The Impact of AI and Cross-Border Data Regulation on International Trade in Digital Services: A Large Language Model Approach*

Ruiqi Sun† Daniel Trefler ‡

This Version: 28 November 2023

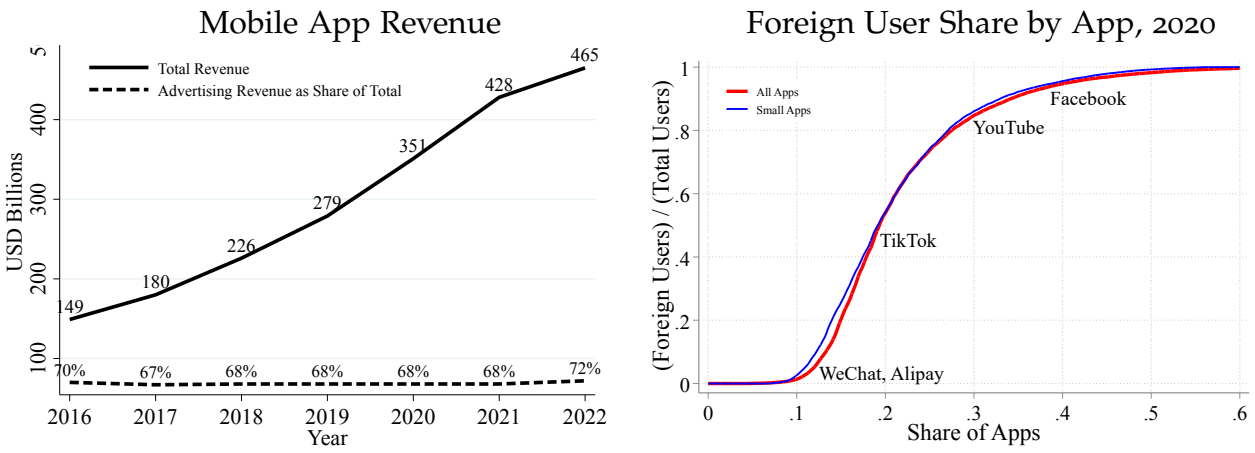
ABSTRACT: The rise of artificial intelligence (AI) and of cross-border restrictions on data flows has created a host of new questions and related policy dilemmas. This paper addresses two questions: How is digital service trade shaped by (1) AI algorithms and (2) by the interplay between AI algorithms and cross-border restrictions on data flows? Answers lie in the palm of your hand: From London to Lagos, mobile app users trigger international transactions when they open AI-powered foreign apps. We have 2015–2020 usage data for the most popular 35,575 mobile apps and, to quantify the AI deployed in each of these apps, we use a large language model (LLM) to link each app to each of the app developer's AI patents. (This linkage of specific products to specific patents is a methodological innovation.) Armed with data on app usage by country, with AI deployed in each app, and with an instrument for AI (a Heckscher-Ohlin cost-shifter), we answer our two questions. (1) On average, AI causally raises an app's number of foreign users by 2.67 log points or by more than 10-fold. (2) The impact of AI on foreign users is halved if the foreign users are in a country with strong restrictions on cross-border data flows. These countries are usually autocracies. We also provide a new way of measuring AI knowledge spillovers across firms and find large spillovers. Finally, our work suggests numerous ways in which LLMs such as ChatGPT can be used in other applications.

*The authors thank our discussants (Reka Juhasz, David Weinstein and Zhe Yuan) as well as Philippe Aghion, Ajay Agrawal, Avi Goldfarb, Gordon Hanson, Kevin Lim, Javier López González, Peichun (Will) Wang, and seminar participants at the Hang Zhou AI and Economics Workshop, Hong Kong University, INSEAD, the Korea Institute for International Economic Policy, Liaoning University, National University of Singapore, the Paris PILLARS workshop, Princeton's IES, Rochester's VITM, Seoul National University, the Trade Policy Research Forum, Yongsai University, University of British Columbia, and University of Toronto. Trefler thanks the Social Sciences and Humanities Research Council of Canada (SSHRC) and the Bank of Canada for generous financial support.

†Rotman School of Management, (ruiqi.sun@rotman.utoronto.ca).

‡Rotman School of Management and Department of Economics, University of Toronto, and NBER (dtrefler@rotman.utoronto.ca).

Figure 1: Mobile Apps: Revenues and Global Reach



Notes: The left panel is mobile app revenue by year (solid red line) and ad revenue as a percentage of this revenue (dashed line). Data are from Statista. The right panel uses data on the share of users who are foreign for each of 35,575 apps in our data. The panel plots the cumulative distribution of these shares. The thick line cumulates over all apps and the thin line cumulates over small apps (apps with below-median numbers of users).

Digital services are the fastest growing component of international trade, already accounting for a quarter of world exports and a third of US exports (OECD, 2023). A dynamic component of this trade is mobile app services. Mobile apps did not exist two decades ago, yet they now dominate the lives of many and are valued by consumers at $2.5 trillion (Brynjolfsson, Collis, Liaqat, Kutzman, Garro, Deisenroth, Wernerfelt, and Lee, 2023). Worldwide mobile app revenue has grown spectacularly, tripling in the last six years to hit $465 billion in 2022. See the left panel of figure 1. While these revenues are not all trade flows, the international trade dimensions of mobile app services are on full display each time we open our smartphones. Users of mobile apps around the world network with friends (Facebook, developed in the USA), stream short videos (TikTok developed in China), listen to music (Spotify, developed in Sweden), navigate (Waze, developed in Israel), and play games (PUBG, developed in South Korea). When the user and developer are located in different countries, the developer is exporting a service to a foreign user. The share of a typical app's users who are foreign is shockingly high. We have computed these shares for the most popular 35,575 mobile apps over 2015–2020 and plotted their cumulative distribution in the right panel of figure 1. Consider TikTok. Despite being low in the distribution, just below the 20th percentile, almost half of its users are foreign. Foreign shares are even higher for YouTube and Facebook. High foreign shares are common throughout the size distribution of apps. Small apps (the thin blue line) have the same profile as all apps (the thick red line). These levels of foreign penetration are *vastly* higher than those for US and French manufacturers (Eaton,

Kortum, and Kramarz, 2004, table 1) and for online web-based services (Alaveras and Martens, 2015, table 2).

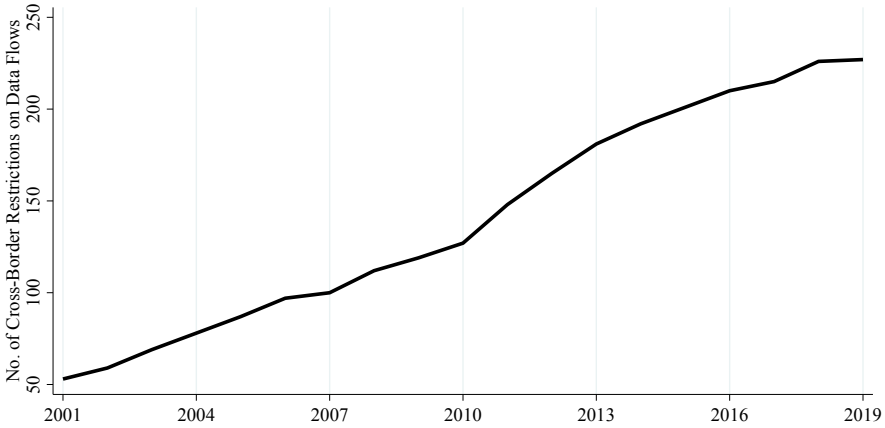
An important feature of many mobile apps is their use of artificial intelligence (AI), both algorithms and data. For example, Facebook tracks its users and feeds the data into algorithms that personalize content, recommend friends, categorize images, translate, chat and more. All of these improve the user experience, which leads to more personal data per user and more users. See Aral (2021) for an in-depth discussion and Sun, Yuan, Li, Zhang, and Xu (forthcoming) for recent evidence on the value of personal data. In this paper, we examine the impact of AI on international trade in mobile app services and how this impact is mitigated by regulations restricting cross-border data flows.

The impact of AI on international trade has received almost no academic scrutiny: Brynjolfsson, Hui, and Liu (2019) study the impact of machine translation on eBay's e-commerce and Beraja, Kao, Yang, and Yuchtman (2023) study the impact of Chinese surveillance policies on China's exports of facial recognition technologies. We study the impact of AI algorithms on a large variety of internationally traded digital services, namely 35,575 mobile apps that cover a wide spectrum of products from games to productivity tools to social networking. The AI algorithms, mostly deep learning, are those described in the 63,679 AI patents owned by app developers. We feed our app and patent data, both of which come with rich text descriptions, into a large language model that identifies which apps use which AI algorithms. We explain how below. Our *first result is that when AI is deployed in a mobile app, the number of foreign users (meaning exports of the app service) increase by orders of magnitude.* Ours is thus the first study of how AI impacts international digital service trade across a wide variety of products. Of course, the decision to deploy AI is endogenous and therefore requires an IV strategy, which we describe below.

The rise of AI in mobile apps has led to a massive migration of data across international borders. Foreign user data is moved to headquarters computing facilities where it is used in increasingly sophisticated AI models. This has led to an explosion of conflicting rules governing cross-border data flows. At one extreme, the US pushes for international agreements promoting freedom of data flows. There are now 72 such international agreements (Nemoto and López González, 2021). At the other extreme, China effectively bans all exports of data and is rapidly exporting its state-based regulatory model worldwide. In between, the EU allows data transfers as long as user rights such as privacy are protected. See Bradford (2023) and O'Shaughnessy (2023). Figure 2 plots the number of restrictions on international data flows by year. The rate of increase was greatest just as mobile apps exploded in popularity.

Despite the rapid rise of conflicting international regulations that promote and restrict

Figure 2: Restrictions on Cross-Border Data Flows



Notes: This figure plots the number of data protection regulations by year. These are regulations relating to international data transfers and local storage requirements (data localization). We are grateful to Casalini and López González (2019, figure 1) for providing the data.

cross-border data flows, these have not been studied empirically by academic international trade economists.[1] We study how regulations restricting cross-border data flows degrade the effectiveness of AI as a tool for improving app quality and hence as a tool for promoting exports of mobile app services. We do this by combining our mobile app data with custom runs of the OECD Digital Services Trade Restrictiveness Index (Ferencz, 2019). *Our second result is that AI's impact on the number of users in foreign country n is halved if country n heavily restricts cross-border data flows.* This explains why Google, Facebook, OpenAI and other firms lobby for digital trade agreements that deregulate cross-border data flows, which in turn explains the pressure on governments to negotiate agreements despite privacy and national security concerns.

IV strategy: Our results are based on regressions of an app's number of foreign users on a measure of the AI deployed in the app. AI deployment is a firm choice variable and so is endogenous. We develop a model of mobile app trade and AI adoption decisions that provides a theory-consistent estimating equation and instrument. The model predicts that AI is adopted (*a*) in countries where AI inputs are relatively inexpensive and (*b*) in industries where AI inputs account for a large share of costs. Examples of mobile app industries are social networking (high AI cost share) and gaming (low AI cost share). This insight leads to a Heckscher-Ohlin-like cost-shifter instrument for AI adoption: AI is deployed in a country-industry pair when (*a*) the country is AI-abundant and (*b*) the industry is AI-intensive.

[1] We are aware of only two non-academic studies, by the USITC (Herman and Oliver, 2022) and the OECD (López González, Sorescu, and Kaynak, 2023).

Summarizing, we show that:

1. AI causally increases international trade in mobile app services by 2.67 log points or by more than 10-fold.
2. This increase is halved by restrictions on cross-border data flows.

Further, these restrictions are most severe in autocracies. There are many other questions we can address e.g., what is the role of gravity. However, due to space limitations, we only examine international trade questions related to AI.

Use of a Large Language Model (LLM): To conduct our analysis we must link apps to AI algorithms. Linking individual patents to individual products is notoriously difficult. It has never been done in the international trade literature or, to our knowledge, in any field of economics.[2] All apps have detailed product descriptions that are used by consumers when choosing apps. Critically, these descriptions are amenable to natural language processing (NLP). Likewise, each patent has text that is amenable to NLP. The novel way we use NLP on app descriptions and patent texts is best explained using a simple scenario. Consider an app developer with a single app a and a single AI patent p. We feed the app description and patent text into Google's LLM, called BERT, and ask it whether the app likely uses the algorithm in the patent. The answer comes back in the form of a cosine similarity ρ_{ap} which is large (small) if the app and patent texts deal with similar (dissimilar) subject matter.[3] The reader who has used ChatGPT, which is based on the same 'transformer' algorithm as BERT, will appreciate the unexpected power of LLMs to extract meaning from text. In the simple scenario above, we measure the AI embodied in app a by $AI_a = \rho_{ap}$. A generalization of this to tens of thousands of apps and patents and 2.4 billion ρ_{ap}, leads to our key regressor measuring the AI deployed in each app.

Externalities: Given that AI algorithms contain nonrival ideas that diffuse across firms, we distinguish between knowledge coming *internally* from the firm's own patents (AI_a) and knowledge coming *externally* from other firms' patents. Building on our simple scenario, suppose there is a second app developer with a single app a' and no patents. Using standard techniques such as patent citations, we would not include the second app developer in the analysis of externalities because it has no patents and hence no citations. In our novel methodology, we are able to measure the external AI embodied in app a'

[2]The linking of patents to industries and product groups has a long lineage e.g., Kortum and Putnam (1997). An excellent recent example is Argente, Baslandze, Hanley, and Moreira (2023), who cluster Nielsen barcode-level products into 400 groups and link these to patents. We note in passing the link of patents to occupations in Webb (2020) and Stapleton and Webb (2023) as well as the link of firms to industries in Hoberg and Phillips (2016) and Pellegrino (2023). None of these papers drill down to the product level and none use LLMs.

[3]App descriptions are consumer-facing while patent texts are engineering- and legal-facing. They thus do not share common words and so cannot be linked using older word frequency techniques such as TFID. An LLM is needed.

using its cosine similarity with the first firm's AI patent. This allows us to expand the set of firms under study to the large mass of firms with no patents. It also allows us to estimate our third result:

3. External AI has economically and statistically significant impacts on international mobile app trade. Externalities are important and affect almost all exporters in the economy.

This provides support for mechanisms emphasized in endogenous growth theory e.g., Grossman and Helpman (1991).

Literature

Little is known about AI's impact on international trade and even less about these impacts specifically on digital service trade. Brynjolfsson *et al.* (2019) show that eBay's introduction of a machine translation system increased its exports by 17.5%. Beraja, Yang, and Yuchtman (forthcoming) show how Chinese government security contracts for facial recognition software provided confidential security data to Chinese firms, data that improved these firms' products. Beraja *et al.* (2023) then show how this increased China's exports of facial recognition technologies. Our paper scales up their excellent research to tens of thousands of products and multiple AI technologies, digs into the adoption decision and cross-firm diffusion mechanisms, and embeds the analysis within an international trade model.

Sun and Trefler (2022) use a subset of the data used here. They do not use an LLM, but instead link patent data to mobile app *industries*. They examine the impact of AI on foreign downloads, the entry and exit of apps, and the welfare gains from mobile app trade. There is thus little overlap with the current paper. Goldfarb and Trefler (2019) and Ferencz, López González, and García (2022) provide qualitative discussions of how AI impacts international trade.[4]

Digital Service Trade Restrictions and Trade: Herman and Oliver (2022) find only weak evidence that trade in services is increased by trade agreements with provisions for free data flows. López González *et al.* (2023, appendix table E.10) investigate the impact of digital trade restrictions on trade using multiple measures of restrictions and trade. When using a broader set of digital restriction than we do, they find large negative effects

[4]There are other related papers that our tangential to our study in that they do not directly measure AI or work with digital service trade. Bailey, Gupta, Hillenbrand, Kuchler, Richmond, and Stroebel (2021) use Facebook data to construct bilateral social connections between countries and show that these are a more powerful determinant of bilateral trade than are traditional determinants such as distance and borders. Unlike our research, they work with trade in goods. As well, there is a vibrant literature on robots and trade. See Stapleton and Webb (2023) and the collection of articles in Yan and Grossman (2022).

of restrictions on trade in goods and services. AI is not part of their study. Goldfarb and Tucker (2019) survey the impact of regulation on the digital economy more generally.[5]

AI, Ads, and the Long Tail of Exporters: Our finding that even small apps have large foreign shares is related to Arkolakis (2010) and Eaton, Kortum, and Kramarz (2011). Advertising plays a large role in this finding and our model draws heavily on Arkolakis (2010).[6]

Externalities and Trade: There is a healthy literature on externalities and trade starting with Coe and Helpman (1997) and most recently by Aghion, Bergeaud, Gigout, Lequien, and Melitz (2021). See Melitz and Redding (2023) for a survey. Firm-level analysis, as in Aghion *et al.* (2021), examines goods trade and patenting. We examine the impact of AI and data restrictions on digital service trade, and offer a novel measure of externalities.

The Internet and Trade: There is a large literature on online services involving a visit to a foreign-hosted website e.g., Freund and Weinhold (2002, 2004), Blum and Goldfarb (2006), Alaveras and Martens (2015), and Lendle, Olarreaga, Schropp, and Vézina (2016).[7] Chen and Wu (2021) and Carballo, Rodriguez Chatruc, Salas Santa, and Volpe Martincus (2022) look specifically at e-commerce websites (platforms) and show that they help producers export more. There is an established literature on mobile apps (e.g., Ershov, forthcoming, Bian *et al.*, 2023), but it does not deal with AI.

This paper is organized as follows. Section 1 provides background on mobile apps and artificial intelligence. Section 2 describes the data on mobile apps and AI patents, and explains the LLM linking procedure. Section 3 lays out the theory underpinning the estimating equation and instrument. Section 4 contains the IV regressions of mobile app service exports on AI (result 1). Section 5 contains the IV regressions involving the interaction of AI with restrictions on cross-border data flows (result 2). Section 6 examines the impact on mobile app service exports of AI knowledge spillovers (result 3).

1. Overview of AI in the Mobile App Industry

How is AI used in mobile apps? This is complicated because there are so many AI algorithms and so many uses. Painting in *very* broad brushstrokes, the app producer wants to maximize profits and typically uses AI to improve revenues rather than reduce costs. Mobile app revenues come from three sources: 10% from streaming services (music

[5]Since that survey, several papers on privacy restrictions have appeared. For examples, see Johnson's (2022) survey of GDPR and the Bian, Ma, and Tang (2023) analysis of changes to Apple's privacy policy.

[6]High foreign shares are also related to the internet literature on long tails (niche products) versus superstars (large products) e.g., Bar-Isaac, Caruana, and Cuñat (2012). Sun *et al.* (forthcoming) ties long tails to AI in an e-commerce setting, but without an international dimension.

[7]Blum and Goldfarb and Lendle *et al.* show that distance matters for online services and matters because of spatially correlated tastes. See Goldfarb and Tucker (2019, section 5) for a broader discussion.

and video), 20% from gaming apps, and 70% from advertising. See the dashed line in figure 1. Revenues thus depend on the number of ads displayed and click through rates per ad, which in turn depend on the number of users, time spent per user, and ad personalization. This paper is about the impact of AI on the number of users.

Some examples illustrate AI's impact on the number of users. Candy Crush does not use AI, but plans on using it to create more frequent content updates. This keeps users engaged. Other potential uses of AI for Candy Crush include calibrating to the user's skill level (avoids user frustration) and personalizing the timing of rewards (keeps the user addicted). In battle royale games, AI is used to personalize interactions with non-player characters. Games typically use less AI than most app categories. The FaceBook social networking app deploys AI for (1) vision and image recognition that categorizes content, (2) recommenders for personalized content such as social interactions with friends and news feeds, (3) content moderation, (4) translation, (4) chatbots for business pages, and much more. All of these improve the user experience, which attracts new users and improves user retention. In short, AI is used in many ways in many apps.

We next describe the mobile app industry. It was born in the second half of 2008 when Apple's App Store was opened with 500 apps for the newly released IPhone 3G. The mobile app market then exploded so that it was well-established by the start of our sample in 2015. There are no systematic data on revenues or profits by app. Even purchase price is usually not available as 94.9% of our apps are free to use.[8] Absent good revenue and profit data, the best available measure of app success is the number of active users. Many apps include code that tracks each time the app is opened and some apps track each time any app is opened. (Yes, you are being tracked.) Data purveyors use this information to calculate the number of people that open the app at least once in a calendar month. This measure of user engagement is called 'monthly active users' or 'users' for short. It is sold commercially to investors and app developers for business analytics, notably as a predictor of revenues. We will use this measure.

An example illustrates the role of user numbers as an indicator of revenues. MIT professor Sinan Aral (Aral, 2021, page 204) describes how the central room in his startup office was dominated by a large screen displaying real-time data on the number of users and time spent per user, "the two metrics that were most critical for managing

[8]Revenue data that are systematically collected, called 'in-app' revenue, miss most advertising revenue and miss most revenues of big, AI-intensive apps. (Recall from the left panel of figure 1 that 70% of revenues are from ads.) For example, in 2020, the Facebook app had almost no in-app revenue despite its parent company Meta reporting $84 billion in digital ad revenues. As Goldfarb and Tucker (2019, page 20) note, "many of the largest online companies — in terms of revenues, profits, and users — are advertising-supported."' Not surprisingly, then, we estimate that in-app revenue is no more than one-third of total revenue and likely closer to one-tenth. We base this on figure 1, Sensor Tower (2021), and Analysis Group (2023).

the business ... That, in essence, determined how much we were worth." Users meant eyeballs and eyeballs meant ad revenue. As the old adage goes, "If you are not paying for the product, you are the product."

Turning to the history of AI, a brief narrative can be built around the contributions of the three 'godfathers of AI.' Their pre-2008 contributions include backpropagation (1986, co-authored by Hinton), convolutional neural networks (1998, co-authored by LeCun and Bengio), and deep learning and generative models (2006, co-authored by Hinton). With very few exceptions, AI was not used in commercial applications before 2012 (Agrawal, Gans, and Goldfarb, 2018) and worldwide funding for AI did not take off until 2013 Aral (2021, figure 3.2). In that year, Hinton's team scored a high-profile success in the Imagenet contest and the next year Hinton and LeCun accepted positions at Google and Facebook, respectively. After 2012, AI diffused slowly into industry. This is documented in figure 3, where the vertical lines at 2015 indicate the start of our sample. Significant ML advances developed by the private sector accelerated starting in 2014. US postings for machine learning jobs accelerated starting in 2015. AI publications accelerated in 2018, after the commercial potential of AI was proven. Finally, the number of parameters in significant ML models, a common measure of AI sunk and fixed costs, rose after 2015 and especially after the appearance of transformer-based LLMs in 2018. The log scale in this panel obscures the fact that between 2019 and 2020 the measure grew 60-fold. This history means that firms investing in machine learning well before 2012 likely did so without a specific mobile app in mind. This will play into our IV strategy where we use an app developer's pre-2008 AI patent filings as a cost-shifter instrument for its post-2015 AI sunk and fixed costs.

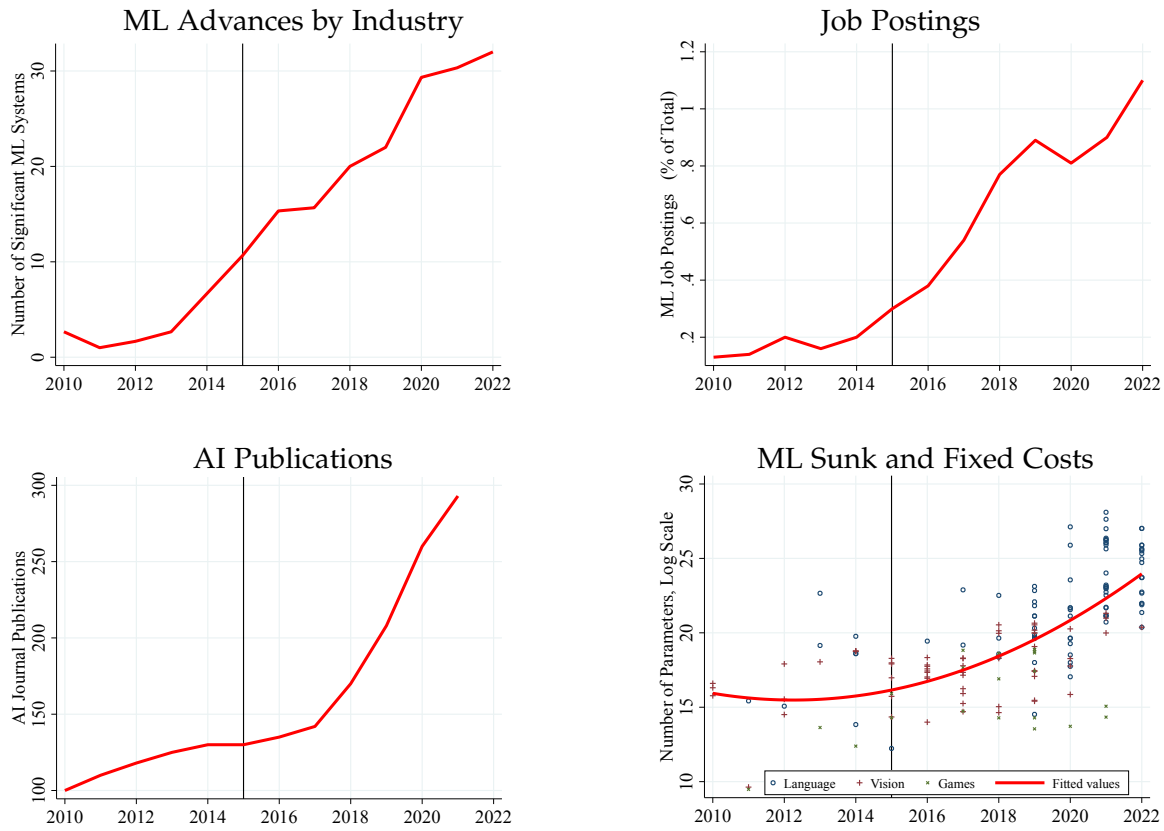
Finally, deep learning was not used in mobile apps until 2016. Pre-deep-learning AI was used in mobile apps after 2012 and continues to be used. For example, recommenders ("If you like this, then you will like that") often use ensemble methods such as clustering, tree classifiers, and nearest neighbours. See Lee and Hosanaga (2019) for recent examples. Thus, this paper is not exclusively about deep learning.

2. Data

2.1. Mobile App Data

We purchased proprietary data from Sensor Tower, which collects data on apps available on the App Store and Google Play Store. At the time of purchase, we concluded that Sensor Tower had the most accurate user data. Sensor Tower tracks monthly active users. We average across months to annualize the data. This average is our dependent variable.

Figure 3: Commercialization of AI



Notes: 'ML Advances by Industry,' 'AI Publications,' and 'ML Sunk and Fixed Costs' are from Maslej, Fattorini, Brynjolfsson, Etchemendy, Ligett, Lyons, Manyika, Ngo, Niebles, Parli, Shoham, Wald, Clark, and Perrault (2023). 'ML Advances by Industry' are defined as significant ML models developed by the private sector. 'ML Sunk and Fixed Costs' are defined as the number of parameters of significant machine learning systems. 'Job Postings' are from Goldfarb, Taska, and Teodoridis (2023) and Lightcast (`https://lightcast.io/resources/blog/lftce-04-13-2023`).

We have data for the top 2,000 app developers (firms) as measured by their 2014–2020 app downloads. It is essential for our study that we link these firms to their AI patents. We laboriously hand-matched each firm with a firm in the Bureau van Dijk (BvD) Orbis IP database, which gives us each firm's patent portfolio. We matched 1,276 of the top 2,000 firms and are very confident that unmatched firms are not in the BvD database. They are either government-owned, dissolved, or small private firms. The latter are mostly tiny game studios with a brief-lived success.[9] Our analysis is at the app level. The 1,276 firms together have 35,575 apps with user data. The median app has 65,000

[9]We initially matched using a variety of well-known algorithms such as FuzzyWuzzy, but these produced poor match rates. We also hand-matched as many of the next 3,000 firms as possible, but match rates were low, again because the firms are likely not in BvD. Note that Sensor Tower unifies and standardizes app ids and firm names, which makes matching easier and more accurate.

users.[10]

We link to Orbis IP in order to track all patents owned by a business group, including all of its subsidiaries e.g., Alphabet owns Google, Waze and DeepMind. Also, Sensor Tower tracks where each app was developed. For example, US-based Google bought Israel-based Waze and Sensor Tower correctly treats Waze as developed in Israel. Sensor Tower data are not always perfect so we improve their data using the Orbis M&A database to allocate apps made by subsidiaries to their country of origin. In the final data, the biggest producers of apps are the US (24% of all apps) and China (11%). More data details appear in Appendix A.

2.2. *AI Patent Data*

The World Intellectual Property Office (WIPO) has invested heavily in tools for identifying AI patents. See WIPO (2019) for an example. We exactly follow the WIPO methodology as described in WIPO (2018). For each patent we check if it meets one of three criteria.

1. The CPC code is an AI algorithm e.g., G06N 3/02 is neural networks.
2. The title or abstract has a keyword identifying it as an AI algorithm, where examples of keywords include deep learning, natural language processing, supervised learning, reinforcement learning, and gradient tree boosting.
3. The CPC code is potentially about AI and the title or abstract has an AI-related keyword e.g., GTL-013 is speech synthesis and is identified as AI if the title or abstract contain the keyword 'embedding.'

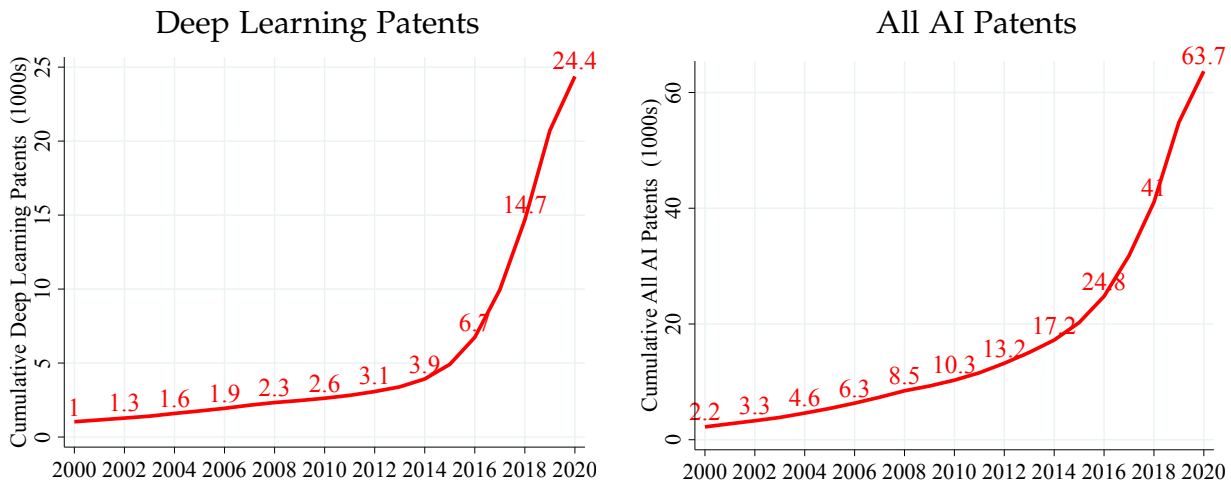
See WIPO (2018) for details. Our 1,276 firms have 10,144,089 patent filings, of which 63,679 are AI patent filings.

A natural question is the extent to which our AI patents cover deep learning as opposed to machine learning more generally. We do not have an exact answer, but a simple check suggests that the majority deal with deep learning. We computed the frequency of keywords in our AI patents and grouped them as deep learning (DL), non-DL, and unclassified. The unclassified patents typically only have the general keywords "artificial intelligence" or "machine learning." Of the classified patents, 90% have keywords that are unambiguously DL.

The left panel of figure 4 plots the cumulative number of patents that are unambiguously deep learning. In 2000, just under 100 such patents were filed annually and the

[10]Online appendix figure A2 shows that, after netting out the top 100 blockbuster apps, our sample of apps is representative of the apps of the top 5,000 app developers.

Figure 4: Rise of Deep Learning



Notes: The figure plots the cumulative number of patent filings by firms in our sample from 1992 to year t. The left panel is for patents that we *unambiguously* classify as deep learning patents. The right panel is for all AI patents.

cumulative filings hit the 1,000 mark. By the end of our sample, over 6,000 were filed annually and the cumulative was 24,370.[11]

2.3. *Linking AI Patents to Apps*

In this subsection, we explain how we link AI patents to mobile apps. We assume that the reader is not overly familiar with LLMs. Details for experts appear in Appendix B. App descriptions are consumer-facing while patent texts are engineering- and legal-facing. They thus do not share common words and so cannot be linked using older word frequency techniques such as TFID. An LLM is needed. Most modern LLMs are based on transformers (Vaswani, Shazeer, Parmar, Uszkoreit, Jones, Gomez, Kaiser, and Polosukhin, 2017). A transformer extracts features from a user-submitted text and represents these features as a numeric vector called an embedding. Think of the vector as an arrow connecting the inputted text to one of the LLM's subject areas. The subject areas are created during the very expensive training stage of the model, which is why LLMs are often called pre-trained models. Note that the embedding is not about the user-submitted text per se, but about where that text is located within the pre-trained model's subject

[11]One objection to patent data is that firms do not patent much and instead rely on trade secrets. This objection is misleading. The right panel of figure 4 shows that there are indeed many AI patent filings. Further, these grew exponentially once the commercial potential of AI became clear. Second, a senior Google executive told us that while Google does not file patents to prevent infringement of its technology, Google definitely files patents to deter other firms from suing Google and to cross-licence in exchange for other firms' technologies. Third, small firms use patents to raise capital (Hochberg, Serrano, and Ziedonis, 2018). Thus, while AI patents may be weak as protection against infringement, they are still of great value to firms large and small and so paint a picture of firms' AI research.

areas. To use a familiar example, consider ChatGPT. 'P' stands for pre-trained and 'T' stands for transformer. The user does not train the model. Instead, the user submits a text query to the pre-trained model, and the model returns a numeric embedding. A novelty of ChatGPT is that the numeric embedding is then rebuilt as text that the user can read.

The LLM we use is Google's BERT. BERT was widely considered to be the best LLM at the time we started our research long before the release of ChatGPT. It is still considered to be a top LLM.[12] As recommended by Google (see Devlin, Chang, Lee, and Toutanova, 2018), and following standard practice, we interpret the cosine of the angle between two embeddings ('cosine similarity') as the degree to which two embeddings point to the same subject area.[13] We can now define what we mean by the AI deployed in an app. Let a and p index apps and patents, respectively, and let ρ_{ap} be the cosine similarity between the embeddings of a's app description and p's patent text. ρ_{ap} is our measure of the AI in patent p deployed by app a.

With 35,575 apps and 63,679 AI patents, we have computed 2.4 billion ρ_{ap}. There is some randomness to LLM responses (think of what happens when you ask ChatGPT to regenerate an answer). We therefore set a threshold $\overline{\rho}$ below which a cosine similarity is deemed to be non-positive i.e., below which we conclude that the app does not use information from the AI patent. At risk of an abuse of notation, we redefine ρ_{ap} so that it is zero whenever $\rho_{ap} < \overline{\rho}$. As our baseline we choose $\overline{\rho} = 0.2$, but we will show that our results hold for $\overline{\rho} = 0.0, 0.1$.

2.4. Key Regressor: The App-Level Measure of AI Deployment

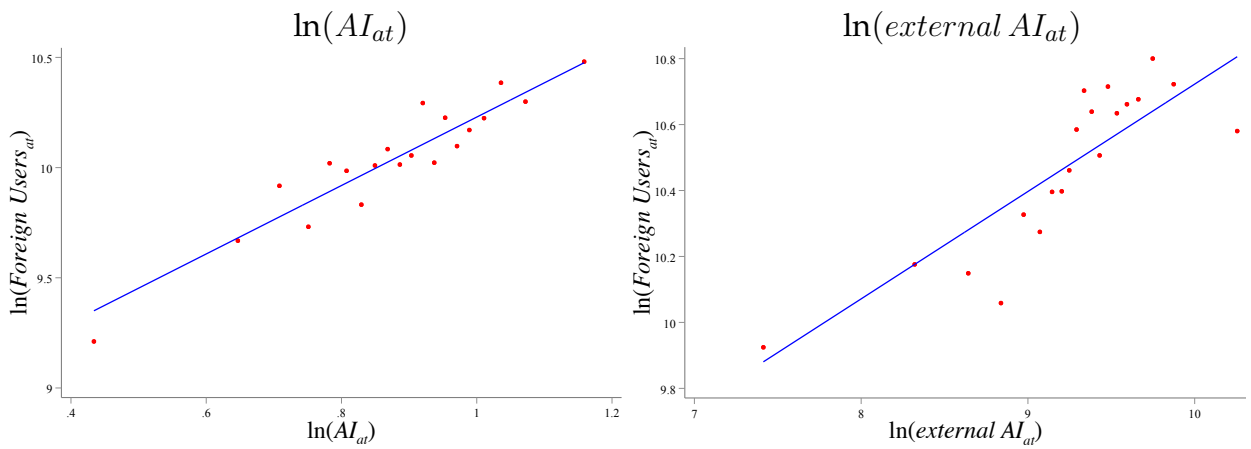
Our interest lies in regressing the number of users of an app on the AI deployed in the app. We therefore need to aggregate the ρ_{ap} across patents to get to the app level. Let $\mathcal{P}_{at}$ be the set of AI patents filed between 1992 and t and owned by the developer of app a. Our measure of the AI deployed in app a in year t is the sum of the ρ_{ap} over this set:

$$AI_{at} = \sum_{p \in \mathcal{P}_{at}} \rho_{ap} \tag{1}$$

[12] ChatGPT has more public recognition than BERT because of the former's user-friendly interface. However, within the industry it is widely known that the transformer technology was invented at Google (Vaswani *et al.*, 2017), was first deployed as BERT, and is the basis for Google's lead in the field. In short, BERT is a major LLM produced by the largest industry player. If Google is less in the spotlight, it is in part because Google has been cautious about releasing a potentially dangerous technology. See Financial Times (2023). Also, in small-scale experiments with ChatGPT, we found that BERT performed slightly better for our purposes.

[13] The cosine similarity between two embedding vectors B and B' with elements b_i and b'_i, respectively, is $\Sigma_{i=1} b_i b'_i / (\Sigma_i b_i^2 \Sigma_i b_i'^2)^{1/2}$ where by construction of the embeddings $\Sigma_i b_i^2 = \Sigma_i b_i'^2 = 1$. That is, cosine similarity is an uncentered correlation.

Figure 5: Bivariate Plots



Notes: The plots are bin scatters from the OLS regression of $\ln(Foreign\,Users_{at})$ on firm fixed effects, industry-year fixed effects and either $\ln(AI_{at})$ or $\ln(external\,AI_{at})$. $\ln(external\,AI_{at})$ is described in section 6 below.

and $AI_{at} = 0$ if the developer has no AI patents. One can interpret (1) as a weighted sum of patents, weighted by the strength of the connection between the app and the patent. Unrelated patents ($\rho_{ap} = 0$) receive zero weight while highly related patents ($\rho_{ap} = 1$) receive a large weight. 27% of apps have positive $AI_{a,2020}$.

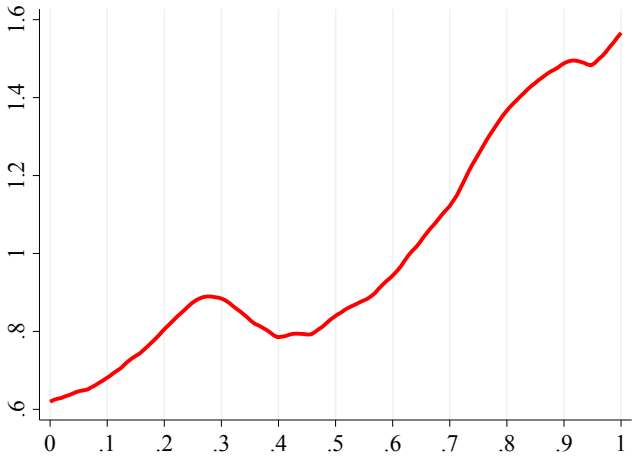
The left panel of figure 5 is a bin scatter from a regression of the log of foreign users on $\ln(AI_{at})$. Firm and industry-year fixed effects are included. As is apparent, AI_{at} is correlated with the number of foreign users.

2.5. *Validation*

In this section we consider validation exercises for our ρ_{ap} and AI_{at}. We start with an anecdote about an exchange we had at the NBER Digital Economics and AI Tutorial. We explained our research to a data scientist working on the Google Map app. His initial reaction was that our ρ_{ap} must be invalid because he never looks at patents. We then showed him the Google patent with the highest cosine similarity to Google Maps. He looked shocked and responded that this was exactly what he was doing!

Our first validation exercise identifies apps known to use AI and shows they have high AI_{at}. Machine learning is rarely done on the phone. Instead, APIs send data and queries to a cloud-based server where the machine learning model is housed and responses are sent back to the phone. This makes it difficult to know with certainty whether an app uses AI. However, a small number of apps do deep learning on-phone and this can be tracked. Each mobile app comes with an installation file (a DMK) and

Figure 6: Ground Truth: On-Phone Deep Learning



within this file are software developer kits (SDKs) that may include deep learning SDKs such as TensorFlow Lite from Google. Xu, Liu, Liu, Lin, Liu, and Liu (2021) examined 16,500 app DMKs available on Google Play in 2018 and found deep learning SDKs on 1.2% of them. Following their method, we downloaded 6,374 DMKs and found deep learning SDKs on 1.1% of them. This gives us ground truth that these apps use AI. We calculated percentiles of the distribution of AI_{at} for these 6,374 apps and, in figure 6, we plot the probability density function (pdf) of these percentiles for the subsample of apps with on-phone deep learning. If AI_{at} were unrelated to on-phone deep learning, the pdf would be a horizontal line at $y = 1$. In fact, the pdf is heavily skewed to the right, which means that apps known to use deep learning have among the highest values of AI_{at}.

An anomaly in figure 6 is the bump at 0.3. This is in part an artifact of using a binary indicator of on-phone deep learning (DL). In figure 6, apps to the right all introduced DL in 2019–2020 while apps to the left mostly introduced DL in 2016-2017 (the earliest years of DL commercialization). This suggests that apps on the left use DL in a rudimentary way, which is consistent with a lower AI_{at} percentile. Nevertheless, some of the anomaly is classical measurement error stemming from randomness in the responses of the LLM. In the presence of classical measurement error, the probability limit of the OLS estimator is biased downward while the probability limit of the IV estimator is unaffected (projecting AI_{at} onto the instrument purges the error). *This will create a tendency for IV to exceed OLS.*

Our second validation exercise exploits the fact that if two app descriptions have embeddings with a high cosine similarity, then they deal with the same subject matter and hence share the same App Store category. Note that an app's description does *not* include its App Store category so the category information does not enter into our embeddings.

We validate using a procedure recommended by OpenAI (https://github.com/openai/openai-cookbook/blob/main/examples/Customizing_embeddings.ipynb). For each pair of apps a and a', we know whether they are in the same App Store category ($d_{aa'} = 1$) or not ($d_{aa'} = 0$). Using our app description embeddings, we construct cosine similarities $\rho_{aa'}$ for each app pair. We then use agglomerative clustering on the $\rho_{aa'}$ to construct a dummy $\hat{d}_{aa'}$ for whether the two apps are in the same cluster. We find that $d_{aa'}$ and $\hat{d}_{aa'}$ agree for 88.1% of app pairs (standard error 0.044%). Details appear in online Appendix A. Thus, the LLM is able to use app descriptions to precisely predict whether or not two apps are in the same app category. By implication, the LLM can also predict AI deployment.[14]

3. Modelling the Estimating Equation and an Instrument for AI_{at}

In this section we develop a model of mobile app trade and AI adoption decisions that provides a theory-consistent estimating equation and instrument.

3.1. Consumers

An app a produced in country i is used by a consumer in country n. A consumer must choose a single app or no app (the outside option). As in our data, apps are free. The utility an individual obtains from app a is $U_{ani} = \ln(\delta'_a) + \epsilon_{ani}$ where $ln(\delta'_a)$ is mean utility and ϵ_{ani} is extreme value I (cumulative distribution is $e^{-e^{\epsilon}}$). The outside option is indexed by $a = 0$ and yields mean utility δ'_{0n}. It is costlessly produced and freely available e.g., a public amenity such as a park. As is well known, the share of country-n consumers choosing app a is[15]

$$\delta'_a/\mathcal{U}_n \quad \text{where} \quad \mathcal{U}_n = \delta'_{0n} + \sum_i \sum_{a \in \mathcal{A}_{ni}} \delta'_a \tag{2}$$

and $\mathcal{A}_{ni}$ is the set of apps produced in i and available in n. In what follows, we drop a subscripts unless needed.

3.2. Firms

A firm's mean utility has two components, $\delta' = \alpha\delta$, where δ is an exogenous component and α results from an endogenous investment in AI that improves mean utility. AI

[14]The need to validate BERT can be exaggerated. In the year following the release of ChatGPT, (1) Geoff Hinton left Google to caution the world that AI is evolving too quickly and (2) Microsoft, Google, and Nvidia added $2 trillion to their combined market cap. Data scientists and investors have voted with their feet on the validity of LLM output.

[15]This share is usually expressed as $e^{\ln \delta'_a} / \left(e^{\ln(\delta'_{0n})} + \Sigma_i \Sigma_{a \in \mathcal{A}_{ni}} e^{\ln \delta'_a} \right)$ with $\delta'_{0n} = 1$ so that $e^{\ln(\delta'_{0n})} = 1$.

adopters use AI scientists for all of their activities, meaning sunk, fixed and variable costs. Each firm incurs an entry sunk cost f_{ei}^A and draws a δ from the Pareto distribution $G(\delta) = 1 - \delta^{-\gamma}$ where $\gamma > 1$. If the firm decides to produce, it also incurs a fixed cost f_i^A. As we saw in the right-hand panel of figure 1, selection into exporting is weak (most firms export) so we assume that if a firm operates, it operates in all markets. As Goldfarb and Trefler (2019) and others have noted, the countries with large numbers of AI scientists are also the countries that develop the large models plotted in the 'ML Sunk and Fixed Costs' panel of figure 3. We capture this in a reduced-form way by assuming that the number of AI scientists used for sunk and fixed costs is increasing in the country's endowment L_i^A of AI scientists:

$$f_{ei}^A = (L_i^A)^\psi f_{ei} \quad \text{and} \quad f_i^A = (L_i^A)^\psi f_i \tag{3}$$

where $\psi \in [0,1)$ controls the strength of the effect and f_{ei} and f_i are positive constants.

A firm can raise its mean utility by hiring AI scientists to improve the quality of the app. Specifically, the firm can raise mean utility from $\delta' = \delta$ to $\delta' = \alpha\delta$ by hiring $\frac{\eta-1}{\eta}\alpha^{\frac{\eta}{\eta-1}}$ AI scientists. We assume $\eta > 1$.

Firms earn revenue by displaying ads. We assume that firms display one ad per user.[16] When advertising to a user in country n, the firm receives revenue p_n per ad and hence per user. The number of users of the app in country n is the total number of consumers ($\mathcal{L}_n$) times the firm's share of consumers ($\alpha\delta/\mathcal{U}_n$). Hence, the firm's revenue in market n is $p_n[\alpha\delta/\mathcal{U}_n]\mathcal{L}_n$. The firm's total revenue is $\alpha\delta P$ where

$$P \equiv \sum_n p_n \mathcal{L}_n / \mathcal{U}_n \,. \tag{4}$$

The firm's profit function is

$$\pi_i^A(\alpha,\delta) = \alpha\delta P - w_i^A \left(\frac{\eta-1}{\eta} \alpha^{\frac{\eta}{\eta-1}} \right) - w_i^A f_i^A$$

where w_i^A is the wage of AI scientists. Maximizing profits with respect to α subject to $\alpha \geq 1$, the interior solution is

$$\alpha(\delta) = \delta^{\eta-1} \left(\frac{w_i^A}{P} \right)^{-(\eta-1)} . \tag{5}$$

That is, the higher is a firm's quality δ, the more it sells and hence the more it invests in AI to improve its app. This is the standard scale effect in innovation (e.g., Schmookler, 1954, Lileeva and Trefler, 2010). Further, the lower are AI wages, the more intensively AI is used. That is, demand for AI inputs slopes downward.[17]

[16] It does not change our results if the number of ads a firm displays is proportional to the popularity of the app e.g., proportional to a power function of δ or $\alpha\delta$.

[17] $\eta f_i^A (L_i^A)^\psi > 1$ is a necessary and sufficient condition for an interior solution. See Appendix C for a proof. We assume this condition holds.

Plugging $\alpha(\delta)$ back into the expression for profits we obtain

$$\pi_i^A(\delta) = w_i^A \left\{ \delta^\eta \left(\frac{w_i^A}{P} \right)^{-\eta} \frac{1}{\eta} - f_i^A \right\}. \tag{6}$$

A firm produces if profits are positive. Defining the zero-profit cutoff δ_i^A implicitly by $\pi_i^A(\delta_i^A) = 0$, a firm produces if $\delta > \delta_i^A$ where

$$\delta_i^A = \frac{w_i^A}{P} \left(\eta f_i^A \right)^{1/\eta}. \tag{7}$$

Higher wages or fixed costs make it tougher to survive.

3.3. Equilibrium in App and Labour Markets

The free entry condition (expected profits are zero) is

$$\int_{\delta_i^A}^{\infty} \pi_i^A(\delta) dG(\delta) = w_i^A f_{ei}^A \tag{8}$$

where we require $\gamma > \eta$ for the integral to be finite. Let M_i^A be the mass of firms who pay the sunk cost. The labour-market clearing condition equates labour supply L_i^A with labour demand:

$$L_i^A = M_i^A \left\{ f_{ei}^A + \int_{\delta_i^A}^{\infty} \left[f_i^A + \frac{\eta - 1}{\eta} \alpha^{\frac{\eta}{\eta-1}} \right] dG(\delta) \right\}. \tag{9}$$

The right-hand side of (9) states that each firm must incur the sunk cost f_{ei}^A and, if it operates ($\delta > \delta_i^A$), the fixed cost f_i^A as well as the AI improvement cost. As in Fajgelbaum, Grossman, and Helpman (2011) we treat M_i^A as if it were a continuous variable to facilitate the exposition. The above equations imply

$$\delta_i^A = \left(\frac{f_i}{f_{ei}} \frac{\eta}{\gamma - \eta} \right)^{1/\gamma}, \qquad \frac{w_i^A}{P} = (L_i^A)^{-\psi/\eta} \kappa, \qquad M_i^A = \frac{(L_i^A)^{1-\psi}}{\gamma f_{ei}} \tag{10}$$

where $\kappa \equiv [\eta f_i]^{1/\gamma - 1/\eta} [(\gamma - \eta) f_{ei}]^{-1/\gamma}$. See Appendix C for a proof. Importantly for our IV strategy, an AI-abundant country (large L_i^A) has a low relative wage w_i^A / P.[18]

Re-introducing an app subscript on δ in order to identify the app, let $User_{an} = [\alpha(\delta_a)\delta_a / \mathcal{U}_n] \mathcal{L}_n$ be the number of country-n users of app a. We now state our main result.

[18] The negative relationship between w_i^A and L_i^A requires $\psi > 0$, which we justified on empirical grounds; however, the negative relationship is intuitive and can be obtained in other ways.

Theorem 1 *Country n's demand for app a is*

$$\ln(Users_{an}) = \ln \alpha(\delta_a) + \ln(\mathcal{L}_n/\mathcal{U}_n) + \ln \delta_a \,. \tag{11}$$

The optimal AI deployed in app a is

$$\alpha(\delta_a) = \left[\delta_a \left(L_i^A\right)^{\psi/\eta} / \kappa\right]^{\eta-1} . \tag{12}$$

See Appendix C for a proof. Equation (11) is the basis for our estimating equation. Equation (12) is the basis for our IV strategy. We know from equation (10) that an AI-abundant country has a low relative wage for AI scientists. Equation (12) states that firms in an AI-abundant country thus use more AI-intensive choice of techniques.

The above is not a general equilibrium model because we have not described the demand for advertising and so have not solved for the price of ads in each country, p_n. In online Appendix B we introduce a CES goods sector that uses the Arkolakis (2010) advertising technology. This allows us to solve for the p_n and close the model.

3.4. The Trade Prediction and a Model-Based Instrument for AI Adoption

We make two generalizations of the model to match important features of the data. First, consumers often have many apps on their phones. We realistically assume that there are many app industries (e.g., social networking and gaming) and that consumers either buy one app or no app within each industry. We index app industries by c. (App industries are called $c$ategories in the App Store.)

Adding c subscripts to the theorem 1 parameters $(\eta_c, \gamma_c, \psi_c, f_{ic}, f_{eic}, \delta_{ac})$, for an app a in industry c we can rewrite (11) as

$$\ln Users_{anc} = \ln \alpha_c(\delta_{ac}) + \lambda_{nc} + \varepsilon_{anc} \tag{13}$$

where $\lambda_{nc} \equiv \ln(\mathcal{L}_n/\mathcal{U}_{nc})$ is a country-industry fixed effect and $\varepsilon_{anc} \equiv \ln \delta_{ac}$ is a residual. When country n is not the producer of app a, this is a regression of foreign users on AI deployment $\alpha_c(\delta_{ac})$.

The endogeneity of $\alpha_c(\delta_{ac})$ is apparent. Firm heterogeneity is due to heterogeneity in mean utilities $\ln(\delta_{ac})$ i.e., heterogeneity in demand. A high level of demand is associated with both high AI deployment and a high residual. We thus need a supply-side cost-shifter of AI as an instrument for AI deployment. Equation (12) is the basis for this instrument. To see this, we note that some app industries are more amenable to AI than others. The marginal cost of improving an industry-c app is $w_i^A \alpha_c^{1/(\eta_c - 1)}$, which is decreasing in η_c. That is, η_c is a measure of how inexpensive it is to improve an app using AI and thus stands in for the AI-intensity of the industry. From theorem

1, $\ln \alpha_c(\delta_{ac})$ is supermodular in (L_i^A, η_c).[19] To interpret this it is useful to think about the Heckscher-Ohlin model in which the AI-abundant country exports the AI-intensive industry. Here, L_i^A is like the AI abundance of country i and η_c is like the AI intensity of industry c. Supermodularity implies that AI investments $\alpha_c(\delta_{ac})$ will be high in industry c in country i when c is *AI-intensive* and i is *AI-abundant*. Thus, $\eta_c \times L_i^A$ is a model-based instrument for AI deployment.

In our model all firms invest in AI, but in the data, many firms have no AI patents. To accommodate this data feature, in online Appendix C we extend the model to allow firms the choice of whether or not to deploy AI and in equilibrium not all firms deploy. This is a second source of endogeneity, namely, the decision of whether or not to do any AI. To instrument for this endogeneity, recall that mobile apps did not exist before mid-2008, machine learning was in its infancy in the 2000s (figure 4), its commercial potential was only beginning to be recognized in 2012 (figure 3), and the first trickle of mobile apps using deep learning appeared only in 2016. We conclude from this that AI patents from 2008 or earlier were not directed at improving the quality of apps during 2015–2020, that is, they do not belong in the second stage and so meet the exclusion restriction. We operationalize this as follows. For any app a, let $f(a)$ denote the firm that developed a. Let $D_{f(a),\tau}$ be a dummy equal to one if the firm producing a had filed an AI patent on or before year τ. Then our instrument becomes

$$\eta_c \times L_i^A \times D_{f(a),\tau}\,.$$

We will show that our results are the same for any of the years $\tau = 2006, \ldots, 2011$. We use 2008 in our baseline specification.

Finally, in online Appendix D, we model $D_{f(a),\tau}$ by introducing incumbents into the model. These are firms who, in a pre-period τ before the model opens, developed AI research capabilities and so already incurred the sunk costs f_{ei}^A. When the model opens they can thus deploy AI in mobile apps without incurring additional sunk costs. We show that this does not alter theorem 1. However, it leads to the new and sensible prediction that incumbents (firms with $D_{f(a),\tau} = 1$) are more likely to be AI adopters. This result provides a theory-consistent rationale for adding $D_{f(a),\tau}$ to our instrument.

3.5. *Construction of the Instrument*

AI intensity η_c: Consider an app a developed by firm $f = f(a)$ and part of industry $c = c(a)$. An obvious definition of industry is the App Store categories such as social networking and games. There are 15 categories. Let $\mathcal{A}_{c(a),t}$ be the set of apps in industry $c(a)$ in year t, excluding all apps developed by $f(a)$. We exclude these apps in order to

[19] $\frac{\partial^2 \ln \alpha_c(\delta_{ac})}{\partial \ln(L_i^A)\partial \ln(\eta_c)} = \frac{\partial}{\partial \ln(\eta_c)}\left[\frac{\partial \ln \alpha_c(\delta_{ac})}{\partial \ln(L_i^A)}\right] = \frac{\partial}{\partial \ln(\eta_c)}\left[\psi(\eta_c - 1)/\eta_c\right] = \psi/\eta_c^2 > 0.$

purge the instrument of any information about app a that might be correlated with the second-stage residual. We construct $\eta_{c(a),t}$ in two steps. For each app $a' \in \mathcal{A}_{c(a),t}$, we average the $\rho_{a'p}$ across all patents filed on or before year t by firm $f(a')$. Then we average across all apps a' in $\mathcal{A}_{c(a),t}$. In simple words and setting aside details, $\eta_{c(a),t}$ *is the average value of the AI embodied in the apps of industry* c.

This way of calculating $\eta_{c(a),t}$ suffers from the fact that App Store categories are very broad. For example, battle royale games such as Call of Duty and Fortnite usually use AI while puzzle games such as Candy Crush Saga and Royal Match usually do not. To deal with this, recall from our second validation exercise that if two apps a and a' have a large cosine similarity $\rho_{aa'}$, they are correctly predicted to be in the same App Store category. We can therefore use the $\rho_{aa'}$ to construct narrower industries. To this end, redefine $\mathcal{A}_{c(a),t}$ as the set of ten apps a' with the largest $\rho_{aa'}$ and calculate $\eta_{c(a),t}$ as before, but with $\mathcal{A}_{c(a),t}$ redefined in this way. Our results are not sensitive to using many more or many less than 10 apps. We use this definition of $\mathcal{A}_t$ in our baseline results, but the results are almost identical when we define $\mathcal{A}_{c(a),t}$ using App Store categories.

To give the reader a sense of the $\eta_{c(a),t}$ we averaged them across all a within App Store category c and across all years. The ranking of categories by this average is, from largest to smallest: Developer Tools, Business, Social Networking, Education, Sports, Books, Entertainment, Finance, Music, Weather, Shopping, Navigation, Lifestyle, Food and Drink, and Games. This accords well with the practitioner literature on mobile apps.

AI abundance L_i^A: We need a measure of the supply of AI scientists or, more generally, a measure of access by app developers to AI expertise in country i. We start with Microsoft Academic Graph (MAG), which is a comprehensive database of academic articles. We extract all articles labelled in MAG as "Artificial Intelligence" and published between 1992 and 2020. For each article, MAG lists the authors' institutional affiliations as well as citations by year. We sum across citations from 1992 to year t and across all institutions located in country i. This gives us a citation-weighted number of AI publications by country and year, which we use to proxy for the number of AI scientists L_{it}^A. The logic is that AI researchers train university students who then go on to work in industry. More researchers means more trained AI scientists in industry. Counting citations rather than authors adjusts for the quality of the author and hence for the quality of training. Online appendix figure A3 plots $L_{i,2008}^A$ against GDP per capita in 2008. As expected, the top countries are USA, China, UK, Germany and Canada.[20]

[20]MAG is a large and complex database so we have not experimented with different measures of L_{it}^A.

Instrument: Putting these elements together, our instrument for AI_{at} is

$$Z_{at} = \underbrace{\eta_{c(a),t}}_{\text{AI intensity}} \times \underbrace{\ln(1 + L^{A}_{i(a),t})}_{\text{AI abundance}} \times \underbrace{D_{f(a),2008}}_{\text{initial AI status}} \tag{14}$$

where $f(a)$ is the developer of app a, $c(a)$ is the industry of app a, and $i(a)$ is the country where app a was developed. $L^{A}_{i(a),t}$ has a small number of zeros so we use $\ln(1 + L^{A}_{i(a),t})$.

4. Result 1: The Impact of AI on Exports of Digital Services

Our first of three key results is about how AI deployment AI_{at} makes an app more attractive to foreign users. Since AI_{at} is at the app-year level, we start with app-year level regressions. To this end, in equation (13) we treat n as a single aggregate foreign country. We also add year subscripts t and drop industry subscripts c unless necessary. Then $Users_{anc}$ becomes $ForeignUsers_{at}$, the users of app a that are located outside of the country where app a was developed. Our regression is

$$\ln(ForeignUsers_{at}) = \beta \ln(1 + AI_{at}) + \lambda_f + \lambda_{ct} + \varepsilon_{at} \tag{15}$$

where there are 35,575 apps and six years $t = 2015, \ldots, 2020$. We only include observations with positive foreign users, leaving us with 125,486 observations.[21] The theory states that there should be industry-year fixed effects λ_{ct} where industry is the 15 App Store categories. We also always include firm fixed effects λ_f. Given that our 'treatment' AI_{at} is at the at level, we cluster at the a level to allow for serially correlated errors.

4.1. Baseline

The results appear in table 1. The panels from top to bottom are OLS, IV, first stage, and reduced form. To fix ideas, consider column 2 of table 1 where we have firm and industry-year fixed effects. In the third panel, the first-stage coefficient on the instrument is 0.60 (0.02) and has the predicted positive sign i.e., a high value of the instrument means a low cost of deploying AI and hence a high value of AI_{at}. The instrument has a large weak-instruments F-statistic (1,384). In the top panel, the OLS coefficient on $\ln(1 + AI_{at})$ is 1.48 (0.14). In the next panel, the IV coefficient is 2.67 (0.44). Both have the predicted positive sign, meaning that AI deployment raises the mean utility of apps. As noted above, classical measurement error likely biases OLS to zero, which explains why OLS

[21] We have an unbalanced panel because many apps were introduced after 2015 and a few apps did not survive until 2020. Only 9% of the unbalanced panel has $ForeignUsers_{at} = 0$ so we do not use PPML and using $\ln(1 + ForeignUsers_{at})$ on the full unbalanced panel makes no difference.

Table 1: Foreign Users and Internal-to-the-Firm AI Deployment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS: ln(*Foreign Users* $_{at}$)						
ln(1+*AI* $_{at}$)	1.58*	1.48*	1.67*	1.69*	1.20*	1.58*	1.24*
	(0.13)	(0.14)	(0.14)	(0.15)	(0.12)	(0.15)	(0.13)
ln(1+*undirected AI* $_{at}$)					-0.65*		-0.64*
					(0.11)		(0.11)
ln(1+*non AI* $_{ft}$)						-0.95*	-0.35
						(0.34)	(0.34)
Obs.	125,486	125,486	125,467	125,024	125,486	125,486	125,486
R^2	0.25	0.25	0.25	0.27	0.25	0.25	0.25
FEs	*f, c, t*	*f, ct*	*f, ct, it*	*f, cit*	*f, ct*	*f, ct*	*f, ct*
	IV: ln(*Foreign Users* $_{at}$)						
ln(1+*AI* $_{at}$)	2.78*	2.67*	2.82*	3.10*	2.44*	2.82*	2.59*
	(0.41)	(0.44)	(0.49)	(0.52)	(0.48)	(0.47)	(0.52)
ln(1+*undirected AI* $_{at}$)					-0.44*		-0.38*
					(0.13)		(0.14)
ln(1+*non AI* $_{ft}$)						-2.16*	-1.77*
						(0.56)	(0.63)
Weak Instrument F (KP)	1,607	1,384	1,176	1,082	1,414	1,245	1,261
	First Stage: ln(1+*AI* $_{at}$)						
Z_{at}	0.63*	0.60*	0.56*	0.56*	0.55*	0.56*	0.51*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
ln(1+*undirected AI* $_{at}$)					-0.14*		-0.16*
					(0.01)		(0.01)
ln(1+*non AI* $_{ft}$)						0.90*	0.98*
						(0.04)	(0.04)
	Reduced Form: ln(*Foreign Users* $_{at}$)						
Z_{at}	1.76*	1.60*	1.59*	1.73*	1.35*	1.59*	1.31*
	(0.26)	(0.26)	(0.28)	(0.29)	(0.26)	(0.27)	(0.26)
ln(1+*undirected AI* $_{at}$)					-0.79*		-0.81*
					(0.11)		(0.11)
ln(1+*non AI* $_{ft}$)						0.38	0.77
						(0.32)	(0.33)

Notes: This table reports estimates of equation (15). Each observation is an app a (35,575 apps) in a year t (2015,...,2020). The dependent variable is the log number of foreign users of app a in year t. The key regressor $\ln(1 + AI_{at})$ is the AI deployed in app a (see equation 1). The instrument for $\ln(1 + AI_{at})$ is the Heckscher-Ohlin-like cost shifter Z_{at} (see equation 14). The four panels are OLS, IV, first-stage, and reduced-form. For the fixed effects, consider an app a in industry c developed by firm f located in exporting country i. Columns 1–4 contain fixed effects for (f,c,t), (f,ct), (f,ct,it), and (f,cit), respectively. Column 5 adds $\ln(1 + undirected\, AI_{at})$, which is the sum of the ρ_{ap} across patents with $\rho_{ap} < 0.2$ i.e., across patents p that are not cosine similar to a. Column 6 adds the count of non-AI patents owned by f. Standard errors are clustered at the app level. * indicates 1% significance.

is smaller than IV. In addition, we show below that heterogeneous responses likely also contribute to OLS being smaller than IV. In the bottom panel of table 1, the reduced-form coefficient on the instrument Z_{at} is 1.60 (0.26) and so also has the correct sign.

Turning to magnitudes, AI_{at} is constructed from the underlying ρ_{ap} and so has no intrinsic meaning outside the context of the LLM. We thus scale $\ln(1 + AI_{at})$ by its interquartile range (iqr) so that β gives the impact of a one iqr change in $\ln(1 + AI_{at})$ on the dependent variable. More generally, all independent variables in what follows are scaled by their iqrs so that their magnitudes are easily interpreted and compared across regressors. *From column 2 of the IV panel, a one iqr increase in the AI deployed in an app leads to a 2.67 log point increase in its foreign users or a more than 10-fold increase ($e^{2.67} = 14$). This is the headline number for our first result.*

In columns 1–4 we introduce various fixed effects. In column 3 we add exporter-year fixed effects (it) where exporter is the country where the app was developed. In column 4 we add industry-exporter-year (cit) fixed effects. Reassuringly, across columns 1–4 the IV coefficients on $\ln(1 + AI_{at})$ vary within the narrow band from 2.67 to 3.10.

Column 4 is useful for thinking about whether our results are driven by AI algorithms or by the inherent scalability of apps that use AI e.g., social networking.[22] If scalability is driven by the demand side (e.g., network externalities as in Rosen, 1981), then it is orthogonal to our supply-side instrument and so is netted out by our IV strategy. If scalability is driven by the supply side either at the category level (e.g., social networking requires cloud computing overhead) or the category-exporter level (social networking apps can only be developed in large countries such as the US or China) then our cit fixed effects absorb scalabilty.

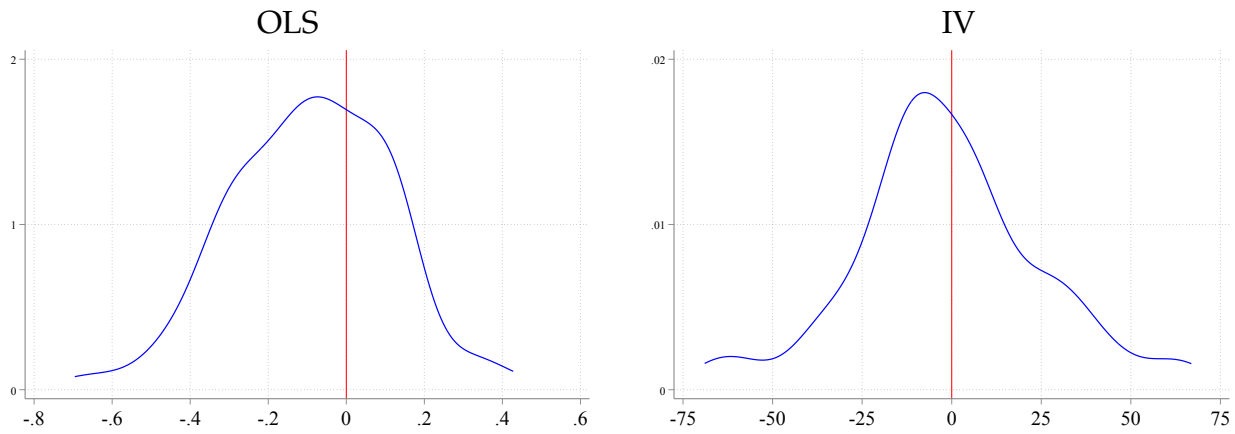
4.2. The Role Played by ρ_{ap}

We next turn to the role played by the ρ_{ap}. Again indexing apps by a and AI patents by p, AI_{at} is the sum of the ρ_{ap} across AI patents that are cosine similar to app a i.e., across patents p with $\rho_{ap} \geq 0.2$. The AI patents with $\rho_{ap} < 0.2$ are not directed toward the mobile app a and so should not have a positive effect on foreign users. We therefore construct a counterpart to AI_{at} based on these undirected patents and call it $undirected\, AI_{at}$.[23] If ρ_{ap} is playing its expected role, then $undirected\, AI_{at}$ should not have a positive effect on users. $\ln(1 + undirected\, AI_{at})$ is added as a regressor in column 5 of table 1. As expected, its coefficient is non-positive, so that the ρ_{ap} are playing their expected role. The negative coefficient on $undirected\, AI_{at}$ has an obvious explanation. If a firm is filing patents in

[22] See Agrawal *et al.* (2018) for a discussion of the link between AI and scalability. See Goldfarb and Trefler (2019) for modelling insights about the demand- and supply-side sources of scalability in AI.

[23] Formally, let $\mathcal{P}_{at}^{undirected}$ be the set of AI patents filed between 1992 and t, owned by the developer of app a, and for which $\rho_{ap} < 0.2$. Then $undirected\, AI_{at} \equiv \sum_{p \in \mathcal{P}_{at}^{undirected}} \rho_{ap}$.

Figure 7: OLS and IV Estimates When ρ_{ap} Are Randomly Drawn



Notes: We randomize the ρ_{ap}, recalculate the AI_{at}, and reestimate the model. We repeat this 100 times. The panels display the distribution of the 100 estimated coefficients on $\ln(1 + AI_{at})$ for OLS (left panel) and IV (right panel). These distributions are centred on zero and the coefficients are rarely statistically significant. This is a placebo test which shows that the ρ_{ap} are playing an important role.

areas unrelated to mobile apps then it is likely redirecting its scarce innovation resources away from mobile apps, thus reducing the attractiveness of its apps and hence reducing the number of foreign users.

It is possible that our result about AI_{at} is less about AI deployment in mobile apps and more about a firm's 'innovativeness' in general. Firm fixed effects capture time-invariant innovativeness. To examine time-varying innovativeness, in column 6 of table 1 we include a regressor that is the count of the non-AI patents filed by the firm between 1992 and t. The coefficient is not positive and its inclusion does not shrink the coefficient on $\ln(1 + AI_{at})$ so our result about AI_{at} is not driven by the overall inventiveness or patenting behaviour of the firm. Again, the negative coefficient is consistent with internal resource re-direction. In column 7 we include both undirected AI and non-AI and arrive at the same conclusions.

4.2.1. Two Placebos

In this subsection we consider two placebo tests. Recall that AI_{at} is a count of patents weighted by the ρ_{ap} (equation 1). We first show that when we replace the ρ_{ap} with randomized ρ_{ap} our results disappear. Consider the subsample involving firms that have AI patents i.e., for which the ρ_{ap} are defined. Let F_ρ be the empirical distribution of the ρ_{ap}. We replace each ρ_{ap} with a random draw from F_ρ, recalculate AI_{at} using the randomized ρ_{ap}, and reestimate the model. We repeat this 100 times and plot the distribution of the estimated coefficients in figure 7. The 100 OLS and IV coefficients are centred on zero. The IV coefficients are *never* statistically significant. The OLS coefficients

Table 2: A Placebo: Randomly Reallocating Patents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	IV: ln(*Foreign Users* $_{at}$)						
ln(1+*AI* $_{at}$)	2.97*	2.85*	2.92*	3.25*	2.61*	2.94*	2.71*
	(0.47)	(0.51)	(0.55)	(0.58)	(0.54)	(0.53)	(0.57)
ln(1+*placebo AI* $_{at}$)	-0.04	-0.04	-0.03	-0.04	-0.04	-0.03	-0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
ln(1+*undirected AI* $_{at}$)					-0.45*		-0.40*
					(0.13)		(0.14)
ln(1+*non AI* $_{ft}$)						-1.98*	-1.56*
						(0.50)	(0.57)
Obs.	125,486	125,486	125,467	125,024	125,486	125,486	125,486
R^2	0.25	0.25	0.25	0.27	0.25	0.25	0.25
FEs	*f, c, t*	*f, ct*	*f, ct, it*	*f, cit*	*f, ct*	*f, ct*	*f, ct*
Weak Instrument *F*	1,436	1,242	1,132	1,039	1,254	1,174	1,180

Notes: This table is identical to the IV panel of table 1 except for the addition of the placebo regressor $\ln(1 + placebo\, AI_{at})$. The OLS, first-stage, and reduced-form estimates appear in online appendix table A1. Standard errors are clustered at the app level. * indicates 1% significance.

are rarely significant and, even when significant, are so small (always less than 0.43) that their 99% confidence intervals never overlap with the 99% confidence interval for the OLS estimate using the actual $\ln(1 + AI_{at})$. We conclude from this placebo test that the ρ_{ap} are playing an important role.

In our second placebo test, we randomize patents. For firms that have patents, we replace a firm's AI patents with the AI patents of all other firms. Letting p' index the patents of all other firms, we compute the $\rho_{ap'}$ and sum across the p' to obtain a new variable $\ln(1 + placebo\, AI_{at})$. It equals zero if the firm has no AI patents. Table 2 reports the IV results when adding $\ln(1 + placebo\, AI_{at})$ to our estimating equation. Except for this addition, the table is identical to the IV panel of table 1. In all specifications, the coefficient on $\ln(1 + placebo\, AI_{at})$ is precisely estimated to be zero.[24] Further, its inclusion leaves unchanged the coefficients on our key regressor $\ln(1 + AI_{at})$.

Summarizing this section, we showed that the ρ_{ap} play a key role. Our results evaporate when we use AI patents with small ρ_{ap} (undirected AI), when we randomize the ρ_{ap}, or when we replace a firm's patents with those of other firms.

[24]The same conclusion holds for OLS. See appendix table A1. It also holds if, instead of adding $\ln(1 + placebo\, AI_{at})$, we replace $\ln(1 + AI_{at})$ with $\ln(1 + placebo\, AI_{at})$.

Table 3: Foreign, Domestic and All Users

	(1)	(2)	(3)
	ln(*Foreign Users* $_{at}$)	ln(*All Users* $_{at}$)	ln(*Domestic Users* $_{at}$)
		OLS	
ln(1+AI_{at})	1.48*	1.67*	1.87*
	(0.14)	(0.13)	(0.15)
Obs.	125,486	125,486	90,409
R^2	0.25	0.21	0.38
FEs	*f, ct*	*f, ct*	*f, ct*
		IV	
ln(1+AI_{at})	2.67*	3.89*	4.93*
	(0.44)	(0.42)	(0.49)
Weak Instrument F (KP)	1,384	1,384	985
		Reduced Form	
Z_{at}	1.60*	2.33*	2.98*
	(0.26)	(0.25)	(0.29)

Notes: This table reports estimates of equation (15), but with the alternative dependent variables listed in the column headers. Column 1 repeats column 2 of table 1. Columns 2 and 3 repeat the specification in column 1, but with the dependent variables being all users (foreign plus domestic) and domestic users. The three panels are for OLS, IV, and reduced form. The first stage appears in column 2 of table 1. If an app is produced in a country which is not one of the 84 importing countries in our sample, we have no domestic user information. Standard errors are clustered at the app level. * indicates 1% significance.

4.3. Domestic versus Foreign Users

The table 1 results are about foreign users, but the theory applies to domestic users as well. An interesting question is whether the results for domestic and foreign users are the same. A number of papers have argued that innovation is more valuable to domestic users than to foreign users e.g., the literatures on product cycles (Vernon, 1966), directed technical change with international technological mismatch (Acemoglu and Zilibotti, 2001, Acemoglu, Gancia, and Zilibotti, 2015), and multinational production (Arkolakis, Ramondo, Rodríguez-Clare, and Yeaple, 2018). To assess these arguments, in table 3 we consider different dependent variables. Column 1 is again for foreign users, column 2 is for all users, and column 3 is for domestic users. The coefficient on $\ln(1 + AI_{at})$ is indeed larger for domestic users than foreign users. This supports these theories. It also suggests that there are heterogeneous responses to AI. If the coefficient on AI is largest for apps that are most likely to deploy AI in response to the cost factors captured by Z_{at},

Table 4: Magnitudes

	(1)	(2)	(3)
	IV: ln(*Foreign Users* $_{at}$)		
ln(1+AI_{at})	2.67*		
	(0.44)		
IHS(AI_{at})		2.82*	
		(0.47)	
ln(AI_{at})			3.78*
			(0.72)
Obs.	125,486	125,486	27,824
FEs	*f, ct*	*f, ct*	*f, ct*
Weak Instrument F (KP)	1,384	1,346	700

Notes: This table reports IV estimates of equation (15), but with alternative measures of AI deployment. Column 1 repeats column 2 of table 1. Columns 2 replaces $\ln(1+AI_{at})$ with the inverse hyperbolic sine of AI_{at}, that is, with the log of $AI_{at}+(AI_{at}^2+1)^{1/2}$. Column 3 uses the restricted sample with $AI_{at}>0$ and replaces $\ln(1+AI_{at})$ with $\ln(AI_{at})$. See online appendix table A2 for the OLS, first-stage, and reduced-form results. Standard errors are clustered at the app level. * indicates 1% significance.

then IV (LATE) will be bumped up relative to OLS. See Card (2001, eqn. 11).[25]

4.4. Coefficient Magnitudes

As is well known, the interpretation of magnitudes can be sensitive to functional form, especially when taking logs of a variable that can be zero i.e., AI_{at}. In table 4 we explore magnitudes for three specifications. Column 1 repeats our baseline IV results (table 1, column 2). Column 2 replaces $\ln(1+AI_{at})$ with the inverse hyperbolic sine of AI_{at}, scaled by its iqr. This makes no difference. Column 3 restricts the sample to observations having positive values of AI_{at} and replaces $\ln(1+AI_{at})$ with $\ln(AI_{at})$. This captures the *intensive margin* of AI deployment and has a larger coefficient. These conclusions based on IV also hold for OLS. See online appendix table A2.

4.5. Sensitivity

In the online appendix we report a large number of alternative specifications and find that these are all consistent with our baseline specification of $\beta^{IV}=2.67$ ($s.e.=0.44$). We conclude this discussion of our result 1 with a few more alternative specifications.

[25]The fact that AI has a bigger impact on domestic users than foreign users also leads one to think about 'gravity' factors that place a wedge between the domestic and foreign impacts. We explore this in section 5 where we show that restrictions on cross-border data flows degrade the benefits of AI to foreign users.

1. In constructing the instrument we defined app a's industry $c(a)$ using similar apps. We can alternatively define $c(a)$ as a's App Store category. This does not change our result ($\beta^{IV} = 2.44$, $s.e. = 0.38$). See online appendix table A3.
2. In constructing AI_{at} we used ρ_{ap}-weighted counts of patents. Alternatively, we can use ρ_{ap}-weighted counts of patent families, in which case $\beta^{IV} = 2.61$ ($s.e. = 0.43$). We can also use ρ_{ap}-weighted counts of patent citations, in which case we obtain the larger result $\beta^{IV} = 5.18$ ($s.e. = 0.88$). See online appendix table A4. The increased size likely reflects the heavy right skew of patent citations.
3. We have not included any other time-varying firm characteristics such as firm revenues or assets. Adding these does not change our results at all. See online appendix table A4.
4. We have not included any other app characteristics. Following Ershov (forthcoming) and Bian *et al.* (2023), we include app age, app price, app rating, a dummy for in-app purchases, a dummy for whether the app displays adds, and a dummy for whether the firm advertises the app. For the subsample with these data, our results are unchanged at $\beta^{IV} = 2.52$ (0.47). See online appendix table A5.
5. When we defined the instrument we included a dummy $D_{f(a),\tau}$ for whether the firm filed AI patents between 1992 and year $\tau = 2008$. We also consider alternative years $\tau = 2006, \ldots, 2011$ and find that β^{IV} varies very little across these, between 2.64 and 2.88. See online appendix table A6.[26]
6. When constructing $AI_{at} = \Sigma_{p \in \mathcal{P}_{at}} \rho_{ap}$, we set $\rho_{ap} = 0$ whenever ρ_{ap} was less than a threshold $\overline{\rho}$. In our baseline, we set $\overline{\rho} = 0.2$. We also consider $\overline{\rho} = 0.0$, 0.1, 0.2, 0.3, 0.4 and 0.5. (99% of our ρ_{ap} are in this range.) While magnitudes are sensitive to the choice of cutoff, the β^{IV} are always economically and statistically very large (β^{IV} between 1.85 ($s.e.$ = 0.22) and 4.63 ($s.e.$ = 0.79)). See online appendix table A7.
7. In our baseline we used all AI patents rather than just deep learning patents because many recommenders use ensemble methods that do not involve deep learning. When we restrict ourselves to deep learning patents, $\beta^{IV} = 2.29$ $(s.e. = 0.44)$, which is still very large. See online appendix table A8.
8. It would be disappointing if our results were driven entirely by blockbuster apps such as Facebook. Blockbuster apps account for about 1% of our sample observations. In online appendix table A9 we drop app-year observations with the largest values of the dependent variable. Dropping the top 1%, 5% and 10% of observations leads to IV results of 2.59 (0.43), 2.54 (0.42) and 2.35 (0.40), respectively. These are

[26] We stop at 2011 because 2012 is the year AI commercialization began. Also, throughout this paper, whether we cumulate patents starting in 1992 or 2000 makes no difference since AI patents were rare in our sample before 2000.

only slightly lower than our baseline result of 2.67 (0.44). Thus, our results are not driven by blockbuster apps.[27]

9. We introduced category-exporter-year fixed effects to control for the inherent scalability of apps. Scalability might still enter through the back door via the instrument. Recall that the instrument includes L_{it}^{A}, which is mildly correlated with country size. In appendix table A10, to purge size we divide L_{it}^{A} by country i's year-t capital stock and repeat our baseline table 1. Again, this makes no difference.

This review of alternative specifications demonstrates the robustness of our result 1.

5. Result 2: Restrictions on Cross-Border Data Flows

AI algorithms with limited data are typically of limited value. Data are needed to train models and personalize predictions, both of which improve app quality. This has led producers of mobile apps to harvest vast amounts of user data. The movement of these data across borders has set off two conflicting trends in the international regulation of digital commerce. The first is the explosion of trade agreements with digital chapters which tend to liberalize data flows. There are 72 such agreements (Nemoto and López González, 2021). In addition, the WTO is very close to an e-commerce plurilateral involving 89 countries that may include provisions promoting cross-border data flows and limiting the use of data localization rules (rules requiring domestic data to be stored domestically). In contrast, there has been a trend towards unilateral national restrictions on the cross-border outflow of citizen data, often out of concern for privacy and national security. The most famous of these is the EU's General Data Protection Regulation (GDPR). See also the EU's 2023 Digital Services Act. Far more restrictive is China's many new laws that include the Cybersecurity Law (2017), the Data Security Law (2021), the Personal Information Protection Law (2021), and Measures on the Standard Contract for the Cross-Border Transfer of Personal Information (2023). Many other countries have related laws e.g., India's Information Technology Act was used to ban over 200 of China's most popular mobile apps. Figure 2 above documented the rapid rise of domestic regulations that restrict cross-border data flows. Despite the conflicting trends towards liberalizing and restricting cross-border data flows, there has been no academic assessment of their impacts.

To assess how regulatory restrictions on cross-border data flows limit the effectiveness of AI, we use the new OECD Digital Services Trade Restrictiveness Index (DSTRI), which inventories regulatory barriers to digital trade. See Ferencz (2019). The OECD was kind enough to provide us with the subcomponents of the index that are most

[27]Note that this is intended as a destructive diagnostic. Stratifying on the dependent variables produces inconsistent estimates.

relevant for AI and mobile apps. These subcomponents fall into two groups. (1) *Data Transfer* is an index of measures restricting cross-border data flows or requiring data to be stored locally. (2) *Data Development* indexes other restrictions that affect trade in digitally enabled services in ways that reduce the number of users and hence the amount of local data produced and available for cross-border transfer. These include restrictions on downloading, streaming, and advertising as well as local performance requirements that impose large compliance fixed costs e.g., commercial local presence requirements.[28] The DSTRI covers 60 of the importers in our data. 2020 data for each country appear in online appendix table A11.

The DSTRI catalogues *discriminatory* regulations. Thus, Data Transfer does not include restrictions on data transfer between two firms within a country and Data Development does not include restrictions that apply equally to domestic and foreign firms. The Data Transfer and Data Development indexes range from zero (no restrictions) to one (very restrictive). Let $Data_{nt} \in [0,1]$ denote the DSTRI index for either Data Transfer or Data Development. Because DSTRI inventories discriminatory regulations, $Data_{nt}$ is relevant for an app produced in country i and used in country $n \neq i$.

We introduce regulations into the model as follows. For an app a produced and used in country n, we have as before that AI is fully effective and the number of domestic users is $\left(\frac{\alpha_{ct}(\delta_{ac})\,\delta_{ac}}{\mathcal{U}_{cnt}}\right)\mathcal{L}_{nt}$ where recall that $\mathcal{L}_{nt}$ is the number of consumers and the term in parentheses is the share of consumers choosing app a from among all apps in industry c available in n. For an app produced in i and used in country $n \neq i$, AI's effectiveness is assumed to be reduced by $\theta Data_{nt}$ so that the number of users is

$$Foreign\,Users_{acnt} = \frac{[\alpha_{ct}(\delta_{ac})]^{1-\theta Data_{nt}}\;\delta_{ac}}{\mathcal{U}_{cnt}}\mathcal{L}_{nt}\,.$$

The larger is $Data_{nt}$, the less access the app developer has to country-n data and so the less effective is AI for country-n users. As before, we measure $\ln(a_{ct}(\delta_{ac}))$ by $\beta\ln(1+AI_{at})$.[29] This leads to

$$\ln(Foreign\,Users_{ant}) = (1-\theta Data_{nt})\cdot\beta\ln(1+AI_{at}) + \ln(\mathcal{L}_{nt}\,/\,\mathcal{U}_{cnt}) + \delta_a$$

where we suppressed the industry subscript $c = c(a)$ whenever there is already an app subscript a. Our estimating equation is thus

$$\ln(Foreign\,Users_{ant}) = \beta\ln(1+AI_{at}) - \theta' Data_{nt}\cdot\ln(1+AI_{at}) + \lambda_{cnt} + \lambda_f + \varepsilon_{ant} \quad (16)$$

[28]Using the categories in Ferencz (2019), Data Transfer is the five subcomponents of 'Infrastructure and Connectivity' dealing with cross-border transfer of personal data and cross-border data flows. Data Development is all seven subcomponents of 'Other barriers affecting trade in digitally enabled services.

[29]With data restrictions, the firm's problem changes and so the optimal level of AI investment $\alpha_{ct}(\delta_{ac})$ changes. However, we do not need to revisit the model and recalculate this because whatever this new optimal level is, its log will still be measured by $\beta\ln(1+AI_{at})$.

where $\theta' = \theta\beta$, $\lambda_{cnt} = \mathcal{L}_{nt} / \mathcal{U}_{cnt}$ are fixed effects, we include firm fixed effects λ_f, and $\varepsilon_{ant} = \delta_a$ is a residual.[30]

We cluster standard errors at the app level and, reassuringly, we will show that our standard errors are similar to those in our at-level regressions.[31]

As before, we instrument $\ln(1 + AI_{at})$ with Z_{at}. We instrument $Data_{nt} \cdot \ln(1 + AI_{at})$ with $Data_{nt} \cdot Z_{at}$. Notice that the fixed effects λ_{cnt} absorb the data regulations $Data_{nt}$ so that, conditional on the fixed effects, the instrument $Data_{nt} \cdot Z_{at}$ is unlikely to be correlated with the residual ε_{ant}.

Table 5 presents our results. Unlike our previous regressions, which were at the app-year at level, the data in table 5 are at the app-importer-year ant level. The top two panels are OLS and IV. Column 1 is a basic specification without the DSTRI variables and shows that the OLS and IV coefficients on AI are largely unchanged from the at-level regressions above.

We consider the two types of data restrictions separately. In column 2 we interact $Data\,Transfer_{nt}$ with AI, in column 3 we interact $Data\,Development_{nt}$ with AI, and in column 4 we include both interactions. In these columns, the OLS and IV coefficients on all the interaction terms are negative, indicating that *data restrictions reduce the effectiveness of AI as a tool for increasing the number of foreign users. This is the second main result of this paper. We are the first to document this finding and are thus able to provide an important input into trade policy.*

Our finding is also important in a domestic setting. In the US Department of Justice case against Google, the central issue is whether Google's dominant position is the result of its algorithms or its data. The Department of Justice argues that Google uses its dominant position to harvest data that then reinforces its position. Google argues that its success is not due to its dominant position but to its algorithms.[32] Our results show that even superior algorithms are limited by the availability of data.

The third, fourth and fifth panels of table 5 report the first stages for $\ln(1 + AI_{at})$, $Datat\,Transfer_{nt} \cdot ln(1 + AI_{at})$ and $Data\,Development_{nt} \cdot ln(1 + AI_{at})$, respectively. One should always have major concerns about an IV model with more than one endogenous variable. We take a quantum of solace in the large weak-instruments F statistics.

[30] In $\varepsilon_{ant} = \delta_a$, the subscripts differ, but this is minor. It is trivial to allow mean utilities to vary by importer, which adds an n subscript to δ_a. It is equally trivial to allow mean utilities to evolve over time, which adds a t subscript to δ_a e.g., improvements in smartphones lead to higher mean utilities over time.

[31] Clustering at an produces much smaller standard errors. Two-way clustering by a and n produces almost identical standard errors to one-way clustering by a.

[32] e.g., *New York Times*, September 18, 2023, "A Key Question in Google's Trial: How Formidable Is Its Data Advantage?"

Table 5: Importer-Level (a,n,t) Regressions: Role of Data Restrictiveness

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS: ln(*Foreign Users* $_{ant}$)					
ln(1+ AI_{at})	1.34*	1.40*	1.45*	1.46*	1.52*	1.66*
	(0.14)	(0.14)	(0.14)	(0.14)	(0.15)	(0.15)
Data Transfer $_{nt}$ × ln(1+ AI_{at})		-1.03*		-0.19		-0.54*
		(0.13)		(0.13)		(0.18)
Data Development $_{nt}$ × ln(1+ AI_{at})			-2.53*	-2.43*		-2.28*
			(0.23)	(0.24)		(0.29)
Obs.	3,575,088	3,575,088	3,575,088	3,575,088	1,678,918	1,678,918
Number of importers	60	60	60	60	25	25
R^2	0.25	0.25	0.25	0.25	0.23	0.23
FEs	*f, cnt*	*f, cnt*	*f, cnt*	*f, cnt*	*f, cnt*	*f, cnt*
	IV: ln(*Foreign Users* $_{ant}$)					
ln(1+ AI_{at})	2.27*	2.42*	2.55*	2.58*	2.60*	3.01*
	(0.53)	(0.54)	(0.55)	(0.55)	(0.54)	(0.55)
Data Transfer $_{nt}$ × ln(1+ AI_{at})		-1.90*		-0.77*		-2.03*
		(0.18)		(0.17)		(0.23)
Data Development $_{nt}$ × ln(1+ AI_{at})			-3.73*	-3.34*		-2.74*
			(0.35)	(0.37)		(0.41)
Weak Instrument F (KP)	670	333	326	217	794	257
	First Stage: ln(1+AI_{at})					
Z_{at}	0.51*	0.50*	0.50*	0.50*	0.51*	0.50*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Data Transfer $_{nt}$ × Z_{at}		0.02*		-0.01*		0.01
		(0.00)		(0.00)		(0.00)
Data Development $_{nt}$ × Z_{at}			0.08*	0.09*		0.09*
			(0.01)	(0.01)		(0.01)
	First Stage: *Data Transfer* $_{nt}$ × ln(1+ AI_{at})					
Z_{at}		0.001		0.000		0.002
		(0.001)		(0.001)		(0.001)
Data Transfer $_{nt}$ × Z_{at}		0.497*		0.496*		0.485*
		(0.006)		(0.006)		(0.006)
Data Development $_{nt}$ × Z_{at}				0.004*		0.003*
				(0.001)		(0.001)
	First Stage: *Data Development* $_{nt}$ × ln(1+ AI_{at})					
Z_{at}			-0.001	-0.001		-0.000
			(0.001)	(0.001)		(0.001)
Data Transfer $_{nt}$ × Z_{at}				-0.003*		-0.008*
				(0.000)		(0.001)
Data Development $_{nt}$ × Z_{at}			0.505*	0.507*		0.504*
			(0.007)	(0.007)		(0.007)

Notes: This table reports estimates of equation (16). Each observation is an app a (35,575 apps) used by consumers in a country n (60 importers) in a year t (2015, . . . ,2020). The dependent variable is the log number of users of app a in country n. $Data\,Transfer_{nt}$ and $Data\,Development_{nt}$ are OECD indexes of digital service trade restrictiveness (higher values are more restrictive). The panels report OLS, IV, and three first-stages. All specification include firm fixed effects (f) and industry-importer-year fixed effects (cnt). Standard errors are clustered at the app level. * indicates 1% significance.

Further, the first-stage results are very sensible in that each instrument targets only its corresponding endogenous regressor. James Bond would be comforted.

A number of our 60 importing countries are small, leading one to wonder whether our results are driven mainly by small economies. To examine the effects for large economies we repeat the analysis for the 25 largest economies in our sample (listed below). Comparing columns 1 and 5, the IV coefficient rises slightly from 2.27 to 2.60, which means our results are not driven solely by small economies. Comparing columns 4 and 6, the coefficients on $\ln(1 + AI_{at})$ and $Data\,Transfer_{nt} \times \ln(1 + AI_{at})$ both grow, while the coefficient on $Data\,Development_{nt} \times \ln(1 + AI_{at})$ shrinks. The shrinkage reflects the fact that Data Development restrictions are less used by large economies. Summarizing, columns 5–6 establish that our results are not driven solely by small economies.

Turning to coefficient magnitudes, we first ask "By how much do data restrictions dampen the impact of AI on foreign users?" Specifically, what happens when a country shifts from having the highest observed level of restrictions to having the lowest observed level of restrictions. Since there are no extreme values of either Data Transfers or Data Development, this thought experiment is well within the sample variation we are using. Also, we restrict ourselves to sample variation among the 25 largest economies. In 2020, $Data\,Transfer_{n,2020}$ was 0.16 for the most restrictive country (Indonesia) and 0 for the least restrictive countries (e.g., Canada and the US). We therefore consider the change $\Delta Data\,Transfer_n = 0.16$. In 2020, $Data\,Development_{n,2020}$ was 0.13 for the most restrictive country (Egypt) and 0 for the least restrictive countries, including the US and all CPTPP partners, so we set $\Delta Data\,Development_n = 0.13$.[33]

From equation (16), AI's impact on foreign users depends on data restrictions via $\partial \ln(Foreign\,Users_{ant}) / \partial \ln(1 + AI_{at}) = \beta - \theta' Data_{nt}$. As restrictions rise from 0 to $\Delta Data_n$ this impact falls from β to $\beta - \theta' \Delta Data_n$ i.e., falls by $\theta' \Delta Data_n$. From column 4 of table 5, this fall is

$$-0.77 \Delta Data\,Transfer_n - 3.34 \cdot \Delta Data\,Development_n = -0.56\,. \tag{17}$$

A useful way to interpret this change is as follows. From the column 4 coefficient on $\ln(1 + AI_{at})$ of 2.58, without any data restrictions the impact of AI is 2.58 and with high restrictions it is $2.58 - 0.56 = 2.02$. Thus, data restrictions reduce the impact of AI from a 13.2-fold impact ($e^{2.58}$) to a 7.5-fold impact ($e^{2.02}$). *The impact of AI on foreign users is halved by data restrictions.* This is a huge effect and is one of our headline numbers. When we

[33] The Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP) has a digital chapter that deals with many of the restrictions in $Data\,Development_{nt}$. Data for all countries appear in online appendix table A11.

look at the 25 largest economies in our sample, the exact same conclusion emerges. The impact of AI on foreign users is halved.[34]

These results explain the heavy lobbying by major platform companies for trade policies that liberalize cross-border data flows.[35]

Table 6 drills down to the impacts for each of the top 25 countries in our sample. Columns 1–2 report $Data\,Transfer_{n,2020}$ and $Data\,Development_{n,2020}$. Columns 3–4 report the calculations in equation (17), but treating the data in columns 1–2 as $\Delta Data\,Transfers_n$ and $\Delta Data\,Development_n$. Equation (17) is implemented in two ways, using coefficients from our 60-country sample (column 4 of table 5) and using coefficients from our 25-country sample (column 6 of table 5). From columns 3–4 of table 6, there is a wide range of impacts of data restrictions on AI effectiveness.

A second thought experiment asks about the impact of data restrictions on users of an app that deploys an average amount of AI. From equation (16), this is $\Delta \ln(Foreign\,Users_{an}) = \overline{\ln(1+AI_{at})}\,(-\theta' \Delta Data_n)$. For the largest 25 countries (column 6 of table 5) this is

$$\Delta \ln(Foreign\,Users_{an}) = \overline{\ln(1+AI_{at})}\,(-2.03\Delta Data\,Trans_n - 2.74 \cdot \Delta Data\,Dev_n)$$

where $\overline{\ln(1+AI_{at})} = 1.41$ is the 2020 user-weighted average value of $\ln(1+AI_{at})$ among firms that deploy AI. The results appear in column 5 of table 6. For China, foreign users are reduced by 0.68 log points. The results are in line with the scant literature on the topic. Sun *et al.* (forthcoming) find experimental evidence that when the Alibaba platform does not use personal data for its recommendations, customer click-through rates drop by 75%, product views drop by 33%, and customer purchases drop by 81%. For EU countries, our impacts range from 0.20 to 0.31 (see column 5 of table 6). This is a little higher than found in Goldberg, Johnson, and Shriver (forthcoming) where, for a diverse set of online firms, GDPR reduced page views and e-commerce revenue by only 12%, though this rises to 20% for the display-ad and social-advertising channels. We find it very reassuring that our results are consistent with both experimental evidence and the GDPR shock.

Finally, column 6 reports the sum of the Polity V autocracy and democracy scores. Autocracies are in italics and have scores below -3. Democracies have scores above 6.

[34]For the 25 largest economies, the calculation is $-2.03 \cdot 0.16 - 2.74 \cdot 0.13 = -0.325 - 0.356 = -0.68$. Thus, as restrictions rise the impact of AI falls from 3.01 to $3.01 - 0.68 = 2.33$ so that data restrictions reduce the impact of AI from a 20.3-fold impact ($e^{3.01}$) to a 10.3-fold impact ($e^{2.33}$). Again, the impact is halved.

[35]See United States International Trade Commission (2019) in the context of USMCA and https://techcrunch.com/2022/01/19/google-lobbies-for-new-privacy-shield/ in the context of EU data regulations.

Table 6: The Impact of AI on Foreign Users: Role of Cross-Border Data Flow Restrictions

Country	OECD DSTRI Index, 2020		Reduction in AI Coefficient			
	Data Transfer $_n$	*Data Development* $_n$	60-Country Specification	25-Country Specification	Reduction in ln(*Foreign Users*)	PolityV Score
	(1)	(2)	(3)	(4)	(5)	(6)
Egypt	0.08	0.13	0.50	0.52	0.74	-4
Indonesia	0.16	0.07	0.34	0.50	0.71	9
China	0.12	0.09	0.39	0.48	0.68	-7
Nigeria	0.08	0.11	0.43	0.46	0.65	7
Pakistan	0.08	0.11	0.43	0.46	0.65	7
Russia	0.12	0.07	0.31	0.42	0.60	4
India	0.12	0.07	0.31	0.42	0.60	9
Turkey	0.08	0.07	0.28	0.34	0.48	-4
South Korea	0.08	0.04	0.21	0.28	0.40	8
Brazil	0.12	0.00	0.09	0.24	0.34	8
Germany	0.08	0.02	0.13	0.22	0.31	10
Belgium	0.08	0.02	0.13	0.22	0.31	8
Italy	0.04	0.04	0.18	0.20	0.28	10
Philippines	0.04	0.04	0.18	0.20	0.28	8
France	0.04	0.04	0.18	0.20	0.28	10
Netherlands	0.04	0.02	0.10	0.14	0.20	10
Argentina	0.04	0.02	0.10	0.14	0.20	9
Spain	0.04	0.02	0.10	0.14	0.20	10
Mexico	0.04	0.00	0.03	0.08	0.11	8
Japan	0.04	0.00	0.03	0.08	0.11	10
Australia	0.04	0.00	0.03	0.08	0.11	10
South Africa	0.04	0.00	0.03	0.08	0.11	9
UK	0.04	0.00	0.03	0.08	0.11	8
USA	0.00	0.00	0.00	0.00	0.00	8
Canada	0.00	0.00	0.00	0.00	0.00	10

Notes: Columns 1 and 2 report the data we use from custom runs of the OECD DSTRI. Columns 1 and 2 report our data on $DataTransfer_{n,2020}$ and $Data\,Development_{n,2020}$. Columns 3 and 4 report how the impact of AI on foreign users responds when a country eliminates its restrictions on cross-border data flows. It is computed as follows. $\partial \ln(ForeignUsers_{ant})/\partial \ln(1+AI_{at}) = \beta - \theta' Data_{nt}$ is the impact of AI on foreign users and $\Delta\partial \ln(ForeignUsers_{ant})/\partial \ln(1+AI_{at}) = -\theta' \Delta Data_n$ is how it responds to cross-border data restrictions. This is what appears in columns 3–4. In column 3 (4), θ' is taken from column 4 (6) of table 5. Column 5 is the impact of cross-border data restrictions on foreign users. Specifically, column 5 is column 4 times $\overline{\ln(1+AI_{at})} = 1.41$, the 2020 user-weighted average value of $\ln(1+AI_{at})$ among firms that deploy AI. Column 6 is the sum of the Polity V autocracy and democracy scores. Countries in italics have scores below -3. All other countries have scores above 6.

Autocracies have the strongest restrictions on cross-border data flows and thus suffer the most from these restrictions. Autocracy is costly.

6. External Sources of AI Knowledge

A remarkable feature of AI has been its rapid diffusion. There are many channels for diffusion, including journal articles, academic presentations, international competitions (ImageNet and Kaggle), repositories of algorithms (GitHub), and AI patents databases (WIPO and Google Patents). It is clear that many mobile app developers deploy AI that they have sourced through these channels. In this last section we explore the impact on mobile app quality of AI sourced from outside the firm. That is, we explore AI knowledge spillovers.

Our approach is novel. Suppose we had a rich corpus of text which described the current state of AI. This corpus is the AI *potentially* available to a firm that does not develop AI internally, but instead relies on external knowledge. For each app we could ask whether its description is cosine similar to this AI corpus. That is, we could develop a measure of the external AI potentially embodied in each app. We could then use this to assess the impact of external AI potential on an app's number of foreign users.

We implement this as follows. We treat the AI corpus at time t as the set of all AI patents filed from 1992 to t by our sample of mobile app producers. Denote this corpus by $\mathcal{C}_t$. This misses AI knowledge available from the other channels listed above, but patents certainly describe a significant piece of the larger AI corpus and, usefully, we have already assembled and used this element of the larger corpus. Let $\mathcal{A}_t$ be the set of apps developed by firms with *no* AI patent filings from 1992 to t. These firms must rely on external AI. For each app a in $\mathcal{A}_t$, we define its potential to draw on the external AI corpus $\mathcal{C}_t$ as

$$external\ AI_{at} = \begin{cases} \sum_{p \in \mathcal{C}_t} \rho_{ap} & \text{if } a \in \mathcal{A}_t \\ 0 & \text{if } a \notin \mathcal{A}_t \end{cases}. \tag{18}$$

This is the ρ_{ap}-weighted sum of all external-to-the-firm patents in our sample. As usual, we scale this by its interquartile range. The right panel of figure 5 above is a bin scatter from a regression of the log of foreign users on $\ln(external\ AI_{at})$ with firm and industry-year fixed effects. The panel shows that $external\ AI_{at}$ is correlated with the number of foreign users.

In what follows we include $external\ AI_{at}$ as a regressor in our foreign-user regressions. However, we first address the issue of endogeneity. $external\ AI_{at}$ is built up from three items: (1) the patents that are being summed; (2) the text of these patents; and, (3) the text of the app description. The first two items involve patents that were *not*

filed by the developer of app a and so are orthogonal to a. The third item, the app description, is potentially endogenous because the product characteristics it describes are potentially endogenous. The most salient source of endogenous product characteristics is an unobservable firm characteristic such as management that is correlated both with the firm's ability to use external AI and with its ability to create high-demand products. Our inclusion of firm fixed effects controls for this source of endogeneity. We therefore do not instrument $external\,AI_{at}$.[36]

Table 7 reports our results. Consider the OLS results in the top panel. The panel has the same format as table 1, but with three differences. First, we include $\ln(1 + external\,AI_{at})$ as a regressor. Second, the sample is smaller because we only include at observations involving apps developed by firms with no AI patent filings as of year t. Third, for this sample, firms have no patents as of year t so that $AI_{at} = 0$ and $undirected\,AI_{at} = 0$. Hence, we do not include these variables. In column 2 of the top panel, which has our baseline firm and industry-year fixed effects, the coefficient on $\ln(1 + external\,AI_{at})$ is 0.33 ($s.e.$ = 0.03). There are thus statistically and economically significant AI externalities. A one iqr increase in $ln(1 + external\,AI_{at})$ leads to a 0.33 log point increase in the number of foreign users. This conclusion is invariant to fixed effects (columns 1–4) and to the inclusion of $\ln(1 + non\,AI_{ft})$.

The second and third panels use the full sample so that we can include AI_{at} and $undirected\,AI_{at}$. The second panel is OLS and the third panel is IV with $\ln(1 + AI_{at})$ instrumented as before with Z_{at}. The coefficient on $\ln(1 + externality\,AI_{at})$ is stable across panels, meaning that our finding of externalities is robust.[37]

There are two economically interesting explanations for why the IV coefficient for external AI is smaller than for AI (in column 2, 0.30 versus 2.25). First, the AI corpus may be less relevant for an app than is a patent developed in-house. Second, the translation of knowledge potential into practice may occur with some probability $Prob$, in which case the benefit of external AI is $0.30 = Prob \times 2.25$. This allows us to back out the probability using $Prob = 0.30/2.25 = 0.13$.

Our robust finding of knowledge spillovers across mobile app developers – spillovers from developers that do patentable AI research to those who do not – is an important conclusion for the vast literature on knowledge spillovers e.g., Grossman and Helpman

[36]Demand estimation with endogenous product characteristics is not part of the trade literature and rarely a part of the industrial organization literature e.g., Berry and Haile (2021, page 17). Restated, the literature typically does not control for the endogeneity of product characteristics.

[37]The OLS and IV estimates of the coefficient on $\ln(1 + AI_{at})$ are slightly smaller than in our table 1 baseline. This is to be expected because the sample variation driving the coefficient relies in part on the contrast between the AI of patenting and non-patenting firms. This contrast is larger when non-patenters are treated as having no AI (table 1, $\beta^{IV} = 2.67$) as opposed to having external AI (table 7, $\beta^{IV} = 2.25$).

Table 7: AI Externalities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS: ln(*Foreign Users* $_{at}$)						
ln(1+*external AI* $_{at}$)	0.33*	0.33*	0.33*	0.32*	0.33*	0.33*	0.33*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
ln(1+*non AI* $_{ft}$)						0.12	0.12
						(0.38)	(0.38)
Obs.	95,301	95,301	95,283	94,836	95,301	95,301	95,301
R^2	0.26	0.26	0.27	0.28	0.26	0.26	0.26
FEs	*f, c, t*	*f, ct*	*f, ct, it*	*f, cit*	*f, ct*	*f, ct*	*f, ct*
	OLS: ln(*Foreign Users* $_{at}$)						
ln(1+*external AI* $_{at}$)	0.34*	0.34*	0.34*	0.34*	0.32*	0.34*	0.32*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
ln(1+*AI* $_{at}$)	1.09*	0.98*	1.13*	1.12*	0.83*	1.02*	0.84*
	(0.13)	(0.14)	(0.15)	(0.15)	(0.13)	(0.15)	(0.14)
ln(1+*undirected AI* $_{at}$)					-0.39*		-0.39*
					(0.11)		(0.11)
ln(1+*non AI* $_{ft}$)						-0.36	-0.02
						(0.34)	(0.35)
Obs.	125,486	125,486	125,467	125,024	125,486	125,486	125,486
R^2	0.25	0.25	0.26	0.27	0.25	0.25	0.25
	IV: ln(*Foreign Users* $_{at}$)						
ln(1+*external AI* $_{at}$)	0.29*	0.30*	0.30*	0.29*	0.29*	0.29*	0.29*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
ln(1+*AI* $_{at}$)	2.39*	2.25*	2.34*	2.62*	2.15*	2.37*	2.28*
	(0.43)	(0.46)	(0.52)	(0.55)	(0.49)	(0.50)	(0.54)
ln(1+*undirected AI* $_{at}$)					-0.20		-0.15
					(0.13)		(0.14)
ln(1+*non AI* $_{ft}$)						-1.68*	-1.54
						(0.58)	(0.64)
Weak Instrument F (KP)	1,762	1,519	1,295	1,185	1,494	1,371	1,333

Notes: Each observation is an app-year pair. The new regressor is $\ln(1 + external\ AI_{at})$ defined in equation (18). The top panel is for the subsample of at observations involving apps in $\mathcal{A}_t$ (roughly, the set of apps developed by firms with no AI patents). The bottom two panels (OLS and IV) are for the full sample and are identical in structure to table 1 except for the inclusion of $\ln(1 + external\ AI_{at})$. To understand the fixed effects, consider an app a in industry c developed by firm f located in exporting country i. Columns 1–4 contain fixed effects for (f,c,t), (f,ct), (f,ct,it), and (f,cit), respectively. In the top panel, columns 4–5 are the same and columns 6–7 are the same. This is because the firms in the top-panel sample have no patents and so have $undirected\ AI_{at} = 0$. (They also have $AI_{at} = 0$.) The first-stage and reduced-form estimates appear in online appendix table A12. Standard errors are clustered at the app level. * indicates 1% significance.

(1991). To our knowledge, this is among the tightest pieces of evidence on spillovers in the international trade literature. Other tight evidence uses patent citation data e.g., Aghion *et al.* (2021). Unlike research based on patent citations, we are drawing bilateral connections involving *firms that have no patents, which is the overwhelming majority of firms.* Our use of an LLM allows us to use a new data source, one that is available even for firms that do not patent, and thus provides a new and widely applicable methodology.

7. Conclusion

Digital service trade is large, rapidly growing, and understudied. The most dynamic element of this trade is mobile apps. Mobile apps did not exist two decades ago, yet they now dominate the lives of many and are valued by consumers at $2.5 trillion (Brynjolfsson *et al.*, 2023). We built a sample of 35,575 apps used in 84 countries and exported from 64 countries over 2015–2020. We showed that foreign users as a share of total users is far higher for mobile apps than for manufactured goods or web-based internet services such as e-commerce.

Many mobile apps rely heavily on AI algorithms and data. We developed a novel method of linking mobile apps to AI by using a large language model to construct a measure of the degree to which a firm's mobile app deploys the AI described in the firm's patent portfolio. This linking of patents to products solved a long-standing problem in the innovation literature dating back at least to Kortum and Putnam (1997). With this tool in hand and using a theory-consistent estimating equation and instrument we showed that:

1. AI causally increases international trade in mobile app services by 2.67 log points or by more than 10-fold.

The value of deploying AI depends critically on the availability of data. The appearance of mobile apps in late 2008 sparked an explosion of conflicting regulations, laws, and trade agreements governing cross-border data flows. These have first-order implications for international trade, as well as for privacy, national security, and foreign interference in elections. Yet academic research by international trade economists has been absent. We showed that:

2. AI's causal impact on mobile app trade is halved by restrictions on cross-border data flows. Further, autocracies have the tightest restrictions on the export of data and this reduces these countries' use of foreign apps by between 50% and 75%.

Finally, we provided a novel method of estimating externalities. A common method exploits patent citations and so is restricted to firms with patents. Most firms have no patents. We used a large language model to estimate an app's potential to use the corpus

of AI patents developed by other firms. We did this for apps whose developers have no patents. We showed that:

3. The greater is an app's potential to use the AI corpus developed by other firms, the more foreign users the app has.

That is, there are knowledge spillovers.

In future research we will explore how our method can be used to study the international diffusion of AI technologies and to inform discussions about international policy coordination for AI algorithms and data.

References

Acemoglu, Daron, Gino Gancia, and Fabrizio Zilibotti. 2015. Offshoring and directed technical change. *American Economic Journal: Macroeconomics* 7(3):84–122.

Acemoglu, Daron and Fabrizio Zilibotti. 2001. Productivity differences. *Quarterly Journal of Economics* 116(2):563–606.

Aghion, Philippe, Antonin Bergeaud, Timothee Gigout, Matthieu Lequien, and Marc Melitz. 2021. Exporting ideas: Knowledge flows from expanding trade in goods. mimeo, Collège de France.

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

Alaveras, Georgios and Bertin Martens. 2015. International trade in online services. JRC working papers on digital economy, European Commission Joint Research Centre.

Analysis Group. 2023. The continued growth and resilience of Apple's App Store ecosystem. URL https://www.apple.com/newsroom/pdfs/the-continued-growth-and-resilience-of-apples-app-store-ecosystem.pdf.

Aral, Sinan. 2021. *The Hype Machine: How Social Media Disrupts Our Elections, Our Economy, and Our Health – and How We Must Adapt*. NY: Crown Publishing Group.

Argente, David, Salomé Baslandze, Douglas Hanley, and Sara Moreira. 2023. Patents to Products: Product Innovation and Firm Dynamics. Working paper, Federal Reserve Bank of Atlanta.

Arkolakis, Costas. 2010. Market penetration costs and the new consumers margin in international trade. *Journal of Political Economy* 118(6):1151–1199.

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple. 2018. Innovation and production in the global economy. *American Economic Review* 108(8):2128–2173.

Bailey, Michael, Abhinav Gupta, Sebastian Hillenbrand, Theresa Kuchler, Robert Richmond, and Johannes Stroebel. 2021. International trade and social connectedness. *Journal of International Economics* 129:103418.

Bar-Isaac, Heski, Guillermo Caruana, and Vicente Cuñat. 2012. Search, design, and market structure. *American Economic Revierw* 102(2):1140–60.

Beraja, Martin, Andrew Kao, David Y. Yang, and Noam Yuchtman. 2023. Exporting the surveillance state via trade in AI. Brookings working papers, Brookings Institution.

Beraja, Martin, David Y. Yang, and Noam Yuchtman. forthcoming. Data-intensive innovation and the state: Evidence from AI firms in China. *Review of Economic Studies* .

Berry, Steven T. and Philip A. Haile. 2021. Foundations of demand estimation. In Kate Ho, Ali Hortacsu, and Alessandro Lizzeri (eds.) *Handbook of Industrial Organization, Volume 4*. Elsevier, 1–62.

Bian, Bo, Xinchen Ma, and Huan Tang. 2023. The supply and demand for data privacy: Evidence from mobile apps. Working Paper, University of British Columbia.

Blum, Bernardo S. and Avi Goldfarb. 2006. Does the internet defy the law of gravity? *Journal of International Economics* 70(2):384–405.

Bradford, Anu. 2023. *Digital Empires: The Global Battle to Regulate Technology*. New York: Oxford University Press.

Brynjolfsson, Erik, Avinash Collis, Asad Liaqat, Daley Kutzman, Haritz Garro, Daniel Deisenroth, Nils Wernerfelt, and Jae Joon Lee. 2023. The digital welfare of nations: New measures of welfare gains and inequality. Working Paper No. 31670, National Bureau of Economic Research,.

Brynjolfsson, Erik, Xiang Hui, and Meng Liu. 2019. Does machine translation affect international trade? Evidence from a large digital platform. *Management Science* 65(12):5449–5460.

Carballo, Jerónimo, Marisol Rodriguez Chatruc, Catalina Salas Santa, and Christian Volpe Martincus. 2022. Online business platforms and international trade. *Journal of International Economics* 137:103599.

Card, David. 2001. Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica* 69(5):1127–1160.

Casalini, Francesca and Javier López González. 2019. Trade and cross-border data flows. OECD Trade Policy Papers 220, OECD.

Chen, Maggie X. and Min Wu. 2021. The value of reputation in trade: Evidence from Alibaba. *Review of Economics and Statistics* 103(5):857–873.

Coe, David T. and Elhanan Helpman. 1997. International r&D spillovers. *European Economic Review* 39(5):859–887.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* .

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz. 2004. Dissecting trade: Firms, industries and export destinations. *American Economic Review Papers and Proceedings* 94(2):150–154.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz. 2011. An anatomy of international trade: Evidence from french firms. *Econometrica* 79(5):1453–1498.

Ershov, Daniel. forthcoming. Variety-based congestion in online markets: Evidence from mobile apps. *American Economic Journal: Microeconomics* .

Fajgelbaum, Pablo, Gene M. Grossman, and Elhanan Helpman. 2011. Income distribution, product quality, and international trade. *Journal of Political Economy* 119(4):721–765.

Feenstra, RC, R Inklaar, and MP Timmer. 2015. The next generation of the Penn World Table. *American Economic Review* 105(10):3150–3182.

Ferencz, Janos. 2019. The OECD digital services trade restrictiveness index. OECD Trade Policy Paper no. 271, OECD.

Ferencz, Janos, Javier López González, and Irene Oliván García. 2022. Artificial intelligence and international trade. OECD Trade Policy Papers 260, OECD.

Financial Times. 2023. Transformers: The google scientists who pioneered an AI revolution. *Financial Times* URL https://www.ft.com/content/37bb01af-ee46-4483-982f-ef3921436a50.

Freund, Caroline and Diana Weinhold. 2002. The Internet and International Trade in Services. *American Economic Review Papers & Proceedings* 92(2):236–240.

Freund, Caroline L and Diana Weinhold. 2004. The effect of the Internet on international trade. *Journal of International Economics* 62(1):171–189.

Goldberg, Samuel G., Garrett A. Johnson, and Scott K. Shriver. forthcoming. Regulating privacy online: An economic evaluation of the GDPR. *American Economic Journal: Economic Policy* .

Goldfarb, Avi, Bledi Taska, and Florenta Teodoridis. 2023. Could machine learning be a general purpose technology? A comparison of emerging technologies using data from online job postings. *Research Policy* 52(1):104653.

Goldfarb, Avi and Daniel Trefler. 2019. Artificial intelligence and international trade. In Ajay Agrawal, Joshua Gans, and Avi Goldfarb (eds.) *The Economics of Artificial Intelligence: An Agenda*, NBER. University of Chicago Press, 463–492.

Goldfarb, Avi and Catherine Tucker. 2019. Digital economics. *Journal of Economic Perspectives* 57(1):3–43.

Grossman, Gene M. and Elhanan Helpman. 1991. *Innovation and Growth in the Global Economy*. Cambridge, Mass. and London: MIT Press.

Herman, Peter R. and Sarah Oliver. 2022. Trade, policy, and economic development in the digital economy. Working Paper No. 2202-03-D, U.S. International Trade Commission.

Hoberg, Gerard and Gordon Phillips. 2016. Text-based network industries and endogenous product differentiation. *Journal of Political Economy* 124(5):1423–1465.

Hochberg, Yael V., Carlos J. Serrano, and Rosemarie H. Ziedonis. 2018. Patent collateral, investor commitment, and the market for venture lending. *Journal of Financial Economics* 130(1):74–94.

Johnson, Garrett. 2022. Economic research on privacy regulation: Lessons from the GDPR and beyond. Working Paper No. 30705, National Bureau of Economic Research.

Kortum, Samuel and Jonathan Putnam. 1997. Assigning patents to industries: Tests of the yale technology concordance. *Economic Systems Research* 9(2):161–176.

Krugman, Paul R. 1980. Scale economies, product differentiation, and the pattern of trade. *American Economic Review* 70(5):950–959.

Lee, Dokyun and Kartik Hosanaga. 2019. How do recommender systems affect sales diversity? A cross-category investigation via randomized field experiment. *Information Systems Research* 30(1):239–259.

Lendle, Andreas, Marcelo Olarreaga, Simon Schropp, and Pierre-Louis Vézina. 2016. There goes gravity: eBay and the death of distance. *Economic Journal* 126(591):406–441.

Lileeva, Alla and Daniel Trefler. 2010. Improved access to foreign markets raises plant-level productivity ... for some plants. *Quarterly Journal of Economics* 125(3):1051–1100.

López González, Javier, Silvia Sorescu, and Pinar Kaynak. 2023. Of bytes and trade: Quantifying the impact of digitalisation on trade. OECD Trade Policy Paper no. 273, OECD.

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green. 1995. *Microeconomic Theory*. Oxford: Oxford University Press.

Maslej, Nestor, Loredana Fattorini, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Helen Ngo, Juan Carlos Niebles, Vanessa Parli, Yoav Shoham, Russell Wald, Jack Clark, and Raymond Perrault. 2023. The AI index 2023 annual report. Working paper, Institute for Human-Centered AI, Stanford University.

Melitz, Marc and Stephen J. Redding. 2023. Trade and innovation. In Ufuk Akcigit and John Van Reenen (eds.) *The Economics of Creative Destruction: New Research on Themes from Aghion and Howitt,*. Cambridge MA: Harvard University Press, 133–174.

Nemoto, Taku and Javier López González. 2021. Digital trade inventory: Rules, standards and principles. OECD Trade Policy Papers 251, OECD.

OECD. 2023. Of bytes and trade: Quantifying the impact of digitalisation on trade. OECD Trade Policy Paper no. 273, OECD.

O'Shaughnessy, Matt. 2023. Lessons from the world's two experiments in AI governance. URL https://carnegieendowment.org/2023/02/14/lessons-from-world-s-two-experiments-in-ai-governance-pub-89035.

Pellegrino, Bruno. 2023. Product differentiation and oligopoly: A network approach. Working paper, University of Maryland.

Reimers, Nils and Iryna Gurevych. 2019. Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. URL http://arxiv.org/abs/1908.10084.

Rosen, Sherwin. 1981. The economics of superstars. *American Economic Review* 71(5):843–858.

Schmookler, Jacob. 1954. The level of inventive activity. *Review of Economics and Statistics* 36(2):183–190.

Sensor Tower. 2021. Global consumer spending in mobile apps reached $133 billion in 2021, up nearly 20% from 2020. URL https://sensortower.com/blog/app-revenue-and-downloads-2021.

Stapleton, Katherine and Michael Webb. 2023. Automation, trade and multinational activity: Micro evidence from Spain. Mimeo.

Sun, Ruiqi and Daniel Trefler. 2022. AI, trade and creative destruction: A first look. In Lili Yan Ing and Gene Grossman (eds.) *Robots and AI: A New Economic Era*. Routledge, 310–345.

Sun, Tianshu, Zhe Yuan, Chunxiao Li, Kaifu Zhang, and Jun Xu. forthcoming. The value of personal data in internet commerce: A high- stakes field experiment on data regulation policy. *Management Science* .

United States International Trade Commission. 2019. U.S.-Mexico-Canada Trade Agreement: Likely impact on the U.S. economy and on specific industry sectors. Working Paper No. 4889, US Government.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (eds.) *Advances in Neural Information Processing Systems*, volume 30.

Vernon, Raymond. 1966. International investment and international trade in product cycle. *Quarterly Journal of Economics* 83(2):324–329.

Webb, Michael. 2020. The impact of artificial intelligence on the labor market. Working Paper, Stanford University.

WIPO. 2018. Data collection method and clustering scheme. Background paper, Geneva: World Intellectual Property Organization.

WIPO. 2019. *WIPO Technology Trends 2019: Artificial Intelligence*. Geneva: World Intellectual Property Organization.

Xu, Mengwei, Jiawei Liu, Yuanqiang Liu, Felix Xiaozhu Lin, Yunxin Liu, and Xuanzhe Liu. 2021. A first look at deep learning apps on smartphones.

Yan, Lili and Gene Grossman. 2022. *Robots and AI: A New Economic Era*. UK: Routledge.

Appendix A. Details of Mobile App Data

The Google Play Store is not available in China. For Chinese user data we therefore use apps available on Apple's App Store and scale up each app's users by $(G_t + A_t)/A_t$ where G_t and A_t are, respectively, the number of Android and iOS smartphones in China in year t. Note that while we are imputing Chinese users (demand side), we are not imputing the foreign users of Chinese apps. As a result, our imputation turns out to be innocuous: All our results are unchanged when we (1) delete China as a user of apps or (2) add industry-country-year fixed effects. The latter absorb $(G_t + A_t)/A_t$ and would absorb any industry-year scaling $(G_{ct} + A_{ct})/A_{ct}$.

Appendix B. BERT

The large language model we use is called BERT, which stands for Bidirectional Encoder Representations from Transformers. We use the `paraphrase-multilingual-mpnet-base-v2` (Reimers and Gurevych, 2019) variant of BERT. This variant is multilingual, which we need because app descriptions are in many different languages. Each embedding is a vector of 768 elements, each between -1 and 1 and with Euclidean length of 1. 768 is a tuning parameter chosen by Google. Because BERT is a transformer, it is trained on sentences rather than words. As is common in industrial applications, rather than compute one embedding for the entire text (app description or patent title+abstract), we compute embeddings for each sentence of the text.[38]

Consider a patent p whose text has N_p sentences and an app a whose description has N_a sentences. Having computed embeddings for each sentence, we then calculate the $N_p \times N_a$ cosine similarities between each pair of app-patent embeddings. We then take the average of the top 5 similarities across sentence pairs, which gives us ρ_{ap}. This is what we use in the main text. In online appendix table A13 we also average across the n largest cosine similarities where $n = 1, 5, 10$. (Texts on average have 10.4 sentences.) The table shows that for $n = 1, 5, 10$, respectively, the OLS results are 1.28 (0.13), 1.48 (0.14), and 1.30 (0.14) and the IV results are 1.98 (0.36), 2.67 (0.44), and 3.55 (0.60). All these results imply very large impacts of AI on foreign users. We do not explore this further because, at the sentence-level, every specification involves the manipulation of the trillions of sentence-level cosine similarities and so is computationally intensive.

[38] It is easier to do just one embeddings for the entire text. However, this leads to relatively little variation in the ρ_{ap}, a well-known defect of LLMs. Data scientists at a major platform company told us that, as a result, they work at the sentence level. An additional oft-ignored problem when computing a single embedding for a lengthy text is token limits. LLMs truncate inputs when the number of tokens in the inputted text hits a maximum value. Truncation obviously compromises the quality of the embedding.

When feeding patent texts into the LLM, we feed in the title plus abstract. We use patents translated into English by Google and downloaded from patents.google.com. Other patent text such as claims and object of invention are often missing for non-US patents so we do not use these.

Appendix C. Proof of Equation (10) and Theorem 1

From (6)–(7), $\pi_i^A = \left[(\delta/\delta_i^A)^\eta - 1\right] w_i^A f_i^A$. Hence, from the Pareto distribution, the left side of (8) equals $(\delta_i^A)^{-\gamma} w_i^A f_i^A \eta/(\gamma - \eta)$. Plugging this into (8) and using (3) yields the expression for δ_i^A in (10). Plugging this δ_i^A into (7) and using (3) yields the expression for w_i^A/P in (10). In (9), the term in braces equals $(L_i^A)^\psi f_{ei}\gamma$. Plugging this into (9) and using (3) yields the expression for M_i^A in (10).

The proof of theorem 1 is trivial. Equation (11) follows from the discussion preceding either theorem 1 or equation (4). Equation (12) follows from plugging (10) into (5).

Finally, we show that $\eta f_i^A (L_i^A)^\psi > 1$ is a necessary and sufficient condition for an interior solution to the choice of α, meaning $\alpha(\cdot) > 1$ for all $\delta \geq \delta_i^A$. From the expression for $\alpha(\delta)$ in theorem 1, $\alpha(\cdot) > 1$ for all $\delta \geq \delta_i^A$ iff $\alpha(\delta_i^A) > 1$. From the expression for $\alpha(\delta)$ in theorem 1 and the definition of κ after equation (10), $\alpha(\delta_i^A) > 1$ iff $\delta_i^A (L_i^A)^{\psi/\eta} / \left[\delta_i^A (\eta f_i)^{-1/\eta}\right] > 1$ or $(L_i^A)^\psi \eta f_i > 1$ as required.

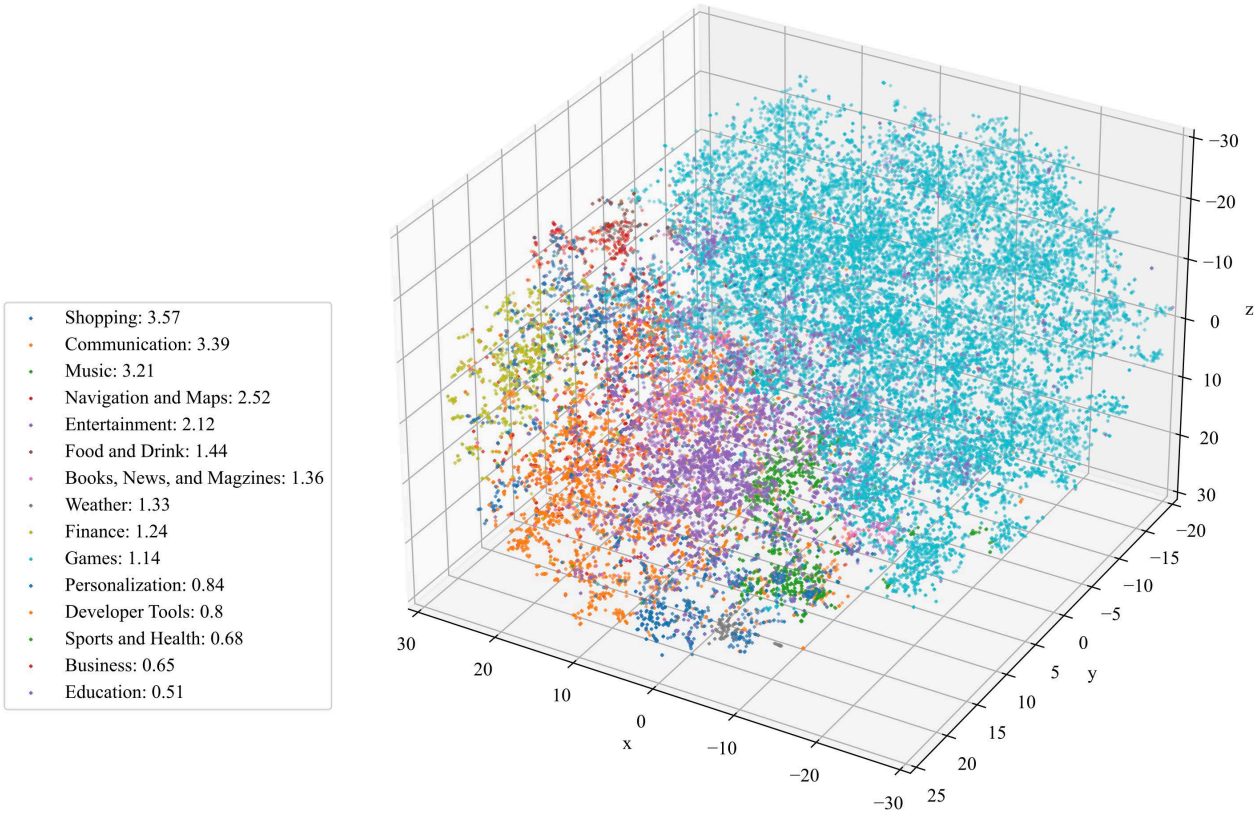
Online Appendix to

"The Impact of AI and Cross-Border Data Regulation on International Trade in Digital Services"

by

Ruiqi Sun and Daniel Trefler

Figure A1: $\rho_{aa'}$ Validation: Predicting App Categories



Appendix A. Validation of our use of an LLM

The details of the subsection 2.5 validation exercise involving $\rho_{aa'}$ are as follows. There are 35,575 apps and hence more than a half-billion app pairs. To reduce the dimensionality of the problem while keeping the most important app pairs, for each App Store category we select the 10 largest apps as measured by average users over 2015–2020. There are 15 app categories and so there are 150 focal apps, which we index by a. For each a, we calculate its cosine similarity with the remaining 35,574 apps, which we index by a'. There are $150 \times 35{,}574 = 5{,}336{,}100$ cosine similarities $\rho_{aa'}$. We assign these to 15 clusters using either k-means or agglomerative clustering (they yield identical results) and generate a binary variable $\hat{d}_{aa'}$ that is one if a and a' are in the same cluster. We compare this to a binary variable $d_{aa'}$ that is one if a and a' are in the same Apple Store category. For the 5,336,100 app pairs, the two binary variables agree 88.1% of the time (standard error is 0.00044).

Figure A1 provides visual verification of this result using a plot that allows the reader to examine all app pairs without the need for clustering. Each embedding is a 768-element vector. We reduce its dimension to a 3-element vector by extracting the largest 3 principal components from the $768 \times 35{,}575$ matrix of embeddings and then by plotting each app embedding as a point in a three dimensional space. See figure A1. Points (apps) that are close together should be in the same App Store category while those far apart should not be. In figure A1, the 15 categories are indicated by colours. Visual inspection

confirms that colours are clustered, meaning that apps with similar embeddings are located close to each other. This validates the performance of our large language model.[39]

Appendix B. The General Equilibrium Model with Advertising

We introduce advertising in the simplest way possible, namely, using a Krugman (1980) model with no trade costs and an Arkolakis (2010) advertising technology. Specifically, we introduce a differentiated consumer good that uses advertising to promote sales. It is produced with unskilled labour having wage $\tilde{w}_i$. Each country is endowed with the same amount of unskilled labour, denoted $\tilde{L}$. The population (and customer base) is $\mathcal{L}_n = L_n^A + \tilde{L}$ where recall that L_n^A is the endowment of AI scientists.

We micro-found advertising almost exactly as in Arkolakis (2010). Let S be the number of ads placed by a firm. Let $n_n(D)$ be the probability that a particular consumer in country n sees the ad at least once after D ads have been sent, where $n_n(0) = 0$. We assume that each ad placed targets a unique consumer, which is motivated by the fact that what an app user sees on her app is placed there either by an auction or an algorithm.[40] We also assume that within a given market, the cost per consumer differs depending on how many consumers have already been reached. In particular, the probability that a new ad is seen by a consumer for the first time is assumed to be $\beta n_n(D)^{(\beta-1)/\beta}$ where $\beta \in (0,1)$. That is, the probability that a new ad is seen for the first time is decreasing in the probability that the consumer has seen the ad.[41]

The marginal change in the number of consumers reached through new ads is $n_n'(D)\mathcal{L}_n$. Under our assumptions, $n_n'(D)\mathcal{L}_n = \beta n_n(D)^{(\beta-1)/\beta}$. Solving this differential equation subject to the initial condition $n(0) = 0$ implies

$$n_n(D) = (D/\mathcal{L}_n)^{\beta}. \tag{A1}$$

Turning to the consumer-goods firm's problem, one unit of the consumer good requires one unit of unskilled labour and so costs $\tilde{w}_i$. Note that, as in Krugman (1980), there is no firm productivity heterogeneity. There is CES demand for consumer goods with elasticity of substitution $\sigma > 1$. Consumer goods are costlessly traded internationally and there are no fixed costs of exporting so that each variety produced in country i is available worldwide at cost $\tilde{w}_i$. Consider a variety ω produced in country i with price $\tilde{p}_i(\omega)$ and quantity demanded in country n of

$$\tilde{q}_{ni}(\omega) = \left(\frac{D_{ni}(\omega)}{\mathcal{L}_n}\right)^{\beta} \frac{\tilde{p}^{-\sigma}(\omega)}{\tilde{P}^{1-\sigma}}\, y_n\, \mathcal{L}_n \tag{A2}$$

where $(D/\mathcal{L}_n)^{\beta}$ is the fraction of consumers that buy the good when the advertising level is D, $\tilde{P} = [\sum_i \int_{\omega\in\Omega_i} \tilde{p}_i^{1-\sigma} d\omega]^{1/(1-\sigma)}$ is the CES price index, Ω_i is the set of varieties

[39] An alternative approach uses UMAP. We did not try this.

[40] In terms of Arkolakis (2010), we are assuming that his scale parameter α equals unity. This is completely unimportant for any of our results.

[41] Arkolakis instead assumes the probability is $(1 - n_n(D))^{\beta}$ where $\beta > 0$. This has some better properties than our assumption, especially its implication that $n_n'(0) > 0$ which is crucial for Arkolakis's point about modelling small amounts of exports. Since we assume that there are no fixed costs of exporting, small amounts of exports will occur in our model even without advertising i.e., we do not need and do not have $n_n'(0) > 0$. Our assumption leads to simpler closed-form expressions.

produced in i, and y_n is per capita income available for spending on consumer goods, $y_n = [w_n^A L_n^A + \tilde{w}_n \tilde{L}]/\mathcal{L}_n$. Profits do not appear in per capita income because, with free entry, profits are used to pay for entry costs.

The optimal price is $\tilde{p}_i(\omega) = [\sigma/(\sigma-1)]\tilde{w}_i$. Firm profits are thus

$$\tilde{\pi}_i(\omega) = \sum_n \left\{ \frac{\tilde{w}_i}{\sigma - 1} \tilde{q}_{ni}(\omega) - p_n D_{ni}(\omega) \right\} - \tilde{w}_i \tilde{f}$$

where $\tilde{f}$ is the sunk cost of operating and recall that p_n was defined in the discussion preceding equation (4) as the per unit price of advertising in market n. Maximizing with respect to $D_{ni}(\omega)$ and using (A2) yields

$$p_n D_{ni}(\omega) = \frac{\beta \tilde{w}_i}{\sigma - 1} \tilde{q}_{ni}(\omega) \qquad \text{for all destinations } \; n \,. \tag{A3}$$

Plugging this into the expression for profits and setting profits to zero (free entry condition) yields

$$\sum_n \tilde{q}_{ni}(\omega) = \frac{\sigma - 1}{1 - \beta} \tilde{f} \,. \tag{A4}$$

Let $\tilde{M}_i$ be the measure of firms from country i. Equating supply and demand for unskilled labour yields

$$\tilde{L} = \tilde{M}_i (\tilde{f} + \sum_n \tilde{q}_{ni})$$

or, substituting in (A4),

$$\tilde{M}_i = \tilde{M} \equiv \frac{\tilde{L}}{\tilde{f}} \frac{1 - \beta}{\sigma - \beta} \,. \tag{A5}$$

This implies that the measure of firms is the same in all countries. Thus, the consumer-goods market and market for unskilled labour are symmetric across countries with the exception of total income $y_n \mathcal{L}_n$. However, this income is spent equally across all varieties produced in all countries so that the production side of the model remains symmetric across both varieties and countries. We therefore consider an equilibrium with $\tilde{w}_i = \tilde{w}$ for some $\tilde{w}$ and all i. Let N be the number of countries. From (A2) and (A3),[42]

$$D_{ni}(\omega) = D_n = \mathcal{L}_n \left\{ (y_n/p_n)(\beta/\sigma)/(N\tilde{M}) \right\}^{1/(1-\beta)} \,. \tag{A6}$$

A firm from country i selling into country n demands D_n ads. The measure of such firms is $\tilde{M}$. Total demand for ads placed in front of country n consumers is $\sum_i \tilde{M}_i D_{ni}$. Plugging in $\tilde{M}_i$ from (A5) and D_{ni} from (A6), the total demand for ads in country n in the symmetric equilibrium is

$$\Sigma_i \tilde{M}_i D_{ni} = N\tilde{M} D_n = (y_n/p_n)^{1/(1-\beta)} \mathcal{L}_n \tilde{\kappa} \tag{A7}$$

where $\tilde{\kappa} \equiv (N\tilde{M})^{-\beta/(1-\beta)} (\beta/\sigma)^{1/(1-\beta)}$. From (A5), $\tilde{\kappa}$ is a constant. This is the demand for apps.

[42] The proof of (A6) is as follows. In the symmetric equilibrium, $\tilde{p}_i = \tilde{p} = \tilde{w}\sigma/(1-\sigma)$, $\tilde{p}_i^{-\sigma}/\tilde{P}^{1-\sigma} = \tilde{p}^{-\sigma}/\Sigma_i \tilde{M} \tilde{p}^{1-\sigma} = 1/(N\tilde{M}\tilde{p})$. Plugging these into (A2), $\tilde{p}\tilde{q}_{ni} = (D_{ni}/\mathcal{L}_n)^{\beta}(y_n \mathcal{L}_n)/(N\tilde{M})$. But from (A3), $\tilde{p}\tilde{q}_{ni} = (\sigma/\beta) p_n D_{ni}$. Equating these two expressions for $\tilde{p}\tilde{q}_{ni}$ and simplifying yields (A6).

Since there is one ad per app per user, the supply of ads is the measure of consumers in n choosing apps over the outside option. Restated, it is the share of consumers choosing apps over the outside option times the measure of consumers:

$$S_n \equiv \frac{\mathcal{U}_n - \delta'_{0n}}{\mathcal{U}_n} \mathcal{L}_n \,. \tag{A8}$$

From equation (2), $\mathcal{U}_n = \delta'_{0n} + \sum_i \sum_{a \in \mathcal{A}_{ni}} \alpha(\delta_a)\delta_a$ where $\mathcal{A}_{ni}$ is the set of apps from i available in n. An app from i with $\delta > \delta_i^A$ is available in all countries. Treating the number of apps as a continuous variable, we can rewrite this welfare expression as

$$\mathcal{U}_n = \delta'_{0n} + \sum\nolimits_i \left\{ M_i^A \int_{\delta_{Ai}}^{\infty} \alpha(\delta)\delta \, dG(\delta) \right\} . \tag{A9}$$

From (10), M_i^A and δ_i^A are constants. From (12), $\alpha(\delta)$ depends on constants and δ. Hence, the right side of (A9) is a constant. $\mathcal{U}_n$ is thus independent of ad prices and, more generally, is a constant. Thus, from (A8), the supply of ads S_n is a constant.

We next use a fixed-point theorem to show that there exists prices $\{\tilde{w}_i, p_i\}_{i=1}^N$ that are positive, sum to unity, and clear the markets for unskilled labour and ads. $D_n - S_n$ is the excess demand for ads in country n. $\tilde{M}_i(\tilde{f} + \sum_n \tilde{q}_{ni}) - \tilde{L}$ is the excess demand for unskilled labour in country i, which depends on the $\tilde{w}_i$ via the $\tilde{q}_{ni}$ (see equation A2 and the markup rule). These excess demand functions satisfy definition 17.B.2 of Mas-Colell, Whinston, and Green (1995). Existence follows from their theorem 17.C.2.

We conclude by informally defining general equilibrium in our setting:

1. Wages $\tilde{w}_i$ that clear national markets for unskilled labour.
2. Prices $\tilde{p}_i(\omega)$ and firm measures $\tilde{M}_i$ that clear international markets for consumer goods and set the profits of consumer-goods firms to zero.
3. Wages w_i^A that clear national markets for AI scientists.
4. Numbers of firms M_i^A that set the profits of app producers to zero.
5. Ad prices p_i that clear markets for ads.

Appendix C. Mobile App Firms that Choose Not to Use AI

In this section we modify the model to allow for the possibility that not all firms use AI. We note that in the data there are both big and small firms that do not use AI so we do not want a simple Melitz selection in which big firms use AI and small firms do not. To this end, we assume that a firm decides whether or not to enter the app market and whether or not to do so using AI. The firm thus faces a familiar technology-choice problem trading off fixed versus marginal costs. We also assume that there is free entry into both AI and non-AI apps so that, even after making the choice of whether or not to use AI, firms earn expected profits of zero.

In order to keep all of the main text model, we segment the labour markets for AI and non-AI labour. Then all of the results in the main text continue to hold. In particular, if a firm chooses to use AI then it hires AI workers at a wage w_i^A. If it chooses not to use AI then it hires non-AI workers at a wage w_i. These non-AI workers are also skilled (they understand programming, hardware, cloud computing etc). Let L_i be the

mass of skilled labour in the economy and let s_i be the share of these workers who are trained in AI so that $s_i L_i = L_i^A$ is the share of skilled labour with AI training and $(1 - s_i)L_i$ is the share that is not. s_i indexes Heckscher-Ohlin comparative advantage. In general equilibrium, countries with high s_i will have relatively low AI wages (low w_i^A/w_i) and hence a comparative advantage in AI-intensive apps. That is, firms in AI-abundant countries are more likely both to adopt AI and to use AI more intensively.

With this set up, the section 3 model continues to describe firms that choose to use AI. Consider a firm from country i that choose not to use AI. The firm pays an entry sunk cost f_{ei} and receives a demand draw δ from a Pareto distribution $G(\delta) = 1 - \delta^{-\gamma}$ where $\gamma > 1$. If the firm decides to produce the app it must pay a fixed costs f_i. It then operates in all markets.

Firm revenues are as in the section 3 model, but with $\alpha = 1$. From the discussion preceding equation (4), revenues are δP and so the firm's profit function is

$$\pi_i(\delta) = \delta P - w_i f_i \,.$$

The cutoff for producing is defined by $\pi_i(\delta_i) = 0$ so that the firm produces if

$$\delta > \delta_i \equiv \frac{w_i}{P} f_i \,.$$

As in section 3, we solve for the mass of firms who pay the sunk cost (M_i) and the wage w_i using the free entry condition

$$\int_{\delta_i}^{\infty} \pi_i(\delta) dG(\delta) = w_i f_{ei} \tag{A10}$$

along with the labour market clearing condition

$$(1 - s_i)L_i = M_i \left\{ f_{ei} + [1 - G(\delta_i)] f_i \right\} . \tag{A11}$$

The right-hand side of equation (A11) states that each firm must pay the sunk cost f_{ei} and, with probability $1 - G(\delta_i) = \delta_i^{-\gamma}$, produce and pay the fixed cost f_i. Solving equations (A10)–(A11) with Pareto is standard and yields

$$\delta_i = \left(\frac{f_i}{f_{ei}} \frac{1}{\gamma - 1} \right)^{1/\gamma} , \quad \frac{w_i}{P} = \frac{\delta_i}{f_i} , \quad \text{and} \quad M_i = \frac{(1 - s_i)L_i}{\gamma f_{ei}} \,. \tag{A12}$$

This is very similar to its AI counterpart equation (10). Likewise, the proof is identical to the appendix Appendix C proof of (10).

To imbed this into the general equilibrium model with advertising (online Appendix B), we need only redefine per capita income: $y_n = [w_n(1 - s_n)L_n + w_n^A s_n L_n^A + \tilde{w}_n \tilde{L}]/\mathcal{L}_n$ where $\mathcal{L}_n = L_n + \tilde{L}$. Everything else goes through as before.

Appendix D. Incumbent Firms and an Instrument for the Extensive Margin of AI

We argued in the main text that before there were smartphones, mobile apps, or much commercialization of AI, a number of firms were engaged in AI research without a mobile app in mind. This suggests a firm-level instrument for AI deployment in mobile apps, namely, a dummy indicating whether or not a firm was developing AI capabilities in some year well before the 2015 start of our sample. This is easy to incorporate directly into our model. Assume that there is an exogenous measure $\widetilde{M}_i^A$ of firms that have already incurred the sunk costs f_{ei}^A of setting up for using AI. $\widetilde{M}_i^A$ is exogenous to the model. Assume further that these firms have a distribution of δ that is Pareto: $G(\delta) = 1 - \delta^{-\gamma}$. Then the only modification to the model is that the labour-market clearing condition (9) now becomes

$$L_i^A = M_i^A f_{ei}^A + \left(M_i^A + \widetilde{M}_i^A\right) \int_{\delta_i^A}^{\infty} \left[f_i^A + \frac{\eta - 1}{\eta} \alpha^{\frac{\eta}{\eta-1}}\right] dG(\delta)\,.$$

As in the baseline model, the integral equals $f_{ei}^A(\gamma - 1)$. Plugging this into the previous equation yields

$$L_i^A = M_i^A f_{ei}^A + \left(M_i^A + \widetilde{M}_i^A\right) f_{ei}^A(\gamma - 1) = M_i^A f_{ei}^A \gamma + \widetilde{M}_i^A f_{ei}^A(\gamma - 1)\,.$$

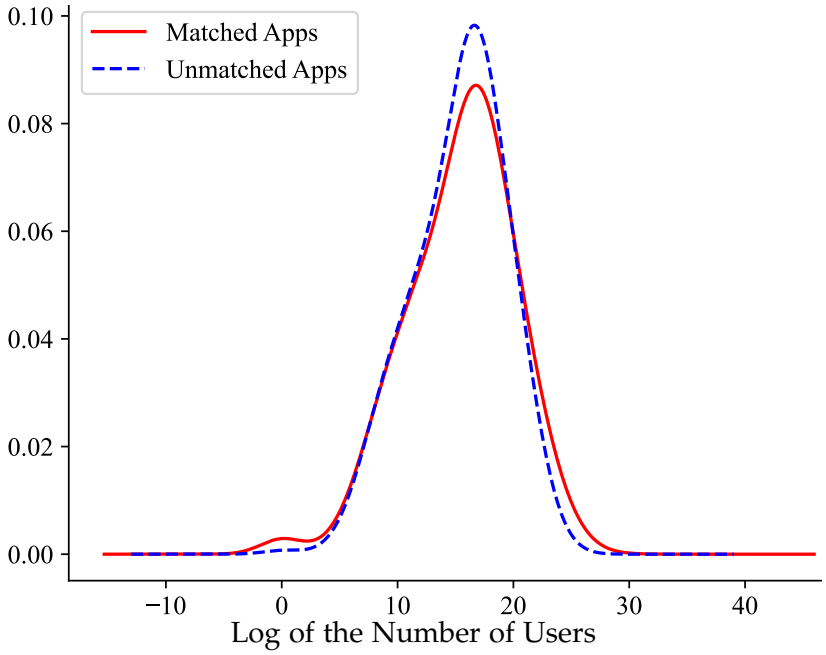
Solving for M_i^A and plugging in $f_{ei}^A = (L_i^A)^\psi f_{ei}$ from equation (3) yields

$$M_i^A = \left[L_i^A - \widetilde{M}_i^A (L_i^A)^\psi f_{ei}(\gamma - 1)\right] / \left[(L_i^A)^\psi f_{ei}\gamma\right]\,. \tag{A13}$$

We assume that the number of incumbents $\widetilde{M}_i^A$ is small enough that $M_i^A > 0$.

Combining this with the extension in online Appendix C we have the following. Incumbents will always choose the AI technology over the non-AI technology. Hence, conditional on being an incumbent and surviving, the firm deploys AI with probability 1. In contrast, recall that the number of firms that do not deploy AI is M_i (equation A12). Hence, conditional on being a non-incumbent and surviving, the probability of deploying AI is $M_i^A[1 - G(\delta_i^A)] / \left\{M_i^A[1 - G(\delta_i^A)] + M_i[1 - G(\delta_i)]\right\} < 1$. In short, incumbents have a higher probability of deploying AI than do non-incumbents.

Figure A2: Distribution of Users Across Apps: Comparison of Samples



Notes: It is of interest to examine the representativeness of our data. To this end, we downloaded the top 5,000 firms along with all their apps and distinguished between the matched sample (1,276 firms with 35,575 apps) and the unmatched sample (5,000 less 1,276 firms with about 42,000 apps). The largest apps are all in the matched sample so that it accounts for 71% of users in the two samples. However, aside from the largest 80 apps (Facebook is #1 and Didi is #80), the two distributions of users are very similar. For example, the median app in the unmatched sample is 59,000 users, which is only a little smaller than the median of 65,000 users in the matched sample. The figure compares the density of users for the two samples and shows how similar they are aside from the right tail.

Figure A3: National Abundance in 2008: Citations of AI Journal Articles by Country

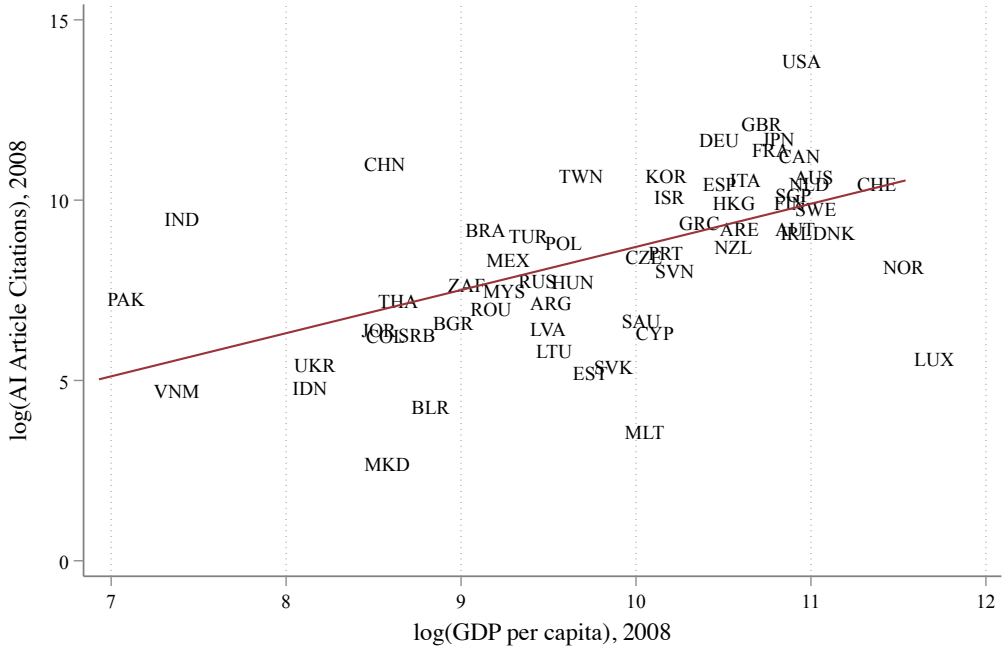


Table A1: A Placebo: Using the Patents of Other Firms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS: ln(*Foreign Users* $_{at}$)						
ln(1+*AI* $_{at}$)	1.57*	1.46*	1.62*	1.65*	1.19*	1.55*	1.23*
	(0.14)	(0.15)	(0.16)	(0.16)	(0.13)	(0.15)	(0.14)
ln(1+*placebo AI* $_{at}$)	0.00	0.00	0.01	0.01	0.00	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
ln(1+*undirected AI* $_{at}$)					-0.65*		-0.64*
					(0.11)		(0.11)
ln(1+*non AI* $_{ft}$)						-0.99*	-0.36
						(0.34)	(0.35)
Obs.	125,486	125,486	125,467	125,024	125,486	125,486	125,486
R^2	0.25	0.25	0.25	0.27	0.25	0.25	0.25
FEs	*f, c, t*	*f, ct*	*f, ct, it*	*f, cit*	*f, ct*	*f, ct*	*f, ct*
	IV: ln(*Foreign Users* $_{at}$)						
ln(1+*AI* $_{at}$)	2.97*	2.85*	2.92*	3.25*	2.61*	2.94*	2.71*
	(0.47)	(0.51)	(0.55)	(0.58)	(0.54)	(0.53)	(0.57)
ln(1+*placebo AI* $_{at}$)	-0.04	-0.04	-0.03	-0.04	-0.04	-0.03	-0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
ln(1+*undirected AI* $_{at}$)					-0.45*		-0.40*
					(0.13)		(0.14)
ln(1+*non AI* $_{ft}$)						-1.98*	-1.56*
						(0.50)	(0.57)
Weak Instrument *F* (KP)	1,436	1,242	1,132	1,039	1,254	1,174	1,180
	First Stage: ln(1+*AI* $_{at}$)						
Z $_{at}$	0.55*	0.52*	0.51*	0.50*	0.49*	0.51*	0.47*
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
ln(1+*placebo AI* $_{at}$)	0.03*	0.03*	0.03*	0.03*	0.03*	0.02*	0.02*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ln(1+*undirected AI* $_{at}$)					-0.12*		-0.14*
					(0.01)		(0.01)
ln(1+*non AI* $_{ft}$)						0.67*	0.76*
						(0.04)	(0.04)
	Reduced Form: ln(*Foreign Users* $_{at}$)						
Z $_{at}$	1.65*	1.49*	1.48*	1.63*	1.28*	1.49*	1.26*
	(0.26)	(0.27)	(0.28)	(0.29)	(0.26)	(0.27)	(0.27)
ln(1+*placebo AI* $_{at}$)	0.04*	0.04*	0.05*	0.05*	0.03*	0.04*	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
ln(1+*undirected AI* $_{at}$)					-0.77*		-0.78*
					(0.11)		(0.11)
ln(1+*non AI* $_{ft}$)						-0.01	0.50
						(0.33)	(0.34)

Notes: This table provides the OLS, IV, first-stage, and reduced-form estimates that correspond to the IV estimates in table 2. Standard errors are clustered at the app level. * indicates 1% significance.

Table A2: Full Version of Table 4 Magnitudes

	(1)	(2)	(3)
	OLS: ln(*Foreign Users* $_{at}$)		
ln(1+AI_{at})	1.48*		
	(0.14)		
lHS(AI_{at})		1.39*	
		(0.13)	
ln(AI_{at})			1.55*
			(0.17)
Obs.	125,486	125,486	27,824
R^2	0.25	0.25	0.20
FEs	*f, ct*	*f, ct*	*f, ct*
	IV: ln(*Foreign Users* $_{at}$)		
ln(1+AI_{at})	2.67*		
	(0.44)		
lHS(AI_{at})		2.82*	
		(0.47)	
ln(AI_{at})			3.78*
			(0.72)
Weak Instrument F (KP)	1,384	1,346	700
	First Stage: ln(1+AI_{at})		
	ln(1+AI_{at})	lHS(AI_{at})	ln(AI_{at})
Z_{at}	0.60*	0.57*	0.45*
	(0.02)	(0.02)	(0.02)
	Reduced Form: ln(*Foreign Users* $_{at}$)		
Z_{at}	1.60*	1.60*	1.70*
	(0.26)	(0.26)	(0.32)

Notes: This table provides the OLS, first-stage, and reduced-form estimates corresponding to the IV estimates in table 4.

Table A3: Sensitivity: Baseline with App Store Category Definition of $\eta_{c(a),t}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS: ln(*Foreign Users* $_{at}$)						
ln(1+*AI* $_{at}$)	1.58*	1.48*	1.67*	1.69*	1.20*	1.58*	1.24*
	(0.13)	(0.14)	(0.14)	(0.15)	(0.12)	(0.15)	(0.13)
ln(1+*undirected AI* $_{at}$)					-0.65*		-0.64*
					(0.11)		(0.11)
ln(1+*non AI* $_{ft}$)						-0.95*	-0.35
						(0.34)	(0.34)
Obs.	125,486	125,486	125,467	125,024	125,486	125,486	125,486
R^2	0.25	0.25	0.25	0.27	0.25	0.25	0.25
FEs	*f, c, t*	*f, ct*	*f, ct, it*	*f, cit*	*f, ct*	*f, ct*	*f, ct*
	IV: ln(*Foreign Users* $_{at}$)						
ln(1+*AI* $_{at}$)	2.64*	2.44*	2.63*	2.88*	2.71*	2.66*	2.90*
	(0.34)	(0.38)	(0.44)	(0.44)	(0.35)	(0.43)	(0.38)
ln(1+*undirected AI* $_{at}$)					-0.39*		-0.33*
					(0.12)		(0.12)
ln(1+*non AI* $_{ft}$)						-2.01*	-2.10*
						(0.53)	(0.51)
Weak Instrument *F* (KP)	2019	1620	1303	1430	2170	1328	1887
	First Stage: ln(1+*AI* $_{at}$)						
Z $_{at}$	0.82*	0.77*	0.72*	0.77*	0.88*	0.69*	0.80*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
ln(1+*undirected AI* $_{at}$)					-0.20*		-0.21*
					(0.01)		(0.01)
ln(1+*non AI* $_{ft}$)						0.88*	0.96*
						(0.04)	(0.04)
	Reduced Form: ln(*Foreign Users* $_{at}$)						
Z $_{at}$	2.17*	1.87*	1.90*	2.21*	2.37*	1.84*	2.32*
	(0.28)	(0.29)	(0.32)	(0.33)	(0.30)	(0.30)	(0.30)
ln(1+*undirected AI* $_{at}$)					-0.94*		-0.95*
					(0.12)		(0.12)
ln(1+*non AI* $_{ft}$)						0.34	0.68
						(0.32)	(0.33)

Notes: This table is identical to table 1 except that the AI intensity of an industry ($\eta_{c(a),t}$) is redefined. In constructing $\eta_{c(a),t}$ we defined industries (c) in two ways. In the main text, an industry is defined as the set of apps that have a high cosine similarity with app a. In this table we define an industry as the set of apps that are in the same App Store category as a (but still excluding app a as well as all apps developed by the developer of a). See section 3.5 for details. *Comparing this table to table 1, the two definitions yield almost identical results.* Standard errors are clustered at the app a level. * indicates significance at the 1% level.

Table A4: Sensitivity to Patent Definitions and Firm Financials

	(1)	(2)	(3)	(4)	(5)
	Baseline Specification	Patent Families	Patent Citations	Sub-sample w. Financial Data	Sub-sample w. Financial Data
	OLS: $\ln(Foreign\ Users_{at})$				
$\ln(1+AI_{at})$	1.48*	1.58*	2.09*	1.44*	1.42*
	(0.14)	(0.14)	(0.22)	(0.15)	(0.15)
$\ln(Assets_{ft})$					1.00*
					(0.26)
$\ln(Revenue_{ft})$					0.46*
					(0.16)
Obs.	125,486	125,486	125,486	62,183	62,183
R^2	0.25	0.25	0.25	0.23	0.23
FEs	*f, ct*	*f, ct*	*f, ct*	*f, ct*	*f, ct*
	IV: $\ln(Foreign\ Users_{at})$				
$\ln(1+AI_{at})$	2.67*	2.61*	5.18*	3.21*	3.17*
	(0.44)	(0.43)	(0.88)	(0.60)	(0.60)
$\ln(Assets_{ft})$					0.73*
					(0.27)
$\ln(Revenue_{ft})$					0.57*
					(0.16)
Weak Instrument F (KP)	1,384	1,475	467	924	928
	First Stage: $\ln(1+AI_{at})$				
Z_{at}	0.60*	0.61*	0.31*	0.49*	0.49*
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
$\ln(Assets_{ft})$					0.16*
					(0.01)
$\ln(Revenue_{ft})$					-0.08*
					(0.01)
	Reduced Form: $\ln(Foreign\ Users_{at})$				
Z_{at}	1.60*	1.60*	1.60*	1.57*	1.56*
	(0.26)	(0.26)	(0.26)	(0.29)	(0.29)
$\ln(Assets_{ft})$					1.22*
					(0.26)
$\ln(Revenue_{ft})$					0.33
					(0.16)

Notes: This table reports alternative specifications to our baseline specification in column 2 of table 1. Column 1 repeats column 2 of table 1. The key regressor $\ln(1 + AI_{at})$ is the AI deployed in app a (see equation 1) and is constructed as the ρ_{ap}-weighted sum of patents. Column 2 replaces patents with patent families. *Comparing columns 1 and 2, this barely changes the OLS and IV estimates.* Column 3 replaces patents with patent citations. *This increases the OLS and IV estimates.* The increased size likely reflects the heavy right skew of patent citations. Columns 4–5 restrict the sample to the half of all observations with financial data. Column 4 is our baseline specification for this subsample. Column 5 adds the log of firm assets and the log of firm revenues to the specification. *Comparing columns 4 and 5, adding financials does not change either the OLS or IV estimates.* Standard errors are clustered at the app level. * indicates 1% significance.

Table A5: Sensitivity to Adding App-Level Controls

	OLS: ln(*Foreign Users* $_{at}$)			IV: ln(*Foreign Users* $_{at}$)			First Stage: ln(1+*AI* $_{at}$)		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
ln(1+*AI* $_{at}$)	1.51*	1.04*	0.93*	2.34*	2.52*	2.65*			
	(0.15)	(0.14)	(0.14)	(0.51)	(0.47)	(0.56)			
ln(1+*undirected AI* $_{at}$)			-0.37*			-0.06			-0.16*
			(0.11)			(0.15)			(0.01)
ln(1+*non AI* $_{ft}$)			-0.44			-2.24*			0.97*
			(0.35)			(0.67)			(0.05)
Age $_{a}$		-0.08*	-0.08*		-0.08*	-0.08*		0.00	0.00
		(0.01)	(0.01)		(0.01)	(0.01)		(0.00)	(0.00)
Price $_{a}$		-0.00*	-0.00*		-0.00*	-0.00*		0.00	0.00
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
Rating $_{a}$		0.80*	0.80*		0.80*	0.80*		0.00	0.00
		(0.02)	(0.02)		(0.02)	(0.02)		(0.00)	(0.00)
In-app-purchase		0.90*	0.90*		0.90*	0.90*		0.01*	0.00*
revenue dummy $_{a}$		(0.05)	(0.05)		(0.05)	(0.05)		(0.00)	(0.00)
Show ads dummy $_{a}$		1.17*	1.17*		1.16*	1.16*		0.00*	0.00*
		(0.05)	(0.05)		(0.05)	(0.05)		(0.00)	(0.00)
Buy ads dummy $_{a}$		1.38*	1.38*		1.37*	1.37*		0.01*	0.01*
		(0.05)	(0.05)		(0.05)	(0.05)		(0.00)	(0.00)
Z_{at}							0.57*	0.57*	0.48*
							(0.02)	(0.02)	(0.02)
Obs.	109,593	109,593	109,593	109,593	109,593	109,593	109,593	109,593	109,593
FEs	*f, ct*	*f, ct*	*f, ct*	*f, ct*	*f, ct*	*f, ct*	*f, ct*	*f, ct*	*f, ct*
R^2	0.26	0.36	0.36						
Weak Instrument *F* (KP)				1,054	1,063	964			

Notes: This table explores the effects of adding app-level characteristics to our baseline specification. Each group of three columns is either OLS, IV, or first stage (see the column headers). Each column 1 repeats column 2 of table 1, but for the smaller sample of apps that have app-level characteristics. Each column 2 adds app-level characteristics. Each column 3 adds undirected AI patents and non-AI patents, as in column 7 of table 1. *From the IV rows, the baseline coefficient of 2.34 changes very little when additional covariates are added.* Standard errors are clustered at the app level. * indicates 1% significance.

Table A6: Sensitivity: $D_{f(a),\tau}$ for $\tau = 2006, \ldots, 2011$

	(1)	(2)	(3)	(4)	(5)	(6)
	2006	2007	2008	2009	2010	2011
	IV: ln(*Foreign Users* $_{at}$)					
ln(1+*AI* $_{at}$)	2.88*	2.66*	2.67*	2.67*	2.64*	2.65*
	(0.47)	(0.44)	(0.44)	(0.44)	(0.43)	(0.42)
Obs.	125,486	125,486	125,486	125,486	125,486	125,486
FEs	*f, ct*	*f, ct*	*f, ct*	*f, ct*	*f, ct*	*f, ct*
Weak Instrument F (KP)	1,188	1,360	1,384	1,398	1,420	1,445
	First Stage: ln(1+*AI* $_{at}$)					
Z_{at}	0.59*	0.60*	0.60*	0.60*	0.61*	0.61*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
	Reduced Form: ln(*Foreign Users* $_{at}$)					
Z_{at}	1.70*	1.60*	1.60*	1.59*	1.60*	1.61*
	(0.28)	(0.27)	(0.26)	(0.26)	(0.26)	(0.26)

Notes: Column 3 repeats our baseline specification from column 2 of table 1. In the equation (14) definition of the instrument Z_{at}, there is a term $D_{f(a),2008}$ which equals one if firm $f(a)$ (the firm which developed app a) had filed an AI patent on or before 2008. 2008 was chosen because the first mobile apps were made available in the second half of 2008. In columns 1–6 we replace 2008 with the years 2006, ..., 2011, i.e., the years before commercial applications of machine learning appeared. The panels are IV, first stage and reduced form. OLS is the same in all columns and appears in column 2 of table 1. *Comparing columns 1–6, the choice of year makes no difference.* Standard errors are clustered at the app level. * indicates 1% significance.

Table A7: Sensitivity to the Choice of Cutoff $\rho_{ap} > 0.2$

	(1)	(2)	(3)	(4)	(5)	(6)
	0.0	0.1	0.2	0.3	0.4	0.5
	OLS: ln(*Foreign Users* $_{at}$)					
ln(1+AI_{at})	1.50*	1.52*	1.48*	1.11*	0.85*	0.59*
	(0.15)	(0.15)	(0.14)	(0.10)	(0.07)	(0.06)
Obs.	125,486	125,486	125,486	125,486	125,486	125,486
R^2	0.25	0.25	0.25	0.25	0.25	0.25
FEs	*f, ct*	*f, ct*	*f, ct*	*f, ct*	*f, ct*	*f, ct*
Share of *ap* pairs	99.95%	96.12%	62.36%	17.49%	2.06%	0.10%
	IV: ln(*Foreign Users* $_{at}$)					
ln(1+AI_{at})	4.63*	4.31*	2.67*	1.85*	1.85*	2.70*
	(0.79)	(0.74)	(0.44)	(0.22)	(0.30)	(0.42)
Weak Instrument *F*	1,571	1,550	1,384	1,516	297	152
	First Stage: ln(1+AI_{at})					
Z_{at}	0.25*	0.29*	0.60*	1.09*	0.65*	0.19*
	(0.01)	(0.01)	(0.02)	(0.03)	(0.04)	(0.02)
	Reduced Form: ln(*Foreign Users* $_{at}$)					
Z_{at}	1.18*	1.24*	1.60*	2.01*	1.20*	0.51*
	(0.20)	(0.22)	(0.26)	(0.24)	(0.19)	(0.07)

Notes: Column 3 repeats our baseline specification from column 2 of table 1. In constructing $AI_{at} = \Sigma_p \rho_{ap}$ we only summed over patents for which $\rho_{ap} > 0.2$. In columns 1–2, we replace 0.2 with 0.0 and 0.1, respectively, and recalculate AI_{at}. *This does not change the OLS results but substantially raises the IV results.* In columns 4–6, we replace 0.2 with 0.3, 0.4, and 0.5, respectively, and recalculate AI_{at}. *This lowers the OLS and IV results, but both remain economically and statistically very large.* The 'Share of ap pairs' reports the share of all possible ap pairs that are above the column's $\overline{\rho}$ threshold. By this measure, columns 1, 2, 5, and 6 are extreme cutoffs. Standard errors are clustered at the app level. * indicates 1% significance.

Table A8: Sensitivity: Deep Learning Patents Only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS: ln(*Foreign Users* $_{at}$)						
	All Patents	Deep Learning Patents					
ln(1+*AI* $_{at}$)	1.48*	1.27*	1.41*	1.45*	1.02*	1.31*	1.01*
	(0.14)	(0.13)	(0.14)	(0.14)	(0.12)	(0.14)	(0.12)
ln(1+*undirected AI* $_{at}$)					-0.70*		-0.70*
					(0.11)		(0.11)
ln(1+*non AI* $_{ft}$)						-0.49	0.06
						(0.33)	(0.33)
Obs.	125,486	125,486	125,467	125,024	125,486	125,486	125,486
R^2	0.25	0.25	0.25	0.27	0.25	0.25	0.25
FEs	*f, ct*	*f, ct*	*f, ct, it*	*f, cit*	*f, ct*	*f, ct*	*f, ct*
	IV: ln(*Foreign Users* $_{at}$)						
	All Patents	Deep Learning Patents					
ln(1+*AI* $_{at}$)	2.67*	2.29*	2.32*	2.53*	2.00*	2.34*	2.05*
	(0.44)	(0.44)	(0.47)	(0.50)	(0.47)	(0.45)	(0.49)
ln(1+*undirected AI* $_{at}$)						-1.34*	-0.87
						(0.49)	(0.54)
ln(1+*non AI* $_{ft}$)					-0.56*		-0.54*
					(0.13)		(0.14)
Weak Instrument F (KP)	1,384	1,070	983	909	1,070	1,038	1,038
	First Stage: ln(1+*AI* $_{at}$)						
	All Patents	Deep Learning Patents					
Z_{at}	0.60*	0.71*	0.69*	0.68*	0.67*	0.69*	0.64*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
ln(1+*undirected AI* $_{at}$)					-0.12*		-0.14*
					(0.01)		(0.01)
ln(1+*non AI* $_{ft}$)						0.78*	0.84*
						(0.04)	(0.04)
	Reduced Form: ln(*Foreign Users* $_{at}$)						
	All Patents	Deep Learning Patents					
Z_{at}	1.60*	1.63*	1.61*	1.73*	1.34*	1.62*	1.31*
	(0.26)	(0.31)	(0.33)	(0.34)	(0.31)	(0.31)	(0.31)
ln(1+*undirected AI* $_{at}$)					-0.80*		-0.82*
					(0.11)		(0.11)
ln(1+*non AI* $_{ft}$)						0.48	0.86*
						(0.32)	(0.33)

Notes: Column 1 repeats our baseline specification from column 2 of table 1. In this column, AI_{at} is constructed using all of the firm's AI patents. In the remaining columns, AI_{at} is defined using only the firm's deep learning patents. *Comparison across columns shows that deep learning has a large impact on foreign users.* Standard errors are clustered at the app level. * indicates 1% significance.

Table A9: Sensitivity to Observations in the Top 1%, 5%, and 10% of Foreign Users

	(1)	(2)	(3)	(4)
		Omit App-Year Pairs in Top x% of User Distribution:		
	Baseline	1%	5%	10%
	OLS: ln(*Foreign Users* $_{at}$)			
ln(1+AI_{at})	1.48*	1.32*	1.18*	1.02*
	(0.14)	(0.13)	(0.13)	(0.13)
Obs.	125,486	124,230	119193	112,895
R^2	0.25	0.24	0.22	0.21
FEs	*f, ct*	*f, ct*	*f, ct*	*f, ct*
	IV: ln(*Foreign Users* $_{at}$)			
ln(1+AI_{at})	2.67*	2.59*	2.54*	2.35*
	(0.44)	(0.43)	(0.42)	(0.40)
Weak Instrument F (KP)	1,384	982	958	926
	First Stage: ln(1+AI_{at})			
Z_{at}	0.60*	0.60*	0.60*	0.61*
	(0.02)	(0.02)	(0.02)	(0.02)
	Reduced Form: ln(*Foreign Users* $_{at}$)			
Z_{at}	1.60*	1.56*	1.53*	1.43*
	(0.26)	(0.26)	(0.25)	(0.24)

Notes: Column 1 repeats our baseline specification from column 2 of table 1. Column 2 omits the 1% of observations with the highest values of $ForeignUsers_{at}$. Columns 3 and 4 repeat this for the highest 5% and 10% of observations. *The table shows that the impacts of AI are felt even for apps that are much smaller than blockbuster apps such as Facebook.* Standard errors are clustered at the app level. * indicates 1% significance.

Table A10: Sensitivity: Baseline with Scaled AI Abundance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS: ln(*Foreign Users* $_{at}$)						
ln(1+*AI* $_{at}$)	1.58*	1.48*	1.67*	1.69*	1.20*	1.58*	1.24*
	(0.13)	(0.14)	(0.14)	(0.15)	(0.12)	(0.15)	(0.13)
ln(1+*undirected AI* $_{at}$)					-0.65*		-0.64*
					(0.11)		(0.11)
ln(1+*non AI* $_{ft}$)						-0.95*	-0.35
						(0.34)	(0.34)
Obs.	125,486	125,486	125,467	125,024	125,486	125,486	125,486
R^2	0.25	0.25	0.25	0.27	0.25	0.25	0.25
FEs	*f, c, t*	*f, ct*	*f, ct, it*	*f, cit*	*f, ct*	*f, ct*	*f, ct*
	IV: ln(*Foreign Users* $_{at}$)						
ln(1+*AI* $_{at}$)	3.08*	2.94*	3.01*	3.24*	2.89*	3.14*	3.09*
	(0.34)	(0.38)	(0.41)	(0.43)	(0.38)	(0.41)	(0.42)
ln(1+*undirected AI* $_{at}$)					-0.36*		-0.29
					(0.12)		(0.13)
ln(1+*non AI* $_{ft}$)						-2.48*	-2.30*
						(0.51)	(0.54)
Weak Instrument *F* (KP)	2,266	1,883	1,662	1,539	2,025	1,653	1,788
	First Stage: ln(1+*AI* $_{at}$)						
Z_{at}	0.99*	0.93*	0.91*	0.91*	0.91*	0.86*	0.83*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
ln(1+*undirected AI* $_{at}$)					-0.16*		-0.18*
					(0.01)		(0.01)
ln(1+*non AI* $_{ft}$)						0.89*	0.97*
						(0.04)	(0.04)
	Reduced Form: ln(*Foreign Users* $_{at}$)						
Z_{at}	3.03*	2.74*	2.73*	2.95*	2.63*	2.71*	2.58*
	(0.33)	(0.35)	(0.37)	(0.39)	(0.34)	(0.35)	(0.35)
ln(1+*undirected AI* $_{at}$)					-0.83*		-0.85*
					(0.11)		(0.11)
ln(1+*non AI* $_{ft}$)						0.32	0.69
						(0.32)	(0.33)

Notes: This table is identical to table 1 except that the AI abundance of a country (L^A_{it}) is redefined. The concern is that the L^A_{it} are correlated with country size and so may be picking up other things that belong in the second stage such as demand-side home-market effects and supply-side inherent scalability of AI. We therefore divide L^A_{it} by country i's capital stock K^A_{it}. K^A_{it} is calculated using constant dollar 2005 prices ($USD) from Feenstra, Inklaar, and Timmer (2015). *Comparing this table with table 1, scaling has only a modest effect on the results.* Standard errors are clustered at the app a level. * indicates significance at the 1% level.

Table A11: OECD Digital Services Trade Restrictiveness Index, 2020

Rank	Country	*Data Transfer* $_{n,2020}$	*Data Develop* $_{n,2020}$	Total	Rank	Country	*Data Transfer* $_{n,2020}$	*Data Develop* $_{n,2020}$	Total
1	Saudi Arabia	0.20	0.09	0.29	31	Denmark	0.04	0.02	0.06
2	Kazakhstan	0.12	0.11	0.23	32	Ireland	0.04	0.02	0.06
3	Indonesia*	0.16	0.07	0.22	33	Lithuania	0.04	0.02	0.06
4	Egypt*	0.08	0.13	0.21	34	Slovakia	0.04	0.02	0.06
5	China*	0.12	0.09	0.21	35	Luxembourg	0.04	0.02	0.06
6	Kenya	0.12	0.09	0.21	36	Netherlands*	0.04	0.02	0.06
7	Nigeria*	0.08	0.11	0.19	37	Argentina*	0.04	0.02	0.06
8	Pakistan*	0.08	0.11	0.19	38	Slovenia	0.04	0.02	0.06
9	India*	0.12	0.07	0.19	39	Peru	0.04	0.02	0.06
10	Russia*	0.12	0.07	0.19	40	Austria	0.04	0.02	0.06
11	Turkey*	0.08	0.07	0.15	41	Spain*	0.04	0.02	0.06
12	Vietnam	0.04	0.09	0.13	42	Finland	0.04	0.02	0.06
13	Greece	0.08	0.04	0.12	43	Thailand	0.04	0.02	0.06
14	Singapore	0.08	0.04	0.12	44	Mexico*	0.04	0.00	0.04
15	South Korea*	0.08	0.04	0.12	45	Colombia	0.04	0.00	0.04
16	Brazil*	0.12	0.00	0.12	46	Australia*	0.04	0.00	0.04
17	Germany*	0.08	0.02	0.10	47	Norway	0.04	0.00	0.04
18	Belgium*	0.08	0.02	0.10	48	Switzerland	0.04	0.00	0.04
19	Philippines*	0.04	0.04	0.08	49	South Africa*	0.04	0.00	0.04
20	Hungary	0.04	0.04	0.08	50	Uruguay	0.04	0.00	0.04
21	Czech Republi	0.04	0.04	0.08	51	UK*	0.04	0.00	0.04
22	Malaysia	0.04	0.04	0.08	52	Israel	0.04	0.00	0.04
23	Italy*	0.04	0.04	0.08	53	Japan*	0.04	0.00	0.04
24	France*	0.04	0.04	0.08	54	Madagascar	0.04	0.00	0.04
25	Poland	0.04	0.04	0.08	55	Ecuador	0.00	0.02	0.02
26	Portugal	0.04	0.04	0.08	56	Dominican	0.00	0.02	0.02
27	Guatemala	0.04	0.04	0.08	57	Chile	0.00	0.00	0.00
28	Sweden	0.08	0.00	0.08	58	USA*	0.00	0.00	0.00
29	New Zealand	0.08	0.00	0.08	59	Canada*	0.00	0.00	0.00
30	Ghana	0.00	0.07	0.07	60	Costa Rica	0.00	0.00	0.00

Notes: This table reports $Data\,Transfer_{n,2020}$ and $Data\,Development_{n,2020}$. The 'total' column is $Data\,Transfer_{n,2020} + Data\,Development_{n,2020}$. An asterisk indicates that the country is in our subsample of the largest 25 economies.

Table A12: First-Stage and Reduced-Form of Table 7: AI Externalities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	First Stage: $\ln(1+AI_{at})$						
$\ln(1+external\ AI_{at})$	0.03*	0.03*	0.03*	0.03*	0.02*	0.03*	0.02*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Z_{at}	0.61*	0.57*	0.53*	0.53*	0.54*	0.54*	0.49*
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
$\ln(1+undirected\ AI_{at})$					-0.12*		-0.14*
					(0.01)		(0.01)
$\ln(1+non\ AI_{ft})$						0.91*	0.98*
						(0.04)	(0.04)
	Reduced Form: $\ln(Foreign\ Users_{at})$						
$\ln(1+external\ AI_{at})$	0.37*	0.37*	0.37*	0.37*	0.34*	0.37*	0.34*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Z_{at}	1.45*	1.29*	1.25*	1.39*	1.16*	1.27*	1.13*
	(0.26)	(0.27)	(0.28)	(0.29)	(0.26)	(0.27)	(0.26)
$\ln(1+undirected\ AI_{at})$					-0.46*		-0.48*
					(0.12)		(0.12)
$\ln(1+non\ AI_{ft})$						0.47	0.69
						(0.32)	(0.33)

Notes: This table reports the first-stage and reduced-form estimates for the IV in table 7. Standard errors are clustered at the app level. * indicates 1% significance.

Table A13: Sensitivity to the Number of Sentence Embeddings Used for ρ_{ap}

	(1)	(2)	(3)
	Top Sentence	Top 5 Sentences	Top 10 Sentences
	OLS: ln(*Foreign Users* $_{at}$)		
ln(1+AI_{at})	1.28*	1.48*	1.30*
	(0.13)	(0.14)	(0.14)
Obs.	125,486	125,486	125,486
R^2	0.25	0.25	0.25
FEs	*f, ct*	*f, ct*	*f, ct*
	IV: ln(*Foreign Users* $_{at}$)		
ln(1+AI_{at})	1.98*	2.67*	3.55*
	(0.36)	(0.44)	(0.60)
Weak Instrument F (KP)	1,550	1,384	1,564
	First Stage: ln(1+AI_{at})		
Z_{at}	0.60*	0.85*	0.37*
	(0.02)	(0.02)	(0.01)
	Reduced Form: ln(*Foreign Users* $_{at}$)		
Z_{at}	1.60*	1.68*	1.31*
	(0.26)	(0.30)	(0.22)

Notes: Column 2 repeats our baseline specification from column 2 of table 1. As described in Appendix B, word embeddings are calculated at the sentence level. To collapse this across sentences we use the following industry standard. For each pair of sentences, one from the app description and one from the patent text, we calculate a cosine similarity. We then average across sentence-level cosine similarities to obtain ρ_{ap}. In our baseline we average across the top 5 sentence-level cosine similarities. In column 3 we average across the top 10 sentence-level cosine similarities. In column 1 we use only the largest sentence-level cosine similarity. *Comparison of columns 1–3 shows that we obtain large impacts for all methods.* Standard errors are clustered at the app level. * indicates 1% significance.