 2005. "Competition and Innovation: An Inverted-U Relationship." *Quarterly Journal of Economics* 120 (2): 701–28.

Aghion, Philippe, Christopher Harris, Peter Howitt, and John Vickers. 2001. "Competition, Imitation and Growth with Step-by-Step Innovation." *Review of Economic Studies* 68 (3): 467–92.

Aghion, Philippe, and Peter Howitt. 2009. *The Economics of Growth*. Cambridge, MA: MIT Press.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Boston, MA: Harvard Business Review Press.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2017. "Concentrating on the Fall of the Labor Share." NBER Working Paper no. 23108, Cambridge, MA.

Azoulay, Pierre, Joshua S. Graff Zivin, and Jialan Wang. 2010. "Superstar Extinction." *Quarterly Journal of Economics* 125 (2): 549–89.

Beron, Kurt J., James C. Murdoch, and Wim P. M. Vijverberg. 2003. "Why Cooperate? Public Goods, Economic Power, and the Montreal Protocol." *Review of Economics and Statistics* 85 (2): 286–97.

Brander, James A., and Barbara J. Spencer. 1981. "Tariffs and the Extraction of Foreign Monopoly Rents under Potential Entry." *Canadian Journal of Economics* 14 (3): 371–89.

Brandt, Loren, and Xiaodong Zhu. 2000. "Redistribution in a Decentralized Economy: Growth and Inflation in Reform China." *Journal of Political Economy* 108 (2): 422–39.

Busch, Marc L. 2001. *Trade Warriors: States, Firms, and Strategic-Trade Policy in High-Technology Competition*. Cambridge: Cambridge University Press.

Davies, Ronald B., and Krishna Chaitanya Vadlamannati. 2013. "A Race to the Bottom in Labor Standards? An Empirical Investigation." *Journal of Development Economics* 103:1–14.

Dobson, Wendy, Julia Tory, and Daniel Trefler. 2017. "Modernizing NAFTA: A Canadian Perspective." In *A Positive NAFTA Renegotiation*, edited by Fred Bergsten, 36–49. Washington, DC: Petersen Institute for International Economics.

Duranton, Gilles. 2011. "California Dreamin': The Feeble Case for Cluster Policies." *Review of Economic Analysis* 3 (1): 3–45.

Duranton, Gilles, and Diego Puga. 2001. "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products." *American Economic Review* 91 (5): 1454–77.

Eaton, Jonathan, and Gene M. Grossman. 1986. "Optimal Trade and Industrial Policy under Oligopoly." *Quarterly Journal of Economics* 101 (2): 383–406.

Ethier, Wilfred J. 1982. "National and International Returns to Scale in the Modern Theory of International Trade." *American Economic Review* 72 (3): 389–405.

Fajgelbaum, Pablo, Gene M. Grossman, and Elhanan Helpman. 2011. "Income Distribution, Product Quality, and International Trade." *Journal of Political Economy* 119 (4): 721–65.

Fredriksson, Per G., and Daniel L. Millimet. 2002. "Strategic Interaction and the Determination of Environmental Policy across U.S. States." *Journal of Urban Economics* 51 (1): 101–22.

Fujii, Hidemichi, and Shunsuke Managi. 2017. "Trends and Priority Shifts in Artificial Intelligence Technology Invention: A Global Patent Analysis." RIETI Discussion Paper Series no. 17-E066, Research Institute of Economy, Trade, and Industry, May.

Goldfarb, Avi, and Catherine Tucker. 2011. "Privacy Regulation and Online Advertising." *Management Science* 57 (1): 57–71.

———. 2012. "Privacy and Innovation." In *Innovation Policy and the Economy*, vol. 12, edited by Josh Lerner and Scott Stern, 65–89. Chicago: University of Chicago Press.

Grossman, Gene M., and Elhanan Helpman. 1989. "Product Development and International Trade." *Journal of Political Economy* 97 (6): 1261–83.

———. 1990. "Trade, Innovation, and Growth." *American Economic Review Papers and Proceedings* 80 (2): 86–91.

———. 1991. *Innovation and Growth in the Global Economy*. Cambridge, MA: MIT Press.

Grossman, Gene M., and Alan B. Krueger. 1995. "Economic Growth and the Environment." *Quarterly Journal of Economics* 110 (2): 353–77.

Grossman, Gene M., and Esteban Rossi-Hansberg. 2010. "External Economies and International Trade Redux." *Quarterly Journal of Economics* 125 (2): 829–58.

———. 2012. "Task Trade between Similar Countries." *Econometrica* 80 (2): 593–629.

Helpman, Elhanan. 1984. "Increasing Returns, Imperfect Markets, and Trade Theory." In *Handbook of International Economics*, edited by Peter B. Kenen and Ronald W. Jones, 325–65. Amsterdam: North-Holland.

Irwin, Douglas A., and Peter J. Klenow. 1994. "Learning-by-Doing Spillovers in the Semiconductor Industry." *Journal of Political Economy* 102 (6): 1200–27.

Klette, Tor Jakob, and Samuel Kortum. 2004. "Innovating Firms and Aggregate Innovation." *Journal of Political Economy* 112 (5): 986–1018.

Krugman, Paul R. 1980. "Scale Economies, Product Differentiation, and the Pattern of Trade." *American Economic Review* 70 (5): 950–59.

———. 1986. *Strategic Trade Policy and the New International Economics*. Cambridge, MA: MIT Press.

Krugman, Paul R., and Anthony J. Venables. 1995. "Globalization and the Inequality of Nations." *Quarterly Journal of Economics* 110 (4): 857–80.

Lim, Kevin, Daniel Trefler, and Miaojie Yu. 2017. "Trade and Innovation: The Role of Scale and Competition Effects." Working paper, University of Toronto.

Manasse, Paolo, and Alessandro Turrini. 2001. "Trade, Wages, and Superstars." *Journal of International Economics* 54 (1): 97–117.

Markusen, James R. 1981. "Trade and the Gains from Trade with Imperfect Competition." *Journal of International Economics* 11 (4): 531–51.

Marx, M., and L. Fleming. 2012. "Non-compete Agreements: Barriers to Entry . . . and Exit?" In *Innovation Policy and the Economy*, vol. 12, edited by J. Lerner and S. Stern. Chicago: University of Chicago Press.

Mayer, Wolfgang. 1974. "Short-Run and Long-Run Equilibrium for a Small Open Economy." *Journal of Political Economy* 82 (5): 955–67.

McLaren, John. 2000. "'Globalization' and Vertical Structure." *American Economic Review* 90 (5): 1239–54.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6): 1695–725.

Miller, A. R., and C. Tucker. 2011. "Can Healthcare IT Save Babies?" *Journal of Political Economy* 119 (2): 289–332.

Mussa, Michael L. 1974. "Tariffs and the Distribution of Income: The Importance of Factor Specificity, Substitutability, and Intensity in the Short and Long Run." *Journal of Political Economy* 82 (6): 1191–203.

Oberfield, Ezra. 2018. "A Theory of Input-Output Architecture." *Econometrica* 86:559–89.

Porter, Michael E. 1990. "The Competitive Advantage of Nations." *Harvard Business Review* 68 (2): 73–93.

PWC. 2017. "Global Top 100 Companies by Market Capitalisation." March 31, 2017, update. https://www.pwc.com/top100. Accessed August 17, 2017.

Rauch, James E. 1999. "Networks versus Markets in International Trade." *Journal of International Economics* 48 (1): 7–35.
Rivera-Batiz, Luis A., and Paul M. Romer. 1991. "Economic Integration and Endogenous Growth." *Quarterly Journal of Economics* 106 (2): 531–55.
Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98 (5): S71–102.
Rosen, Sherwin. 1981. "The Economics of Superstars." *American Economic Review* 71 (5): 845–58.
Ruffin, Roy J., and Ronald W. Jones. 1977. "Protection and Real Wages: The Neoclassical Ambiguity." *Journal of Economic Theory* 14 (2): 337–48.
Saxenian, AnnaLee. 1994. *Regional Advantage Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.
Simcoe, Timothy. 2012. "Standard Setting Committees: Consensus Governance for Shared Technology Platforms." *American Economic Review* 102 (1): 305–36.
Sutton, John, and Daniel Trefler. 2016. "Capabilities, Wealth and Trade." *Journal of Political Economy* 124 (3): 826–78.
Uyarra, Elvira, and Ronnie Ramlogan. 2012. "The Effects of Cluster Policy on Innovation Compendium of Evidence on the Effectiveness of Innovation Policy Intervention." Technical Report, Manchester Institute of Innovation Research Manchester Business School, March.