Does Machine Translation Affect International Trade? Evidence from a Large Digital Platform*

Erik Brynjolfsson
MIT and NBER

Xiang Hui
Washington University in St. Louis and MIT

Meng Liu†
Washington University in St. Louis and MIT

August 1, 2018

Abstract

Artificial intelligence (AI) is matching or surpassing human performance in a growing number of domains. However, there is limited empirical evidence of its economic effects. Using a major e-commerce platform as the empirical context, we study an application at the leading frontier of AI: machine translation. We find that the introduction of a machine translation system has already had a significant effect on international trade on this platform, increasing export quantity by 17.5%. Furthermore, heterogeneous treatment effects are consistent with a substantial reduction in buyers' translation-related search costs due to the introduction of this system. Our results provide causal evidence that language barriers significantly hinder trade and that AI has already substantially improved overall economic efficiency.

Keywords: Artificial Intelligence, International Trade, Machine Translation, Machine Learning, Digital Platforms

*We thank David Atkin, Andrey Fradkin, Avi Gannamaneni, and Avi Goldfarb for helpful discussions. We also acknowledge the support of the MIT Initiative on the Digital Economy (http://ide.mit.edu/) and thank eBay, especially Brian Bieron, for providing the data for this project.

†erikb@mit.edu, hui@wustl.edu, mengliu@mit.edu

1 Introduction

Artificial intelligence (AI) is one of the most important technological advances of our era. Recent progress of AI and, in particular, machine learning (ML), has dramatically increased predictive power in many areas such as speech recognition, image recognition, and credit scoring (Agrawal et al. [2016], Brynjolfsson and Mitchell [2017], Mullainathan and Spiess [2017]). Unlike the last generation of information technology that required humans to codify tasks explicitly, ML is designed to learn the patterns automatically from examples (Brynjolfsson and Mitchell [2017]). This has opened a broad new frontier of applications and economic implications that are, as yet, largely undeveloped. AI has been called a general-purpose technology, like the steam engine and electricity, whose capabilities span beyond specific applications. If this is true, then AI should ultimately lead to fundamental changes in work, trade and the economy.

Nonetheless, empirical evidence documenting concrete economic effects of using AI is largely lacking. In particular, contributions from AI have not been found in measures of aggregate productivity. Brynjolfsson et al. [2017] argue that the most plausible reason for the gap between expectations and statistics is due to lags in complementary innovations and business procedure reorganization. If the gap is indeed due to lagged complementary innovation, the best domains to empirically assess AI impacts are settings where AI applications can be seamlessly embedded in an existing production function because complementary innovations are already in place. In particular, various digital platforms are at the forefront of AI adoption, providing ideal opportunities for early assessments of AI's economic effects.

In this paper, we provide evidence of direct causal links between AI adoption and economic activities by analyzing the effect of the introduction of eBay Machine Translation (eMT) on eBay's international trade. As a platform, eBay mediated more than 14 billion dollars of global trade among more than 200 countries in 2014. The focal AI technology, eMT (from here on also referred to as the policy), is an in-house machine learning system that statistically learns how to translate among different languages. We exploit the discrete introduction of the policy for several language pairs, most notably English-Spanish, as a natural experiment, and study its consequences on U.S. exports on eBay via a difference-in-difference (DiD) estimation strategy. The identification compares the post-policy change in U.S. exports for the treated countries with that of the control countries (i.e., all other countries that U.S. sellers export to on eBay). For instance, we find that eMT increases U.S. exports to Spanish-speaking Latin American countries by 17.5%-20.9% on eBay, depending on the length of the pre- and post-policy time windows we evaluate. To mitigate potential spillover effects, we also use a second control group: offline U.S. exports to the same set of countries treated with eMT, for the DiD estimation. The results are similar, and the comparisons of the policy with either of the two control groups are statistically indistinguishable from each other. In the online appendix, we use U.S. exports to Brazil as a third control group, and also study the two rollouts of eMT in the EU. In each case, the results remain qualitatively unchanged.

Furthermore, we study heterogeneous treatment effects of the policy across different types of products and consumers. We find that the effect of eMT is more pronounced for:

(1) differentiated products,

(2) products with more words in listing titles,

(3) cheaper products, and

(4) less experienced buyers.

Each of these effects are consistent with a large reduction in translation-related consumer search costs. Products and buyers with higher search costs experience a greater benefit from eMT and therefore a larger increase in trade. In the online appendix, we provide a simple model of the effects of eMT and show the robustness of these heterogeneous treatment effects by (1) using the export value in dollars as the outcome variable, (2) using different estimation windows, and (3) repeating all analyses for the cases of EU and Russia. The results are qualitative very similar.

Our DiD results and heterogeneous effects, as well as our knowledge of the technology introduction and application itself, suggest a causal relationship between the introduction of machine translation and an export increase on eBay. More generally, our results may be a harbinger of more widespread effects of not only machine translation, but of related types of AI and ML. As these technologies are adopted and diffuse, we may see comparably large effects in other applications.

1.1 Related Literature and Contribution

1.1.1 Language Barriers in International Trade

Empirical studies using gravity models, as specified in Anderson and Van Wincoop [2003], have established the existence of a robust positive correlation between bilateral trade and reduced language barriers. Typically, researchers regress bilateral trade on a dummy variable for whether the two countries share the same language, and find that this coefficient is strongly positive (e.g., Melitz [2008], Egger and Lassmann [2012], and Melitz and Toubal [2014]). However, these cross-sectional regressions are vulnerable to endogeneity biases. For example, the fact that two countries share the same official or spoken language may be correlated with other shared characteristics or relationships that also affect trade, even after controlling for the usual set of variables in the gravity equation.

A key contribution of our paper, therefore, is that it exploits a natural experiment on eBay to identify the effect of changing language barriers on international trade. The online marketplace provides us with a uniquely-powerful laboratory to study the consequences on bilateral trade after an exogenous decrease in language barriers. Our finding that even a moderate quality upgrade of machine translation could increase export by 17% to 20% is consistent with Lohmann [2011] and Molnar [2013], who argue that language barriers may be far more trade-hindering than suggested by previous literature.

1.1.2 AI, Productivity, and Economic Welfare

The current generation of AI represents a revolution of prediction capabilities (e.g., Agrawal et al. [2016] and Mullainathan and Spiess [2017]). The recent exploding growth in prediction has been enabled by enormously increased data, significantly improved algorithms, and substantially more powerful computer hardware over the past few years (Brynjolfsson and McAfee [2017]).

There has a been a recent surge in interest in artificial intelligence, especially the subfield of machine learning, with significant increases in AI-related papers published, course enrollments, start-ups, start-up funding, and job openings, according to data collected by the AI Index. [1] These are in large part driven by recent breakthroughs in ML, especially supervised learning systems using deep neural networks, which have made possible substantial improvements in many technical capabilities. For instance, when benchmarked against a large data set of images (Imagenet), the best machine vision systems had an error rate of 28.5% in 2010 and now have an error rate of less than 2.5%, surpassing human error rates that are about 5% on the same data set. Similarly, the best speech recognition systems improved from over 15% error rates in 2011 to 5% error rates in 2017, and are now comparable to human error rates. Recently, machines have also surpassed humans at tasks as diverse as playing the game Go (Silver et al. [2016]) and recognizing cancer from medical images (Esteva et al. [2017]). There is active work converting these breakthroughs into practical applications such as self-driving cars, substitutes for human-powered call-centers, and new roles for radiologists and pathologists, but the complementary innovations required are often costly and time-consuming (Brynjolfsson et al. [2017]).

[1] http://cdn.aiindex.org/2017-report.pdf

Machine translation has also experienced significant improvement due to advances in machine learning. For instance, the best score at the Workshop on Machine Translation for translating English to German improved from 15.7 to 28.3 according to a widely-used comparison metric (the BLEU score).[2] Much of the recent progress in MT has been a shift from symbolic approaches towards statistical and deep neural network approaches. For our study, an important characteristic of eMT is that replacing human translators with MT or upgrading MT is typically relatively seamless. For instance, for product listings and descriptions on eBay, users simply consume the output of the translation system, but otherwise need not change their buying or selling process. While they care about the quality of the translation, it makes no difference whether the translation was produced by a human or machine. Thus, adoption of MT can be very fast and its economic effects, especially on digital platforms, can be seen quickly. While so far much of the work on the economic effects of AI has been theoretical (Acemoglu and Restrepo [2018], Aghion et al. [2017], Korinek and Stiglitz [2017], Sachs and Kotlikoff [2012]), and notably Goldfarb and Trefler [2018] in the case of global trade, the introduction of improved MT on eBay gives us an early opportunity to assess the economic effects of AI using a plausible natural experiment.

1.1.3 Peer-to-Peer Platforms and Matching Frictions

Einav et al. [2016] and Goldfarb and Tucker [2017] provide great surveys on how digital technology has reduced matching frictions and improved market efficiency. Reduced matching frictions affect price dispersion, as evidenced in Brynjolfsson and Smith [2000], Brown and Goolsbee [2002], Overby and Forman [2014], Ghose and Yao [2011], and Cavallo [2017]. These reduced frictions also mitigate geographic inequality in economic activities in the case of ride-sharing platforms (Lam and Liu [2017] and Liu et al. [2017]), short-term lodging platforms (Farronato and Fradkin [2018]), crowd-funding platforms (Catalini and Hui [2017]), and e-commerce platforms (Blum and Goldfarb [2006], Lendle et al. [2016], Fan et al. [2016], and Hui [2018]). We contribute to this literature by documenting the significant matching friction between consumers and sellers who speak different languages. Specifically, we find that efforts to remove language barriers provide substantial increases to market efficiency as well as platform profit.

[2] See reference at Euronext: http://matrix.statmt.org/matrix

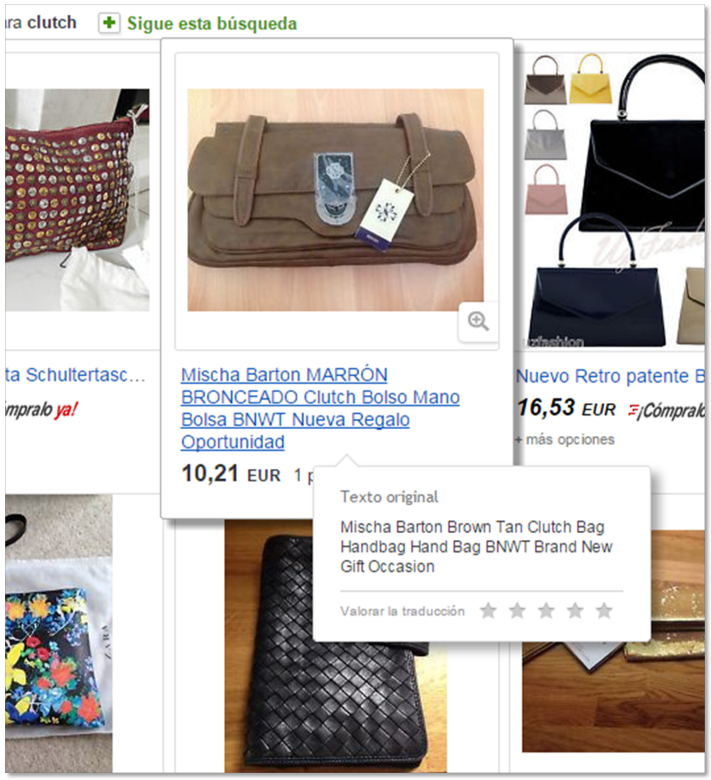


Figure 1: Example of eBay Machine Translation in Search Results Page

Notes: In this example, a Spanish buyer saw a listing from a UK seller on the search result page, and the item title is translated from English to Spanish by eMT.

2 eBay Machine Translation

The primary goal of eMT is to support international trade by making it easier for buyers to search for and understand the features of items that are not listed in their language. In a nutshell, eMT uses statistical models for phrase-to-phrase translations. These machine learning models are trained on both eBay data and other data automatically scraped from the Web to learn translation statistically. Some hand-crafted rules are applied, such as preserving named entities (e.g. numbers and product brands), so that eMT is more suited for the existing eBay environment. eBay also developed systems for post-editing of the outputs by human language experts, also known as machine-assisted human translation (MAHT), which further improved the translation quality. eMT is optimized to work in real-time, yielding high-quality translations within milliseconds.

In 2014, eBay rolled out eMT in three regions in different months: Russia (January), Latin America (May), and the European Union (July). In our main analyses we focus on the rollout in Latin America, because it is the first rollout of eMT that is not contaminated by major political events.[3] In the appendix, we also analyze the rollouts of eMT in the EU and Russia as robustness checks.

To shop on eBay, buyers in Latin America visit www.ebay.com, because there are no local eBay sites in these countries. eBay recognizes their IP addresses as being from Latin America, and the website is automatically localized by translating all its pages to buyers' local language. Note that this part was not affected by eMT, as the translations of the website pages, such as the translations of different product categories, buying formats, and advertisements, are fixed and not translated by eMT. Instead, eMT translates buyers' search queries and item titles. In particular, when buyers type in keywords in Spanish in the search box, eMT quickly translates that Spanish query into English and retrieves listings in the search results page based on the matching relevance between the translated query and listing titles in English. Given the set of listings in the search results page, the second task of eMT is to translate into and show these titles in Spanish. As a result, buyers from Latin America have a localized experience because it is as if they visited a website in Spanish, searched in Spanish, and the engine returned search results with Spanish titles. Figure 1 provides an example of eMT where an item title is translated from English into Spanish.[4]

Prior to eMT, eBay used Bing Translator for query search translation and item title translation. Therefore, the policy treatment here is an improvement in translation quality, which is measured by the human acceptance rate (HAR). eBay selected the top 500 most frequent queries, gave the translation results by eMT and by Bing to three experts in linguistics, and asked for their binary judgment of whether a translation is correct.[5] eBay then computed HAR values using majority votes (share of query translations that received at least two affirmative votes). Using this metric, the HAR for eMT was 91.4% while for Bing it was 84.4%. Thus, we consider eMT to be a moderate quality upgrade over Bing Translator.

3 Data and Empirical Strategy

This paper uses administrative data from eBay. The data include detailed listing attributes, product characteristics, buyer history, seller history, and reputation and feedback. We restrict the reporting

[3] The rollout in Russia was followed immediately by Russia's annexation of Crimea, which prompted international sanctions, and therefore changes in exports in that case could be confounded by political factors.

[4] The screenshot is an example of eMT between English and Spanish in the EU, which we found in an old internal document on eBay. Unfortunately, we could not find a screenshot for eMT in Latin America.

[5] For example, for the query "ropa femenina", eMT returned "women's clothing" while Bing returned "clothes female". Similarly, for the query "celulares", eMT returned "cell phones" while Bing returned "cellular". In both cases, eMT received unanimous affirmative votes and Bing received negative votes.

of summary statistics to comply with eBay's data policy. The offline data of monthly bilateral exports among countries come from the UN Comtrade Database.

To estimate the effect of eMT, we adopt the difference-in-difference (DiD) estimation in the following format:

$$log(Y_{ct}) = \beta T_c \times Post_t + XR_{ct} + \eta_c + \xi_t + \epsilon_{ct}, \tag{1}$$

where Y_{ct} is the export to country c at time t; T_c is the dummy for whether importing country c is in the treatment group (i.e., Spanish-speaking Latin American countries); $Post_t$ is the dummy for the introduction of eMT; XR_{ct} is the average daily bilateral exchange rate in month t; η_c are importing country fixed effects; and ξ_t are month fixed effects. The coefficient β represents the average treatment effect of eMT across all treated countries. Note that we only index importing countries c, because U.S. is the only exporter in our main analyses. Throughout the analyses, the standard errors are clustered at the country level to account for serial correlations of imports from the U.S. for each country.

The identification of the policy effect comes from comparing the intertemporal change in exports in the treatment group (countries that become eligible for eMT) against the baseline intertemporal change in exports in the control group (countries that remain ineligible for eMT). The DiD methodology allows us to control for two types of unobservables: (1) time-invariant country-specific trade propensities (e.g., U.S. exports to Canada are differ from exports to Peru) and (2) time-specific trade propensities that are the same across countries (e.g., exports are different in holiday seasons than in non-holiday seasons).[6]

Note that the DiD methodology does not control for serially-correlated unobservable errors. For the unbiasedness of the DiD estimator, we assume that these errors do not simultaneously correlate with both Y_{ct} and T_{ct}. In other words, eBay does not roll out eMT in countries with certain trade propensities. To indirectly test for this assumption, we plot the average monthly exports to the treated and non-treated countries in the 12 months before and after the introduction of eMT, and find that the parallel trend assumption is likely to hold. We return to this point in more detail in Section 4.1 and provide additional robustness checks when we assess a set of heterogeneous treatment effects. In addition, we follow Autor [2003] and perform leads–lags analyses, where the

[6] Note that our results are also robust to the inclusion of country-specific monthly trends. However, Borusyak and Jaravel [2016] recommend not to include unit-specific time trends in any difference-in-difference specifications because this exacerbates the bias of OLS for short-run impact (see Section 5.1.2. in their paper for details).

test results further strengthen the parallel trend assumption. Details of the test are provided in Appendix A.

The second identification assumption that we make is that the control group remained valid after the introduction of eMT. This assumption would be violated if U.S. exports have limited capacity at the aggregate level or, equivalently, the spillover effect of exports across countries is large. Imagine a scenario where the increase in U.S. exports to Mexico comes partially from a decrease in U.S. exports to China due to substitution. A comparison of exports to these two countries will over-estimate the policy effect because U.S. exports to China would have been higher had eMT not been introduced.

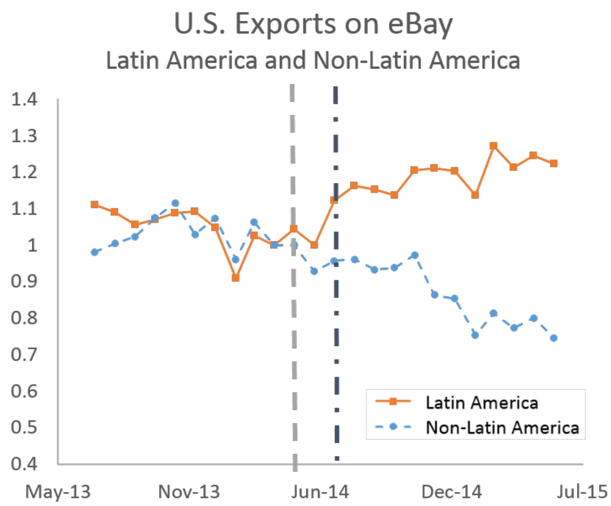
To mitigate this concern, we also use a different control group, which is the overall U.S. exports (online *and offline*) to the treated countries during the same period. Hui [2018] has estimated that eBay accounts for 1.38% of total U.S. exports in categories of products that are sold on eBay. Therefore, export on eBay is not large enough to alter the overall U.S. export pattern, making the second control group less subject to spillover effects.

4 Results

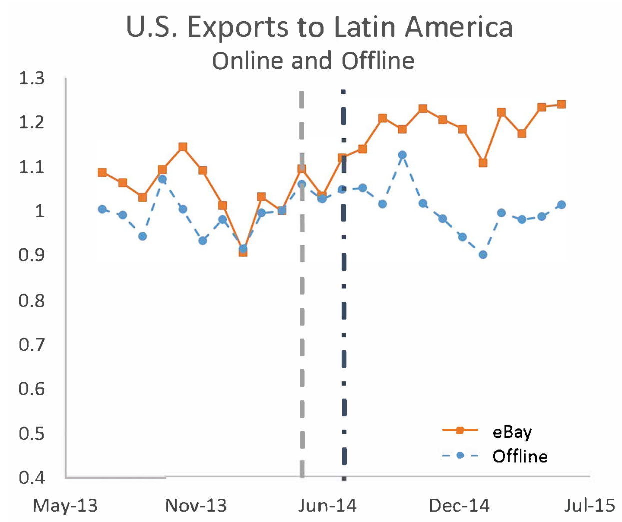
We first estimate the effect of eMT on U.S. exports to the treated countries. Next, motivated by a simple theoretical framework in Appendix A, we study how this effect differs along the following dimensions: (1) homogeneous versus differentiated products, (2) expensive versus inexpensive products, (3) listings with different numbers of words in the title, and (4) different levels of buyer experience on eBay.

4.1 Overall Policy Effect

Before we perform the DiD estimation, we plot average monthly U.S. exports on eBay for the treated and non-treated countries. In Figure 2a, we plot the normalized U.S. exports, measured in quantity, to Latin American countries and to other countries. The dashed and dot-dashed vertical lines refer to the introduction of query translation in May 2014 and item title translation in July 2014, respectively. Export quantities are normalized relative to the level in April 2013 (one month before query translation was introduced). Figure 2a suggests that the pre-trend assumption holds in the year before the policy change, and the policy promotes U.S. exports to Latin America in the year after its introduction.



(a)



(b)

Figure 2: Export Trends Diverge After Introduction of Machine Translation

Notes: Exports in figure 2a are measured in quantity and are normalized to the level in April 2013. Exports in figure 2b are measured in dollars and are normalized to the level in April 2013. The dashed and dot-dashed vertical lines refer to the introduction of query translation and item title translations, respectively.

To account for potential spillover effects, we plot a similar graph in Figure 2b but this time using offline U.S. exports to Latin America as a separate control group. Exports are measured only in U.S. dollars because data on offline export quantities are unavailable. We adopt the same normalization rule as seen in Figure 2a. Figure 2b suggests that the validity of the pre-trend assumption using this control group holds as well.

We apply equation (1) to estimate the policy effect. In Table 1, we estimate the equation using

Table 1: Overall Policy Effect

	Control Group 1				Control Group 2	
	log(Export Quantity)		log(Export Value)		log(Export Value)	
	(1)	(2)	(3)	(4)	(5)	(6)
	+/-6 months	+/-12 months	+/-6 months	+/-12 months	+/-6 months	+/-12 months
T*Post	0.175***	0.209***	0.131***	0.170***	0.118***	0.133***
	(0.007)	(0.007)	(0.020)	(0.017)	(0.031)	(0.025)
R^2	0.99	0.99	0.99	0.99	0.99	0.99

Notes: We control for country and month fixed effects, and monthly exchange rate according to specification (1). Standard errors are clustered at the country level.
*** indicates significance at p = 0.01; ** indicates p = 0.05.

data from +/- 6 months and +/- 12 months around the policy change for both control groups.[7] When we use the control group on eBay, our results show that the introduction of eMT increases U.S. exports on eBay to Latin America by 17.5%–20.9%.

In columns (3) and (4), we use the logarithm of export values in dollars as the dependent variable, and we see that the estimated effects are now 13.1% and 17%. While significantly different from zero, the effect on value is smaller than the effect on quantity. This reflects a small decrease in the average selling prices of U.S. exports (that we observe in the data), possibly due to higher competition among U.S. exporters. Our results suggest that consumers benefit from eMT more than sellers do, at least during the period we examine, because consumers gain both from reduced language frictions and also from lower prices.

Finally, in columns (5) and (6), we use the offline control group, and the estimated size of the policy effect is 11.8%–13.3%. These two estimates are not significantly different from those in columns (3) and (4), suggesting the control group of eBay U.S. exports to non-treated countries is valid.

One concern is that eBay may have advertised more to Latin American consumers after the introduction of eMT. To mitigate this concern, we study how the number of new registered buyers changed in Latin America relative to that in non-Latin American countries and Brazil. The results in the online appendix show no statically significant difference between the treatment and control groups, suggesting that the increased sales do not simply come from more intense advertising.

In the analyses that follow, we report the results on export quantity based on control group 1 and data from +/-6 months around the policy change. We focus on changes in export quantity to purge away eMT's effect on price, because we are mainly interested in its effect on exporting activities. Additionally, using control group 1 allows us to exploit eBay's rich data to understand the heterogeneous effects of the policy. Lastly, using a narrower window reduces other contemporaneous factors that potentially contaminate the estimates. In the online appendix, we also report the estimation results based on export revenue and different window lengths.

4.2 Different Products

Since eMT reduces translation-related search costs, Proposition 1 in Appendix A states that the increase in exports should be proportionately larger for products that had larger search costs before the policy change than for other products. We divide products into two categories: homogeneous

[7] Since U.S. exports to different countries are never zero, the logarithms of exports are always defined.

Table 2: Heterogeneous Effects by Product Type

By Homogeneity of Products		By Degree of Catalogization		By Product Value	
	log(Q)		log(Q)		log(Q)
T*Post	0.187***	T*Post	0.174***	T*Post	0.193***
	(0.016)		(0.015)		(0.015)
T*Post*Homog	-0.062***	T*Post*Q2	-0.017	T*Post*Value[10,50):	-0.019
	(0.022)		(0.02)		(0.022)
		T*Post*Q3	-0.048***	T*Post*Value[50,200)	-0.056***
			(0.02)		(0.022)
		T*Post*Q4	-0.111***	T*Post*Value≥200	-0.059***
			(0.02)		(0.022)
R^2	0.99	R^2	0.96	R^2	0.99

Notes: We use data from six months before and after the introduction of the eMT for estimation. We control for all variables and fixed effects according to specification (1), as well as product category fixed effects. Standard errors clustered at the country level.

*** indicates significance at p = 0.01; ** indicates p = 0.05.

products (e.g., cellphones and books, which are mass produced and have standard identifiers) and differentiated products (e.g. antiques and clothing which have more variation in product attributes). Translation-related search costs should be higher for differentiated products because of higher language requirements (and hence higher translation costs) of translating the specifics of these products into local languages.

We distinguish homogeneous products from differentiated products on eBay by identifying whether a product is assigned a "Product ID" on eBay. These Product IDs are the most fine-grained catalogs on eBay and are defined mainly for homogeneous products. For instance, an "Apple iPhone 8–256 GB–Space Gray–AT&T-GSM" has a unique Product ID that is different from that of an iPhone of a different generation, a different color, internal memory, or carrier. For books or CDs, these Product IDs are ISBN codes. On the other hand, Product IDs are rarely defined for products such as fashion products, clothing, art, and jewelry, because these products have many variations and are often unique.

In the left panel of Table 2, we repeat our DiD regression for the two types of products using exports that are aggregated at the country–product type–month level. We control for product type fixed effects in addition to the controls in specification (1). The results show that, consistent with the theory prediction, the export increase for homogeneous products is smaller than that for differentiated products, and the 6.2% difference is statistically significant.

To further test this heterogeneity, we estimate the policy effect for each of the 36 meta-categories, a mega-category being the highest-level catalog inclusive of all items listed on eBay. The estimates

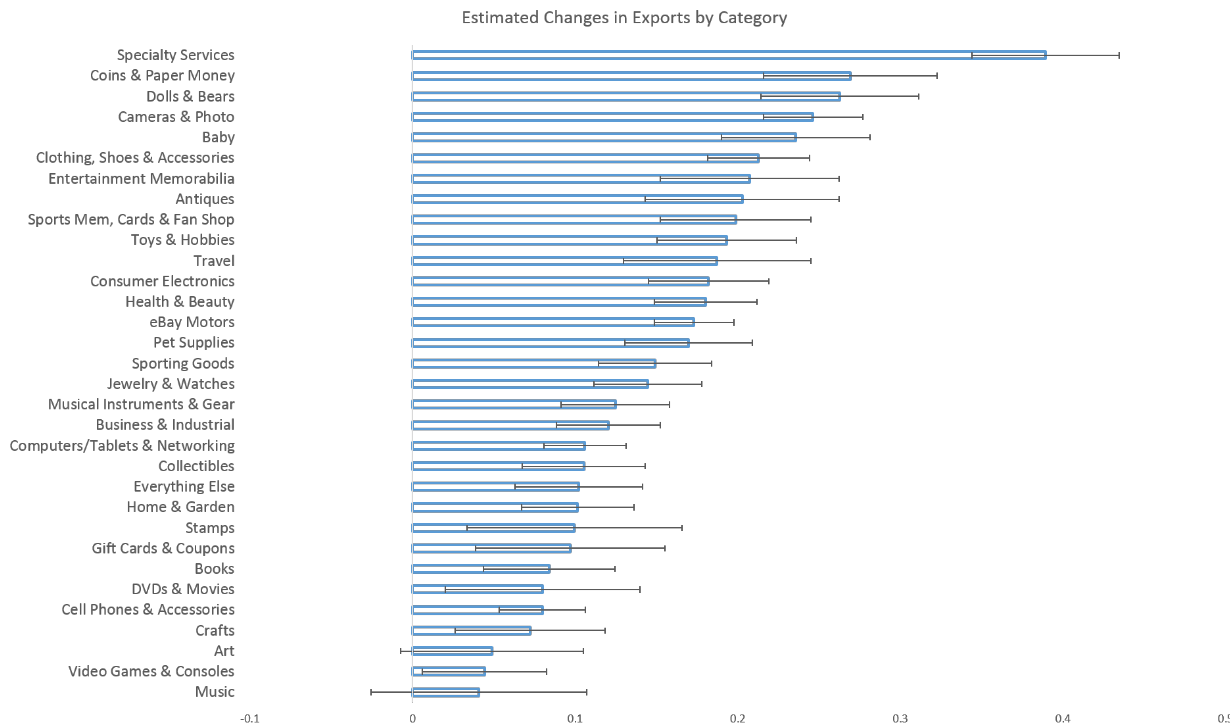


Figure 3: Export Increase by Categories

Notes: We use data from six months before and after the introduction of the eMT to estimate the policy effect on exports for each meta-category on eBay according to specification (1). The rectangles indicate estimated coefficients and the bars represent 95% confidence interval of these estimates.

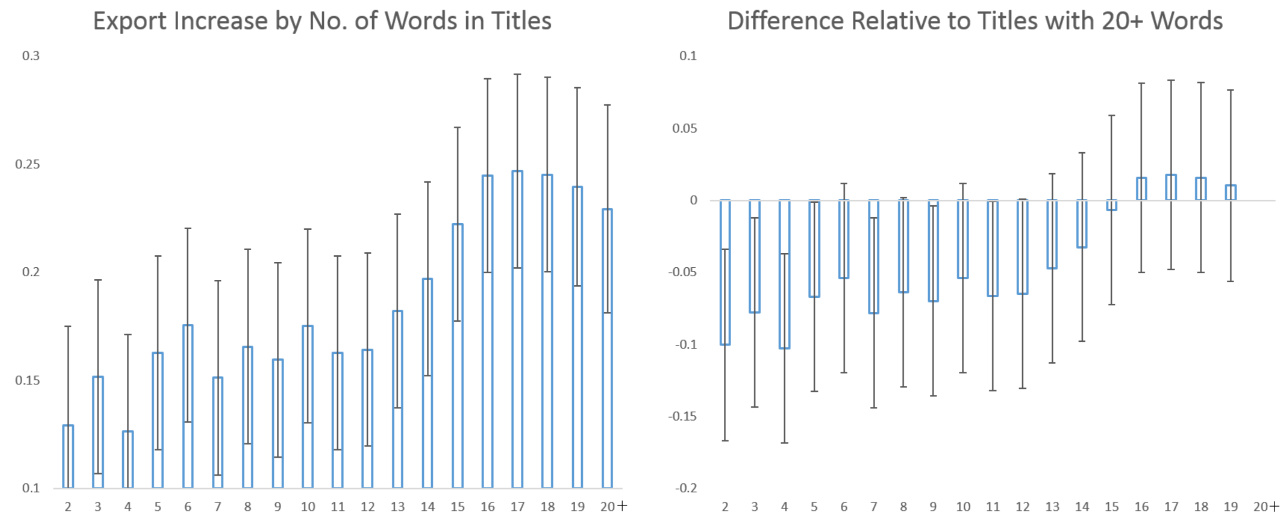


Figure 4: Export Increase by Number of Words in Listing Titles

Notes: We use data from six months before and after the introduction of the eMT to estimate the policy effect on exports for listings with different number of words in the titles (split regressions for different word counts in the left graph and a pooled regression with interaction terms with dummies for different word counts in the right graph. The rectangles indicate estimated coefficients and the bars represent 95% confidence interval of these estimates. The x-axis is capped at '20+' which includes listings with 20 or more words in the titles.

are represented using horizontal bars in Figure 3, and the intervals are the 95% confidence interval around the estimates. A visual inspection of the results is consistent with our theory that the export increase is larger for categories that have more variation in product attributes (e.g., Specialty Services, Coins & Paper Money, Dolls & Bears).

To formalize these observations in regressions, we label meta-categories by their degree of "catalogization", which is measured as the share of listed items with Product IDs on www.ebay.com in 2014. For example, suppose there are 100 listings in both Video Games & Consoles and Coins & Paper Money categories; in Video Games & Consoles 50 of these listings have been assigned Product IDs, while in Coins & Paper Money only 5 listings have Product IDs. Therefore, the degree of catalogization is 50% and 5%, respectively. We then divide all the meta-categories into four quartiles based on this measure, and interact these dummies with the treatment dummies. The estimates in the middle panel of Table 2 show that export increase decreases with the degree of catalogization. In particular, the export increase of the bottom quartile of catalogization (i.e., the most differentiated products) is 11.1% greater than that of the top quartile (i.e., the most homogeneous products).

Next, we hypothesize that the export increase is larger for listings with more words in their

Table 3: Heterogeneous Effects by Buyer Experience

Dependent Variable: log(Export Quantity)			
	(1)	(2)	(3)
	All Products	Homogeneous Products	Differentiated Products
T*Post	0.176***	0.202***	0.194***
	(0.014)	(0.014)	(0.026)
T*Post*Experienced	-0.061***	-0.084***	-0.123***
	(0.019)	(0.020)	(0.037)
R^2	0.99	0.99	0.97

Notes: We use data from six months before and after the introduction of the eMT for estimation. We control for all variables and fixed effects according to specification (1), as well as the standalone dummy variable 'Experienced'. Standard errors are clustered at the country level.
*** indicates significance at p = 0.01.

item titles, because translation-related search costs should generally increase with the number of words. Figure 4 plots the policy effects for listings with a different number of words in listing titles. On the right side, we compare the policy effect relative to that of listings with at least 20 words in their titles. We see that items with fewer than 13 words in their titles experienced 6%–10% smaller export growth, and that change is statistically significant at the 5% level for most cases.

Lastly, we explore how the policy effect differs by product value. Proposition 2 in Appendix A predicts a larger increase in export for cheaper items, because the search cost as a fraction of item value is higher for these items. To test this hypothesis, we divide items into four value bins: $[0, \$10)$, $[\$10, \$50)$, $[\$50, \$200)$, and \$200 and above. Following Einav et al. [2015] and Hui et al. [2017], product value is defined as the average Buy It Now price in the 6 months preceding the policy change. In the right panel of Table 2, we see that products with \$50 or less value have experienced a 19.3% increase in exports, and the increase is about 5.6% smaller for more expensive items than it is for less expensive items.

4.3 Different Buyer Experience

We have established that the increase in export is higher for products associated with relatively higher translation-related search costs, whether proxied by degree of standardization, word count or price. According to Proposition 1 in Appendix A, this result should also hold for *buyers* with a higher search cost. We do not directly observe buyers' search costs, but according to Hui et al. [2016], buyers' experience on eBay is a good proxy for search cost. In particular, buyers with low search costs are generally deal-seekers and spend more time and money on eBay relative to inexperienced buyers. Note that lower search costs, in the context of language barriers, could mean that these buyers have better knowledge of English or have used other translation tools and services.

We define experienced buyers as those who have purchased at least $2,500 worth of items in the previous year, which roughly corresponds to eBay's definition. In column (1) of Table 3, we see that the increase in export is 6.1% larger for inexperienced buyers, consistent with them having higher translation-related search costs than experienced buyers do.

We can also look at the combined effects of buyer experience and product differentiation. In columns (2) and (3), we find that the differential response of experienced buyers persists for both homogeneous and differentiated products, measured by whether items are assigned a Product ID.[8] The persistence of this gap suggests that translation-related search costs vary by buyer types, which is a new margin distinct from product types.

5 Conclusion

In this paper, we exploit a natural experiment on eBay to study the effect of an AI-based machine translation tool on international trade. We show that a moderate quality upgrade increases exports on eBay by 17.5%. The increase in exports is larger for differentiated products, cheaper products, listings with more words in their title. Machine translation also causes a greater increase in exports to less experienced buyers. These heterogeneous treatment effects are consistent with a reduction in translation-related search costs, which comes from two sources: (1) an increased matching relevance due to improved accuracy of the search query translation and (2) better translation quality of the listing title in buyers' language.

Our results have two main implications.

First, language barriers have greatly hindered trade. This is true even for digital platforms where trade frictions are already smaller than they are offline. In our study, the quality upgrade in machine translation is moderate–a 7% increase in human acceptance rate–yet this upgrade generated an export increase of 17.5% in quantity and 13.1% in revenue. To put our result in context, Hui [2018] has estimated that a removal of export administrative and logistic costs increased export revenue on eBay by 12.3% in 2013, which is similar to the effect of eMT. Additionally, Lendle et al. [2016] have estimated that a 10% reduction in distance would increase trade revenue by 3.51% on eBay. This

[8]Note that the effect for inexperienced buyers for all products is not exactly a weighted average of the effects for homogeneous and differentiated products, because the country and month fixed effects have different meanings across the three columns. In order to get an exact weighted average, one would control for a "Homogeneous" dummy and its interaction with all existing variables in column (1).

means that the introduction of eMT is equivalent of the export increase from reducing distances between countries by 37.3%. These comparisons suggest that the trade-hindering effect of language barriers is of first-order importance. Machine translation has made the world significantly smaller and more connected.

Second, AI has great potential to increase economic activity and productivity. In November 2016, Google announced its neural machine translation (NMT) system based on deep learning could reduce translation errors by 55%–85% compared to the previous generation of Google Translate. The estimates in this paper suggest that the effect of NMT on cross-border trade could be large if NMT passes some quality thresholds for different languages and purposes. Besides machine translation, AI applications are also flourishing in other fields such as speech recognition, computer vision, and recommendation systems, as well as an increasing number of applications such as hiring decisions, medical diagnoses, and self-driving vehicles. There will undoubtedly be new opportunities to assess the economic impact of these technologies via natural experiments like the one we examined in this paper, providing further opportunities to assess the economic contributions of AI.

References

Daron Acemoglu and Pascual Restrepo. Artificial intelligence, automation and work. Technical report, National Bureau of Economic Research, 2018.

Philippe Aghion, Benjamin F Jones, and Charles I Jones. Artificial intelligence and economic growth. Technical report, National Bureau of Economic Research, 2017.

Ajay Agrawal, J Gans, and Avi Goldfarb. Exploring the impact of artificial intelligence: Prediction versus judgment. *University of Toronto*, 2016.

James E Anderson and Eric Van Wincoop. Gravity with gravitas: A solution to the border puzzle. *The American Economic Review*, 93(1):170–192, 2003.

David H Autor. Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics*, 21(1):1–42, 2003.

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

Kirill Borusyak and Xavier Jaravel. Revisiting event study designs. 2016.

Jeffrey R Brown and Austan Goolsbee. Does the internet make markets more competitive? evidence from the life insurance industry. *Journal of Political Economy*, 110(3):481–507, 2002.

Erik Brynjolfsson and Andrew McAfee. Whats driving the machine learning explosion? *Harvard Business Review*, pages 3–11, 2017.

Erik Brynjolfsson and Tom Mitchell. What can machine learning do? workforce implications. *Science*, 358(6370):1530–1534, 2017.

Erik Brynjolfsson and Michael D Smith. Frictionless commerce? a comparison of internet and conventional retailers. *Management Science*, 46(4):563–585, 2000.

Erik Brynjolfsson, Daniel Rock, and Chad Syverson. Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. Technical report, National Bureau of Economic Research, 2017.

Christian Catalini and Xiang Hui. Can capital defy the law of gravity? investor networks and startup investment. 2017.

Alberto Cavallo. Are online and offline prices similar? evidence from large multi-channel retailers. *American Economic Review*, 107(1):283–303, 2017.

Peter H Egger and Andrea Lassmann. The language effect in international trade: A meta-analysis. *Economics Letters*, 116(2):221–224, 2012.

Liran Einav, Theresa Kuchler, Jonathan Levin, and Neel Sundaresan. Assessing sale strategies in online markets using matched listings. *American Economic Journal: Microeconomics*, 7(2): 215–247, 2015.

Liran Einav, Chiara Farronato, and Jonathan Levin. Peer-to-peer markets. *Annual Review of Economics*, 8:615–635, 2016.

Andre Esteva, Brett Kuprel, Roberto A Novoa, Justin Ko, Susan M Swetter, Helen M Blau, and Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115, 2017.

Jingting Fan, Lixin Tang, Weiming Zhu, and Ben Zou. The alibaba effect: Spatial consumption inequality and the welfare gains from e-commerce. 2016.

Chiara Farronato and Andrey Fradkin. The welfare effects of peer entry in the accommodation market: The case of airbnb. Technical report, National Bureau of Economic Research, 2018.

Anindya Ghose and Yuliang Yao. Using transaction prices to re-examine price dispersion in electronic markets. *Information Systems Research*, 22(2):269–288, 2011.

Avi Goldfarb and Daniel Trefler. Ai and international trade. Technical report, National Bureau of Economic Research, 2018.

Avi Goldfarb and Catherine Tucker. Digital economics. Technical report, National Bureau of Economic Research, 2017.

Xiang Hui. Facilitating inclusive global trade: Evidence from a field experiment. 2018.

Xiang Hui, Maryam Saeedi, Zeqian Shen, and Neel Sundaresan. Reputation and regulations: Evidence from ebay. *Management Science*, 2016.

Xiang Hui, Maryam Saeedi, Giancarlo Spagnolo, and Steve Tadelis. Certification, reputation and entry: An empirical analysis. *Unpublished Manuscript*, 2017.

Anton Korinek and Joseph E Stiglitz. Artificial intelligence and its implications for income distribution and unemployment. Technical report, National Bureau of Economic Research, 2017.

Chungsang Tom Lam and Meng Liu. Demand and consumer surplus in the on-demand economy: the case of ride sharing. 2017.

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

Meng Liu, Erik Brynjolfsson, and Jason Dowlatabadi. Technology, incentives, and service quality: the case of taxis and uber. 2017.

Johannes Lohmann. Do language barriers affect trade? *Economics Letters*, 110(2):159–162, 2011.

Jacques Melitz. Language and foreign trade. *European Economic Review*, 52(4):667–699, 2008.

Jacques Melitz and Farid Toubal. Native language, spoken language, translation and trade. *Journal of International Economics*, 93(2):351–363, 2014.

Alejandro Molnar. Language barriers to foreign trade: evidence from translation costs. *Nashville: Vanderbilt University*, 2013.

Sendhil Mullainathan and Jann Spiess. Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106, 2017.

Eric Overby and Chris Forman. The effect of electronic commerce on geographic purchasing patterns and price dispersion. *Management Science*, 61(2):431–453, 2014.

Jeffrey D Sachs and Laurence J Kotlikoff. Smart machines and long-term misery. Technical report, National Bureau of Economic Research, 2012.

David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

A. A Simple Theoretical Framework

We provide a simple Nash bargaining framework between buyers and foreign sellers to illustrate the role of eMT. For simplicity, we model the international market to be independent of the domestic market, so that buyers' purchase from foreign sellers is independent of that from domestic sellers. On the demand side, there is a unit-measure of buyers whose highest willingness to pay for product j is w_j, where $w_j > 0$. Buyers differ in translation-related search cost s_i, which follows a continuous distribution F_i. s_i captures different search costs across buyers; for example, buyers differ in their language skills and it is more costly for some buyers to translate and understand titles listed in foreign language.

On the supply side, there is a unit-measure of foreign sellers, and each seller sells a different product. They all have zero marginal cost of production, i.e., $c_j = 0$. Products differ in their search costs, s_j, which follows a continuous distribution F_j. s_j represents different translation-related costs across products. Let buyers' Nash bargaining power be α, where $\alpha \in (0, 1)$ is constant across all products. Let the market competition be such that buyers capture α of their highest willingness to pay, i.e., $p_j = (1 - \alpha)w_j$. Under this model setup, the number of foreign transactions for product j equals the share of buyers who choose to purchase from the foreign seller, or $Pr(w_j - p_j - s_i - s_j > 0)$. Given that $w_j = \frac{p_j}{1-\alpha}$, the number of foreign transactions can be written as $Pr(s_i + s_j < \frac{\alpha p_j}{1-\alpha})$. Note that although we write number of transactions as a function of p_j (for direct comparative statics in Proposition 3), this is equivalent with writing number of transactions as a function of the model primitive w_j due to the deterministic relationship that $p_j = (1 - \alpha)w_j$.

We assume eMT reduces both the product and buyer margins of translation-related search cost to 0 for simplicity. In this case, all buyers buy from all foreign sellers because $w_j > p_j > 0$. Then the increase in number of international transaction is $1 - Pr(s_i + s_j < \frac{\alpha p_j}{1-\alpha})$. Given this, products and buyers with higher search costs should expect a larger export response when search costs are removed. Conversely, products and buyers with low search costs should not expect as much of an effect of eMT, because international trade in these categories were high even before the introduction of the eMT. Propositions 1 and 2 formalize these arguments for different products and different buyer types, holding product value constant.

Proposition 1 *Suppose there are two products with search costs s_j^1 and s_j^2. All else equal, the increase in exports due to eMT is higher for the product with search cost s_j^1 if $s_j^1 > s_j^2$.*

Proof. Recall that the number of foreign transactions for a particular product is $Pr(s_i < \frac{\alpha p_j}{1-\alpha} - s_j)$, which can be written as $F_i(\frac{\alpha p_j}{1-\alpha} - s_j)$. Hence the policy effect on export is $1 - F_i(\frac{\alpha p_j}{1-\alpha} - s_j)$. Since $s_j^1 > s_j^2$, we have that $1 - F_i(\frac{\alpha p_j^1}{1-\alpha} - s_j^1) \geq 1 - F_i(\frac{\alpha p_j^2}{1-\alpha} - s_j^2)$, if p_j^1 and p_j^2 are similar. ■

Proposition 2 *Suppose there are two buyers with search costs s_i^1 and s_i^2. All else equal, the increase in exports due to eMT is higher for the buyer with search cost s_i^1 if $s_i^1 > s_i^2$.*

Proof. Recall that the number of foreign transactions from a particular buyer is $Pr(s_j < \frac{\alpha p_j}{1-\alpha} - s_i)$, which can be written as $F_j(\frac{\alpha p_j}{1-\alpha} - s_i)$. Hence the policy effect on export is $1 - F_j(\frac{\alpha p_j}{1-\alpha} - s_i)$. Since $s_i^1 > s_i^2$, we have that $1 - F_j(\frac{\alpha p_j^1}{1-\alpha} - s_i^1) \geq 1 - F_j(\frac{\alpha p_j^2}{1-\alpha} - s_i^2)$, if p_j^1 and p_j^2 are similar. ■

Having established the relationship between policy effect and search costs by holding product price/value constant, we now study the relationship between policy effect and product price/value by holding search costs constant. Based on the Nash bargaining framework, the consumer surplus is given by $w_j - p_j - s_i - s_j = \frac{\alpha}{1-\alpha} p_j - s_i - s_j$. As a result, consumer surplus of the cheaper product is lower, and therefore buyers were less likely to purchase cheaper product from foreign sellers before the introduction of the eMT. Therefore, the effect of removing translation-related search cost s on foreign sales should be greater for the cheaper product than for the more expensive product. The next proposition illustrates how the eMT effect differs between cheap and expensive products:

Proposition 3 *Suppose there are two products with price p_1 and p_2, where $p_1 < p_2$. All else equal, the increase in exports due to eMT is higher for the product category with p_1.*

Proof. Recall the probability that consumer i buys product j is $Pr(s_i + s_j < \frac{\alpha p_j}{1-\alpha})$. Denote the distribution of $(s_i + s_j)$ as H (a function of the underlying distributions F and G). The policy effect on exports is $1 - H(\frac{\alpha p_j}{1-\alpha})$. Since $p_1 < p_2$, it follows that $1 - H(\frac{\alpha p_1}{1-\alpha}) \geq 1 - H(\frac{\alpha p_2}{1-\alpha})$, if s_j^1 and s_j^2 are similar. ■

This above theoretical framework and propositions lead to a set of testable implications:

- The export increase due to eMT is larger for products with higher translation-related search costs. In particular,
  - Export increase is larger for differentiated products, because translating the specifics of these products requires higher language skills and hence higher translation costs.
  - Export increase is larger for products with more words in the listing titles, because translation costs should increase in number of words.

- The export increase due to eMT is larger for buyers with higher translation-related search costs. In particular,
  - Export increase is larger for inexperienced buyers on eBay, because these buyers spend less time searching on eBay and likely have higher search costs.
- The export increase due to eMT is larger for cheaper products.

B. Leads–Lags Analyses

Similar to Autor [2003], we perform leads–lags analyses to test the parallel trend assumption:

$$log(Y_{ct}) = \sum_{l=-12}^{11} \beta_l T_c \times Post_t(t = k + l) + XR_{ct} + \eta_c + \xi_t + \epsilon_{ct}, \tag{2}$$

where k is the policy introduction month, and β_l is the coefficient on the l-th lead or lag. One test is $\beta_l = 0$ for $l < 0$, i.e., there should not be pre-treatment differences between the treatment and control groups. Moreover, β_l for $l \geq 0$ might be different if there is temporal variation in the treatment effect. The data used are the same as in Table 1. We see from Table 4 that the parallel trend assumptions largely hold for the two control groups except for the seventh and eighth lags for the offline control. Moreover, the policy effects on U.S. export augment and become more significant after the third month from the introduction of eMT.[9]

Table 4: Leads–Lags Analyses

	Online	t-statistics	Offline	t-statistics
Pre-Treatment (t-11)	-0.027	-0.566	-0.034	-0.659
Pre-Treatment (t-10)	0.089*	-1.837	-0.036	-0.695
Pre-Treatment (t-9)	0.055	-1.139	0.023	0.457
Pre-Treatment (t-8)	0.011	0.233	0.133***	2.605
Pre-Treatment (t-7)	0.012	0.254	0.104**	2.027
Pre-Treatment (t-6)	-0.023	-0.476	0.032	0.634
Pre-Treatment (t-5)	-0.025	-0.526	-0.021	-0.414
Pre-Treatment (t-4)	-0.017	-0.362	0.027	0.519
Pre-Treatment (t-3)	-0.013	-0.262	0.004	0.085
Pre-Treatment (t-2)	-0.016	-0.334	0.012	0.232
Pre-Treatment (t-1)	0.023	0.470	0.007	0.139
Post-Treatment (t)	0.056	1.170	0.082	1.610
Post-Treatment (t+1)	0.051	1.063	0.123**	2.405
Post-Treatment (t+2)	0.050	1.037	0.084	1.634
Post-Treatment (t+3)	0.047	0.974	0.114**	2.237
Post-Treatment (t+4)	0.149***	3.102	0.303***	5.927
Post-Treatment (t+5)	0.171***	3.552	0.277***	5.407
Post-Treatment (t+6)	0.160***	3.328	0.241***	4.703
Post-Treatment (t+7)	0.147***	3.055	0.134***	2.627
Post-Treatment (t+8)	0.224***	4.642	0.249***	4.864
Post-Treatment (t+9)	0.251***	5.202	0.184***	3.603
Post-Treatment (t+10)	0.233***	4.843	0.191***	3.726
Post-Treatment (t+11)	0.282***	5.848	0.173***	3.377

Notes: For estimation, we use data from six months before and after the policy change using specification (2). Standard errors are clustered at the country level.

[9] This is due to the supply side, namely that sellers did not increase their international listings, especially on differentiated products, until after the first three months from the introduction of eMT.

Online Appendix
Does Machine Translation Affect International Trade? Evidence from a Large Digital Platform

Erik Brynjolfsson
MIT and NBER

Xiang Hui
Washington University in St. Louis and MIT

Meng Liu*
Washington University in St. Louis and MIT

August 1, 2018

We test for the robustness of our estimates on the impact of eBay Machine Translation (eMT) on U.S. exports to Latin America on eBay. First, we use U.S. exports to Brazil, where the main language is Portuguese, as our control group for U.S. exports to Spanish-speaking Latin American countries. We could not use U.S. exports to Guyana, Suriname, and French Guiana as our control group because U.S. exports to these countries on eBay are too small. The estimates in Table A1 are consistent with the trade-promoting effect of eMT and show a 20.3% increase in export quantity from the U.S. to Spanish-speaking Latin American countries in the six months after the eMT rollout.

One concern is that the increase in U.S. exports to Latin American countries could be driven by eBay's increased advertisement spending in those countries. To deal with this, we use the same DiD specification to study how the number of new registered buyers changed in Latin American countries relative to non-Latin American countries and Brazil. The results in Table A2 show no statistically significant difference between the treatment and control groups, suggesting that the larger effect is not merely coming from more intense advertising.

In the main analyses, we use the logarithm of export quantity as the dependent variable, and we use data from 6 months before and after the policy change for the DiD estimation. In

*erikb@mit.edu, hui@wustl.edu, mengliu@mit.edu

the first set of robustness checks, we replicate our results using the logarithm of export value in dollars as the outcome variable, as well as extending the estimation window to 12 months before and after the policy change. In Table A3, we see that our previous heterogeneous treatments remain qualitatively the same: The export increase is larger for differentiated goods, in categories with a larger share of products catalogued, and for cheaper products. Even though a few specifications are no longer statistically significant, the sign and magnitude of the estimates are consistent with lower search costs from eMT. In Figure A1, we repeat our analyses based on the number of words in listing titles using the logarithm of export value in dollars as the outcome variable, and the results remain qualitatively the same. Lastly, in Table A4, we see again that the export increase is larger for inexperienced buyers on the platform, consistent with the search cost story.

Subsequently, we test the robustness of our results using the other two rollouts of eMT on eBay. In July 2014, eBay rolled out eMT of item titles and search queries in the European Union to promote intra-EU trade (i.e., items listed on the U.S. sites are not translated). In particular, item titles in English were translated to French, Italian, and Spanish, to facilitate British and Irish exports to France, Italy, and Spain (FRITES). Additionally, in January 2014, eBay had a rollout of eMT that translated English into Russian to facilitate exports from English-speaking countries to Russia. However, the effects of this rollout were contaminated by Russia's annexation of Crimea, which prompted international sanctions. Although the estimates of total changes in exports in the case of Russia could be confounded by political factors, the *relative* changes in exports for different types of products and buyer experiences could still be informative of the mechanism of eMT if the political factors affected exports in a similar way across the different groups.

In the case of FRITES, we use UK and Irish exports to non-FRITES countries as the control group for their exports to FRITES countries on eBay. We first plot the average monthly British and Irish exports on eBay to the FRITES and non-FRITES countries. In Figure A2a, exports on eBay are measured in quantity and are normalized relative to the level in June 2014. The dashed vertical lines refer to the introduction of query translation and item title translation in July 2014. The figure suggests that the pre-trend assumption holds in the year before the policy change, and the policy promotes British and Irish exports to

FRITES in the year after its introduction. Similarly, in Figure A2b we use offine British and Irish exports to FRITES as a second control group. Exports are measured in U.S. dollars. The figure suggests that the pre-trend assumption holds for the second control group, and the policy seems effective in promoting trade.

In the case of Russia, we use exports to Russia from non-English speaking countries as the first control group for exports to Russia from English-speaking countries. For the second control group, we use offline exports to Russia from English-speaking countries. We plot the graphs using the two different control groups in Figure A2c and A2d. Figure A2c suggests that the parallel trend assumption holds and that there is a positive effect of the policy on exports to Russia from English-speaking countries on eBay, relative to the null effect of exports to Russia from non-English-speaking countries on eBay. However, a visual inspection of Figure A2d suggests that the second control group is not valid, because we see a kink in exports to Russia on eBay around summer 2013, which is non-existent in offline trade.

In Table A5, we use specification (1) to estimate the policy effect on exports to FRITES and Russia. In Panel A, we control for both fixed effects for seller countries (U.K. and Ireland) and buyer countries. When we use the first control group, we estimate that the introduction of eMT increased U.K. exports to FRITES on eBay by 13.9% and 13.6% in the 6 months and 12 months after the policy change, respectively. When we use offine exports as the control group, we find the estimated policy effects are 18.7% and 11.9%, depending on the time window. We have checked that the difference in the policy effect in dollar terms on eBay and offline is not statistically significant. Therefore, the rollout of eMT in Europe confirms the finding that machine translation increases exports on eBay.

In the case of Russia, when we use the first control group (online exports to Russia from non-English speaking countries on eBay), we see a 17.1% increase in exports to Russia on eBay in the 6 months after the policy change (Panel B), and this number shrinks to 12% when we look at a 12-month window. However, the estimates become statistically insignificant when we use the offline control group. We remain skeptical about our ability to estimate the overall effect of the policy on exports to Russia, because of the inconsistency between our estimates when we use the online and offline control groups.

In Table A6, we study the heterogeneous treatment effects of the impact of eMT in the cases of FRITES and Russia. In the left section of both Panel A and Panel B, we see significant negative coefficient estimates of the triple interaction terms "T*Post*Homogeneous", indicating that the export change is more positive for differentiated products.[1] In the middle section of both panels, we see that markets (i.e., meta-categories on eBay) where the degree of catalogization is in the bottom quartile (i.e., less homogeneous) experienced a positive change in exports, whereas markets on the other end of the catalogization spectrum experienced negative changes relative to the first group. Lastly, in terms of product values, we see in the right section of both panels that export increases are positive for items that are sold for less than $10, and that the increase is smaller or non-existent as the average sales price of a product increases. All of these heterogeneities are consistent with lower search costs and similar to our main analyses in the case of Latin America.

Subsequently, we study how the effect of eMT varies with the number of words in listing titles. In Figure A3, we see that the export change generally increases with the number of words in listing titles. This is consistent with the fact that translation costs increase with the number of words, and therefore the reduction in translation costs is larger for items with longer titles.

Lastly, we study the policy effect on buyers with different experience on eBay. The definition of experienced buyers is the same as before. In Table A7, we see that the estimated export increases are smaller for experienced buyers. This is similar to our finding in the case of Latin America, and is consistent with the fact that, relative to experienced buyers, inexperienced buyers have higher translation-related search costs. We find no statistically significant difference of this gap for homogeneous and differentiated products, suggesting that buyer experience is a different margin of heterogeneity than product types.

[1] In the case of Russia, a negative change in exports for homogeneous products could be a result of political factors, but the sign of this triple interaction term is the same as in the case of Latin American and FRITES.

Table A1: Overall Policy Effect using Brazil as Control

	log(Export Quantity)		log(Export Value)	
	(1)	(2)	(3)	(4)
	+/-6 mo.	+/-12 mo.	+/-6 mo.	+/-12 mo.
T*Post	0.203***	0.553***	0.146*	0.411***
	(0.079)	(0.073)	(0.086)	(0.074)
R^2	0.99	0.99	0.99	0.98

Notes: We control for fixed effects according to specification (1). Standard errors clustered at the country level.

*** indicates significance at p = 0.01; ** indicates p = 0.05.

Table A2: Policy Effect on New Buyer Registration in Latin America

	Dependent Variable:log(Number of New Buyer Registration)			
	Control: Non-LatAm Countries		log(Control: Brazil)	
	(1)	(2)	(3)	(4)
	+/-6 mo.	+/-12 mo.	+/-6 mo.	+/-12 mo.
T*Post	-0.273	-0.434	-0.334	-0.222
	(0.264)	(0.763)	(0.396)	(0.277)
R^2	0.34	0.10	0.83	0.82

Notes: We control for fixed effects according to specification (1). Standard errors clustered at the country level.

*** indicates significance at p = 0.01; ** indicates p = 0.05.

Table A3: Heterogeneous Effects by Product Type: Latin America, Export Value as Outcome, and Different Window Lengths

Panel A. Outcome Variable: log(Rev)

By Homogeneity of Products		By Degree of Catalogization		By Product Value	
T*Post	0.134***	T*Post	0.154***	T*Post	0.186***
	(0.023)		(0.024)		(0.018)
T*Post*Homogeneous	-0.071**	T*Post*Q2	-0.061**	T*Post*Price[10,50):	-0.018
	(0.033)		(0.031)		(0.026)
		T*Post*Q3	-0.055*	T*Post*Price[50,200)	-0.050**
			(0.031)		(0.026)
		T*Post*Q4	-0.064	T*Post*Price>200	-0.076***
			(0.054)		(0.026)
R^2	0.99	R^2	0.92	R^2	0.99

Panel B. +/- 12 Months; Outcome Variable: log(Q)

By Homogeneity of Products		By Degree of Catalogization		By Product Value	
T*Post	0.221***	T*Post	0.270***	T*Post	0.231***
	(0.028)		(0.039)		(0.015)
T*Post*Homogeneous	-0.040	T*Post*Q2	-0.026	T*Post*Price[10,50):	-0.009
	(0.040)		(0.052)		(0.021)
		T*Post*Q3	-0.105**	T*Post*Price[50,200)	-0.045**
			(0.051)		(0.021)
		T*Post*Q4	-0.085	T*Post*Price>200	-0.037*
			(0.089)		(0.021)
R^2	0.99	R^2	0.97	R^2	0.99

Panel C. +/- 12 Months; Outcome Variable: log(Rev)

By Homogeneity of Products		By Degree of Catalogization		By Product Value	
T*Post	0.181***	T*Post	0.235***	T*Post	0.227***
	(0.055)		(0.027)		(0.017)
T*Post*Homogeneous	-0.055	T*Post*Q2	0.040	T*Post*Price[10,50):	-0.005
	(0.078)		(0.036)		(0.024)
		T*Post*Q3	-0.038	T*Post*Price[50,200)	-0.038
			(0.035)		(0.024)
		T*Post*Q4	-0.042	T*Post*Price>200	-0.079***
			(0.062)		(0.024)
R^2	0.98	R^2	0.98	R^2	0.99

Notes: We control for all variables and fixed effects according to specification (1), as well as product category fixed effects. Standard errors clustered at the country level.

*** indicates significance at p = 0.01; ** indicates p = 0.05.

Table A4: Heterogeneous Effects by Buyer Experience: Latin America, Export Value as Outcome, and Different Window Lengths

Panel A. Dependent Variable: log(Rev)			
	Overall	Homog. Product	Diff. Product
T*Post	0.146***	0.116***	0.152***
	(0.022)	(0.025)	(0.046)
T*Post*Experienced	-0.067**	-0.065*	-0.056
	(0.031)	(0.035)	(0.065)
R^2	0.97	0.97	0.89
Panel B. 12 months before and after, Qty			
	Overall	Homog. Product	Diff. Product
T*Post	0.167***	0.214***	0.224***
	(0.015)	(0.016)	(0.025)
T*Post*Experienced	-0.022	-0.030	-0.085**
	(0.021)	(0.023)	(0.035)
R^2	0.98	0.97	0.94
Panel C. 12 months before and after, Rev			
	Overall	Homog. Product	Diff. Product
T*Post	0.152***	0.203***	0.206***
	(0.020)	(0.022)	(0.039)
T*Post*Experienced	-0.022	-0.063**	-0.107**
	(0.028)	(0.031)	(0.055)
R^2	0.99	0.95	0.85

Notes: Standard errors clustered at the country level.
*** indicates significance at p = 0.01.

Table A5: Overall Policy Effect for FRITES and Russia

Panel A. Exports to FRITES				
	Control Group 1 log(Export Quantity)		Control Group 2 log(Export Value)	
	(1)	(2)	(3)	(4)
	+/-6 mo.	+/-12 mo.	+/-6 mo.	+/-12 mo.
T*Post	0.139***	0.136***	0.187**	0.119***
	(0.048)	(0.044)	(0.083)	(0.056)
R^2	0.99	0.99	0.99	0.99
Panel B. Exports to Russia				
	Control Group 1 log(Export Quantity)		Control Group 2 log(Export Value)	
	(1)	(2)	(3)	(4)
	+/-6 mo.	+/-12 mo.	+/-6 mo.	+/-12 mo.
T*Post	0.171***	0.120**	0.058	0.009
	(0.027)	(0.026)	(0.105)	(0.068)
R^2	0.99	0.99	0.99	0.99

Notes: We control for fixed effects according to specification (1). Standard errors clustered at the country level.

*** indicates significance at p = 0.01; ** indicates p = 0.05.

Table A6: Heterogeneous Treatment Effects for FRITES and Russia

Panel A. Exports to FRITES					
By Homogeneity of Products		By Degree of Catalogization		By Product Value	
	log(Q)		log(Q)		log(Q)
T*Post	0.167***	T*Post	0.220***	T*Post	0.209***
	(0.018)		(0.038)		(0.051)
T*Post*Homogeneous	-0.062***	T*Post*Q2	0.060	T*Post*Price[10,50):	-0.031
	(0.022)		(0.051)		(0.072)
		T*Post*Q3	-0.089*	T*Post*Price[50,200)	-0.013
			(0.051)		(0.072)
		T*Post*Q4	-0.229***	T*Post*Price(200+)	-0.120*
			(0.089)		(0.072)
R^2	0.99	R^2	0.97	R^2	0.99
Panel B. Exports to Russia					
By Homogeneity of Products		By Degree of Catalogization		By Product Value	
	log(Q)		log(Q)		log(Q)
T*Post	0.164***	T*Post	0.206***	T*Post	0.333***
	(0.035)		(0.061)		(0.025)
T*Post*Homogeneous	-0.343***	T*Post*Q2	-0.229***	T*Post*Price[10,50):	-0.348***
	(0.050)		(0.066)		(0.036)
		T*Post*Q3	-0.278***	T*Post*Price[50,200)	-0.377***
			(0.066)		(0.036)
		T*Post*Q4	-0.318***	T*Post*Price(200+)	-0.391***
			(0.068)		(0.036)
R^2	0.99	R^2	0.99	R^2	0.99

Notes: We control for fixed effects according to specification (1). Standard errors clustered at the country level.

*** indicates significance at p = 0.01; ** indicates p = 0.05.

Table A7: Heterogeneous Effects by Buyer Experience in FRITES and Russia

Panel A. Exports to FRITES

Dependent Variable: log(Export Quantity)

	(1) All Products	(2) Homogeneous Products	(3) Differentiated Products
T*Post	0.187***	0.106	0.189***
	(0.045)	(0.076)	(0.057)
T*Post*Experienced	-0.087*	-0.062***	-0.062
	(0.049)	(0.096)	(0.062)
R^2	0.99	0.99	0.95

Panel B. Exports to Russia

Dependent Variable: log(Export Quantity)

	(1) All Products	(2) Homogeneous Products	(3) Differentiated Products
T*Post	0.201***	-0.188**	0.202***
	(0.048)	(0.074)	(0.040)
T*Post*Experienced	-0.171**	-0.212***	-0.161*
	(0.080)	(0.050)	(0.087)
R^2	0.99	0.99	0.97

Notes: Standard errors clustered at the country level.
*** indicates significance at p = 0.01.

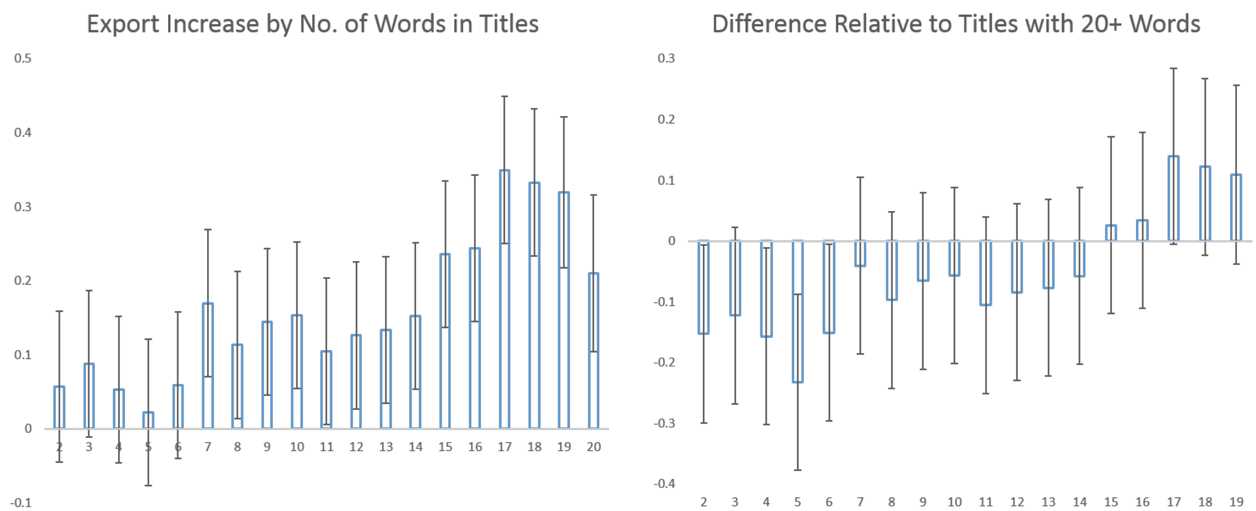
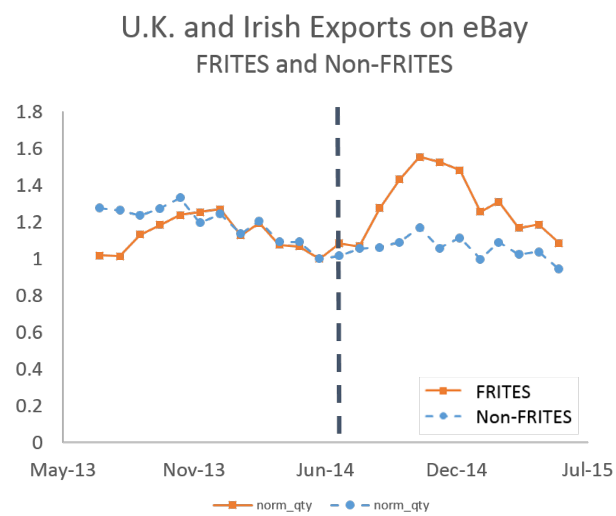
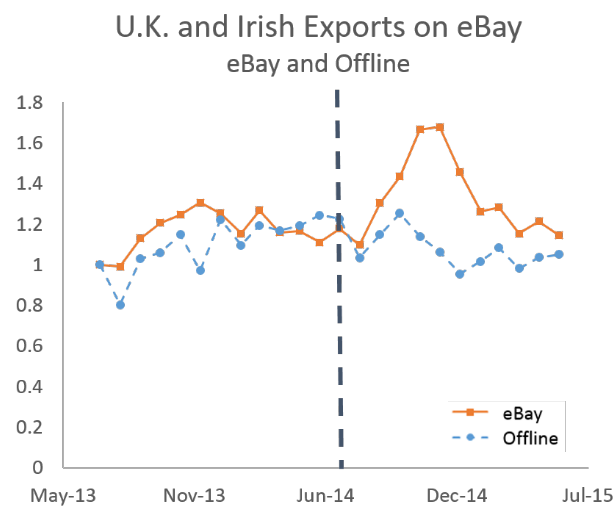


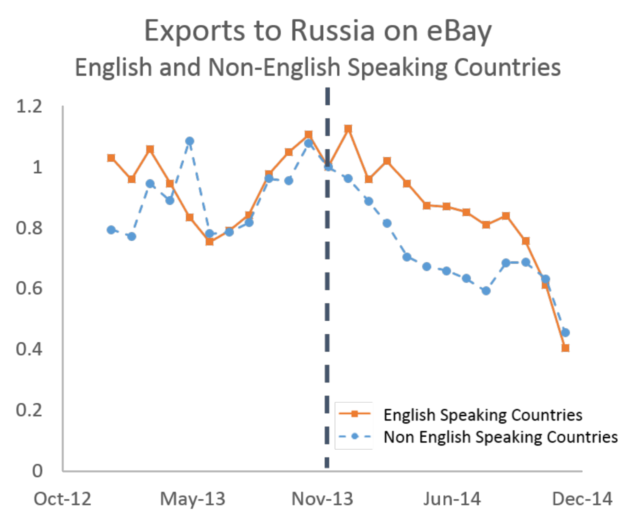
Figure A1: Export Increase by Number of Words in Listing Titles: Latin America, Export Value as Outcome



(a)



(b)



(c)

(d)

Figure A2: Parallel Trend Assumptions

Notes: Exports in Figure A2a are measured in quantity and are normalized to the level in June 2014. Exports in Figure A2b are measured in dollars and are normalized to the level in June 2013. Exports in Figure A2c are measured in quantity and are normalized to the level in December 2013. Exports in Figure A2d are measured in dollars and are normalized to the level in November 2013. The dashed vertical lines refer to the introduction of query translation and item title translation.

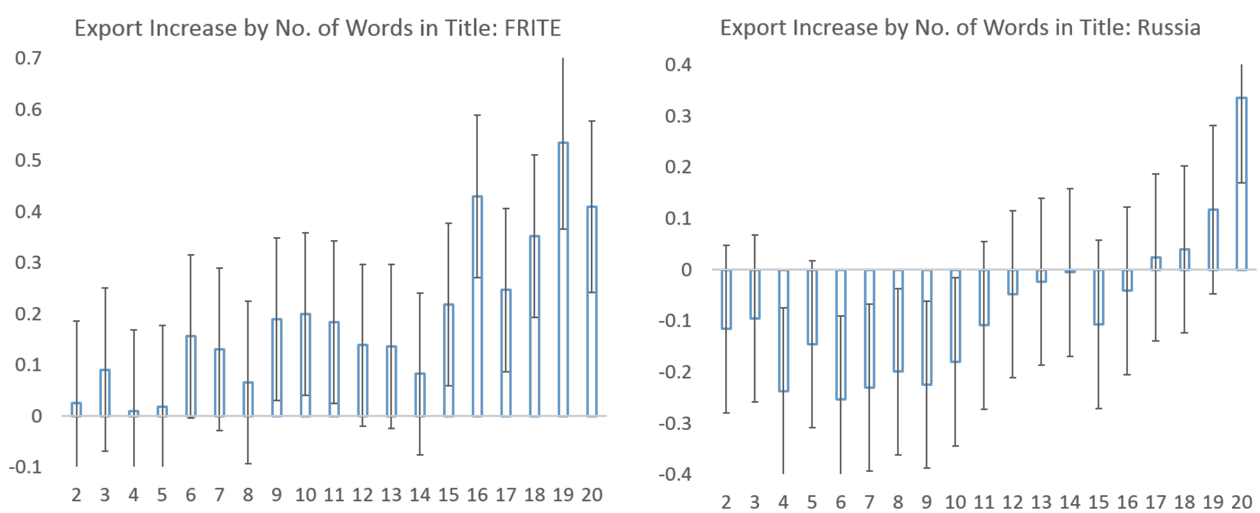


Figure A3: Export Increase by Number of Words in Listing Titles for FRITES and Russia