Can Digital Distribution Defy the Law of Gravity?

Han Yang[*] Yuta Watabe[†] Eugene Kanasheuski[‡]

February 3, 2023

Abstract

Online delivery of digital media content is expected to nullify geographical friction. However, we provide evidence that physical distances still matter for sales on digital distribution platforms. Using data from country-level game sales collected from Steam, the world's largest computer game distribution platform, we show that consumers are less likely to buy games developed in distant countries. The effect is driven by changes in quantity, not by changes in price. Further analysis shows that information friction partially explains the tendency. By developing a model, we quantify the significance of this effect on gains from free trade.

Keywords: Digital Economics, International Trade, Gravity Equations
JEL codes: F15, F23, L96

[*]Institute of Economics, Academia Sinica, han.yang271@gmail.com.
[†]Xiamen University, yutawatabe@gmail.com.
[‡]Solsten, evgen.kanashevsky@gmail.com.

1 Introduction

Digital distribution drastically transformed the trade in digital goods. Companies such as Netflix, Spotify, and Youtube have built platforms that deliver digital content internationally without any additional cost for consumers. The digital nature of these companies' content distribution strategies allows them to overcome distance. For example, streaming Japanese anime in the US is as easy as streaming Mexican dramas in India. With the adoption of the internet and digitization, these services seem to embody a flat world as depicted by Friedman (2007), where anyone can consume digital content from any country around the world. Nevertheless, the conventional gravity model of international trade implies that international sales tend to fall as trading partners become more distant from each other. Whether this phenomenon, which people often refer to as a law of gravity, holds true in digital content has yet to be discovered.

We investigate whether the distance effect perishes or thrives in a purely digital world. Digital distribution is an extreme case of how the internet and digitization transformed the nature of trade. Studying gravity in a purely digital environment is a significant step toward understanding the geographical barriers in international trade and how it is affected by digitization. If distance still matters in the pure digital world, where there is no physical cost of trade, the force of gravity is ubiquitous. Furthermore, by precluding the trade cost explanation, we will better understand the underlying mechanisms of trade frictions.

In this paper, we study effects of geographical distance on trade using a novel dataset from Steam, which is the world's largest digital distributor of computer games. The platform is completely virtual, hence physical distances do not play any practical role in the availability and the delivery of games. We collect sales, prices, genres and the developer's location for games curated by Steam. Using our data, we then perform gravity regressions, and find a robust negative effect of distance on game sales. Our results suggest that, even in the pure digital environment, online sales of PC games on Steam decline as developers and consumers are further apart, indicating distance still matters for trade in digital world. In our specifications, the distance elasticity ranges between -0.1 to -0.2, which are about 10% to 20% of the conventional estimates of trade of physical goods in the literature Disdier and Head (2008). The results suggests that, even in a completely digital environment, there are enduring trade frictions.

To further dissect the driving forces of gravity in digital trade, we perform various regressions to understand the underlying mechanism for the effect of distance on digital sales. Specifically, we test the cost channel and the information channel. To test the cost explanation, we first decompose the effect of distance on sales into prices and quantities.

We find that, as sellers and consumers become more distant, total sales decline mostly through changes in quantity sold rather than changes in the pricing. The results indicate that the cost, which should be reflected in the price, is not driving the gravity.

We look into information channels. Consumers may be ill-informed about games developed in distant locations, hence, it may require longer time for games from distant countries to be acknowledged by gamers. We investigate this channel by grouping games by their release dates. Since older games had more time to become known, they are less likely to be subjected to information friction through distance. Our regression results show that sales of older games are indeed less affected by physical distance, while sales of newer games are more affected by distance. Similar contrasts are found by separating games into non-indie games and indie games. Non-indie developers, which are more affluent in the advertising budget, usually tend to utilize offline advertising, which is more likely to be subjected to distance. On the other hand, indie developers rely more heavily on online advertising.[1] Our results shows that, compared with sales of indie games, sales of non-indie games decline more intensively with increased distance between consumers and developers. Overall, the results indicate that information friction can be potential contributing factors of gravity in digital international trade.

We also perform additional regressions by adding different proximity measures, such as language and cultural distance. Our results are robust; the inclusion of these proximity measures does not change the coefficients of physical distance on game sales. We then examine whether the distance effect is coming from regional commonality and find that distance is still significant even when we control for the common region dummy.

As we highlight the existence of the gravity force present in the digital economy, a natural question arises: how quantitatively important are these enduring frictions? To answer this, we include the enduring frictions to the Armington model and calibrate the model using both the conventional commodity trade data and our Steam dataset to calibrate physical trade costs and the residual frictions separately. The key difference between the two calibrations is that while physical trade costs can be eliminated by reducing trade costs, the latter remains in the economy even if the trade is free. Using our calibrated model, we perform counterfactual experiments to evaluate the welfare implications of residual friction. We consider two cases: the total elimination of both the physical trade barrier and residual friction and the elimination of only the physical trade barrier while keeping residual friction intact. The former counterfactual experiment ignores the fact that the residual friction cannot be eliminated. We compare the

[1]Examples of offline advertising are participation in trade events (e.g., E3, Tokyo Game show) and street advertising.

welfare gains from both counterfactual scenarios to elucidate the quantitative importance of residual frictions. The result suggests there are quantitatively substantial differences in trade potential when the residual frictions are considered. For example, the welfare gains for India are 60% smaller when the residual friction is incorporated into the model. The result also shows that the effects of friction are unevenly distributed; some countries even have higher welfare gains by keeping the residual frictions.

As IT technology and computer become more widely adopted, research on digitization and the economy is also growing. As discussed in Goldfarb and Tucker (2019), one of the most important consequences of digitization is the reduction of trade costs. To the best of our knowledge, our paper is the first study to quantify the effect of geographical distances on international transaction in the purely digital world.

Our paper contributes to the literature on gravity and digital economy. Blum and Goldfarb (2006) offer insights on digital contents and geography. They study the internet browsing activities of US users, and show that viewers are less likely to visit websites that are farther away from the US. They also find that distance matters only to taste related websites, and conclude that the taste differences correlated with distances can be explained by the law of gravity. Our paper also sheds light on the role of internet and geography. Moreover, our dataset provides an unique vantage point to examine online transaction of digital goods, where the goods are coherently organizes by one unified digital platform. Compared to website visits on the internet, cross-border purchase of digital goods provides a more striking resemblance to cross-country consumption of commodities. Hence, our results can be extended to conventional mode of international trade in commodities.

This paper examines a new aspect of how the internet changes the nature of trade. Freund and Weinhold (2004) study on how increased internet access induced more trade. Subsequent study shows that the internet makes trade more sensitive to distance (Akerman, Leuven and Mogstad, 2022). Various studies show how E-commerce reduce distance-related trade costs (Hortaçsu, Martínez-Jerez and Douglas (2009), Lendle et al. (2016), Fan et al. (2018)). We complement these studies by looking at international trade in digital goods, where physical trade costs are absent. While E-commerces such as eBay and Alibaba reduced trade costs via the utilization of the internet, transactions in these platforms still involves transportation and shipping. In contrast, Steam completely extinguishes physical activities associated with trade. Instead of focusing on trade costs reduced by the internet, we look at frictions that endure in the completely digital platform.

Our paper also highlights the increasing role of information on trade. Allen (2014) demonstrates the significance of information friction on trade and our results insinuate that information play a role in explaining the effect of distance on trade. Several studies suggest that information friction plays important roles in online e-commerce platforms (Chen and Wu, 2021), and e-commerce platforms make an active effort to alleviate the relevant information friction (Carballo et al. (2022) and Lendle et al. (2016)). In this paper, we show that even in Steam — an unified and organized digital distribution platform — information is likely to be a possible determining force of gravity.

Finally, we believe our study provides an interesting example of trade in cultural goods (Disdier et al., 2010). Digital distribution platform is well suited for consumption of cultural goods like media such as music, movies, and games. One noticeable study is Ferreira and Waldfogel (2013), which examines patterns of global music consumption using popular music chart data. They show that domestic consumption is still relatively more substantial (and even increasing) in recent decades.[2] Our study adds to the literature by showing that trade in PC games, a modern example of cultural goods, is subject to gravity.

In our quantitative exercises, we borrow the concept of trade potential, which is the welfare gains from completely eliminating trade costs, as defined in Waugh and Ravikumar (2016). Based on the theoretical observation that free trade equilibrium must equalize trade shares across destinations, they construct a sufficient statistic approach to study the welfare implication of trade potential. We complement their analysis by demonstrating that the implications of trade potential can be drastically different if there are enduring frictions that persist in a free trade environment.

The remainder of the paper is as follows. In section 1, we introduce our data from Steam and provide some descriptive statistics. In section 2, we perform gravity regressions to test whether the global sales of games are affected by geography. Finally, in section 3, we develop a quantitative model to demonstrate the role of enduring frictions we unveil in the Steam data.

2 Data

Our analysis relies on unique sales data from Steam, which is a digital distribution platform of video games developed by Valve Corporation. The video games distributed by Steam only include PC games; console games and mobile games are not available on Steam. The platform is compatible with major computer operating systems, including Microsoft Windows, Apple Mac OS, and Linux. Steam was first launched in 2003 and was originally developed to deliver updates or patches for Valve's own games, including the Half-Life series and Counter-Strike. In 2005, Valve started negotiating

[2] In their study, they do not specify how these music are traded (e.g., streaming, internet, CD, ...).

with third-party publishers and developers to publish their games on Steam. Although Steam started with only seven games in 2004, by May of 2007, even though there were only 150 games available on Steam, 13 million user profiles had already been created. Steam grew rapidly and became the largest digital distribution platform for PC gaming. In 2013, Steam had around 75% of the market share of the digital distribution platform in PC games. By 2019, Steam curated more than 30,000 games and had over 95 million active monthly users.

The virtual environment we have chosen has several novel features that are distinctive from traditional merchandise trade.[3] The most unique characteristic of Steam is that goods being traded are completely digital. Games being purchased are delivered via the internet; users do not get physical copies of contents they purchase. There are various studies on digital platforms, which utilize data from Alibaba, Ebay, and Amazon, etc. Most merchandise being traded on those platforms are physical, while in our dataset, both the distribution platform and traded goods are digital. The second distinctive feature of Steam is that its markets are geographically segregated. Users from different regions can only purchase content in local currencies, and each developer also sets different prices for different regions. Most digital retail platforms, such as Amazon and Alibaba, only allow sellers to set one price pegged to one single currency for each product. Steam has geographical differentiation in both pricing and currencies. Meanwhile, Steam also implements strong measures to enforce geo-blocking, making it very difficult for users to switch regions. These novel features of the Steam dataset provide a unique vantage point to analyze international trade in digital goods.

For the sales data on games, we use the dataset created by Steam Spy, which is a website founded by Sergey Galyonkin and launched in 2015. Steam Spy estimates the number of copies sold and average playtime for game titles available on Steam by using the platform's application programming interface (API), which allows programmers to pull information on user profiles, including titles in the inventory, location of users and total playtime for each title. The data from Steam Spy records the monthly number of copies sold for each country at title level. For example, sales of The Witcher 3: Wild Hunt developed by CD PROJEKT RED. Steam Spy was deemed as the most reliable data source for PC gaming. PCGameN reports that Steam Spy is accurate to within 10% of actual sales for games. Many developers use data from Steam Spy to navigate their business decisions. The dataset we collect from Steam Spy covers sales for more than 9,000 unique games in 75 destination countries between February of 2017 to March of

[3]However, we believe this environment is close to merchandise trade rather than service trade. These games are usually purchased outright, and consumers do not communicate or negotiate with sellers. The transaction resembles buying an e-book on Amazon.

2018.

We also utilize Steam's API to extract detailed game-level information, including genres, prices and developers' name. Games in our sample include variety of genres, i.g. first-person shooter (FPS), role-playing games (RPG), fighting games, etc. To construct bilateral trade flow in game sales. We utilize the prices of games across all destination countries, the names of developers/studios. Our data include titles developed by almost 6,000 developers. We use the developer's name to match each studio to the country of its headquarters manually. Our cleaned sample includes measured bilateral sales from 67 origin countries to 75 destination countries for each title and developer.[4]

ISO-3	Frequency	Percentage		Frequency	Percentage
USA	3,060	30.50543	CHN	147	1.465457
GBR	1,026	10.22829	NLD	135	1.345828
RUS	772	7.696142	CZE	109	1.086631
CAN	592	5.901705	HUN	82	.8174659
JPN	565	5.632539	ROM	69	.6878676
DEU	560	5.582694	HRV	64	.6380221
FRA	363	3.618782	DNK	60	.5981457
POL	285	2.841192	IND	54	.5383312
AUS	249	2.482305	MEX	53	.5283621
SWE	231	2.302861	TUR	52	.518393
ESP	173	1.724654	AUT	50	.4984548
UKR	172	1.714684	KOR	48	.4785166
ITA	164	1.634932	NOR	42	.418702
BRA	157	1.565148	ZAF	39	.3887947
FIN	157	1.565148	BEL	38	.3788256

Table 1: Top 30 countries for headquarters' location

	Mean	s.d.	p25	p50	p75
Sales	238746	1327088	3510.54	27469.62	120653.2
Price	10.77393	11.76527	4.99	7.99	14.99
Quantity	14593.73	47898.17	1111	4243	13538

Table 2: Summary of sales, price and quantity sold in the US

We show descriptive statistics of the data. Table 1 shows the top 30 countries that host the most game titles. Games developed in the US take the largest share of games in

[4]The list of countries are in the appendix.

	Mean	s.d.	p10	p50	p90
Countries sold	73.00432	5.567719	66	75	75

Source: Calculated by authors

Table 3: Number of countries sold

our data but there are games from various countries. Table 2 presents summary statistics of the sales, price, and quantity sold in the US market. The mean and median sales in the US are 238,746 USD and 27,469 USD respectively. The mean and the median price in the US is 10.7 USD and 7.99 USD[5]. The mean and median quantity sold in the US is about 14,600 copies and 4,243 copies, respectively, and the larger mean compared to the median indicates that there are some superstar triple A titles in our data which sell great number of copies. Table 3 shows the number of countries these games are available. We consider a game being available in a country if we can observe the game's price in the country's digital storefront. The majority of games in our sample are available in all 75 destination countries in our data set, indicating that Steam's global digital storefront assures the access to games regardless of users' location.

3 The Empirical Model

First, we aggregate our data on cross-country game-level level sales into country-level total sales. Then we run gravity regressions with various conventional bilateral resistance variables, including bilateral distance, to outline the effect of geography in trade of digital goods. The specification is as follows:

$$\begin{aligned}\ln X_{in} =& \beta_{dist} ln(Distance_{in}) + \beta_{cont} Contiguity_{in} + \beta_{colony} Colony_{in} \\ &+ \beta_{commonlang} Common\ language_{in} + \beta_{HMK} Home\ market_{in} + \xi_i + \zeta_n + \varepsilon_{in},\end{aligned}$$

where subscript i is origin country (developer's location), subscript n is the destination country, and X_{in} is the total sales in country n for games developed in i. Our main focus is on β_{dist}, which is the coefficient on physical distance between i and n. Other variables include whether the countries are contiguous, whether they were in colonial relationships, whether they have common official languages, and whether the origin and the destination are the same. These geographical variables between countries are from CEPII GeoDist (Mayer and Zignago, 2011). The last two terms, ξ_i and ζ_n, are dummy variables for the origin country i and the destination country n.

[5]The price of games are much lower than typical AAA games since Steam sells indie games developed in a smaller scale.

Table 4: Gravity: Country level sales

	(1) OLS	(2) OLS	(3) PPML	(4) PPML
Distance	-0.12***	-0.12***	-0.20**	-0.20**
	(0.042)	(0.043)	(0.099)	(0.095)
Contiguity	0.12	0.12	-0.41*	-0.40*
	(0.078)	(0.079)	(0.25)	(0.23)
Colonial relationship	-0.020	-0.018	0.18**	0.16***
	(0.082)	(0.080)	(0.073)	(0.057)
Common official languages	0.053	0.062	0.049	0.11
	(0.072)	(0.069)	(0.098)	(0.081)
Home market		0.36***		-0.27
		(0.11)		(0.22)
Observations	3088	3132	4940	5007

Standard errors in parentheses

Origin / Destination fixed effects are included and standard errors are clustered

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The results are shown in Table 4. Where Columns (1)-(2) show the OLS regression results that exclude zero sales. And the columns (3)-(4) include observations containing zero sales using Poisson Pseudo Maximum Likelihood Silva and Tenreyro (2006). For each specification, we show the results with both exclusion and inclusion the domestic sales.

In our context, international trade of digital goods is accomplished by digital delivery, physical trade costs are completely absent, hence we expect β_{dist} to be both very small and statistically insignificant. However, we find that the coefficients on distance are statistically significant and non-negligible in its magnitudes, and the results are robust across various specifications. Compared with exiting estimates in the literature using conventional trade data of physical goods, the distance coefficients typically range around -1.1. Meanwhile, in our estimation using cross-country sales of PC games, the coefficients on distance range from -0.12 to -0.2. The distance elasticity in our environment of digital trade facilitated by digital platform is about only 10% to 20% of the size of the estimates using conventional commodity trade. Our distance coefficients from e-commerce data are smaller than the existing estimates that also use data from e-commerce platforms; Fan et al. (2018) utilize data on Chinese inter-provincial trade in Alibaba, and find that the distance coefficient is -0.47. Lendle et al. (2016) use international transaction data from eBay, and their estimate on the distance coefficient is -0.44. Our estimated distance effect is about 25% to 50% of the estimates using other e-commerce platform that mostly focus on trade in physical goods. The coefficients of other bilateral geographical variables, in both magnitude and significance, do not vary much across specifications. The signs of the coefficients on contiguity and the colonial relationship changes when zero sales are included. The common language dummy, which we expect to have strong impacts on sales, is surprisingly insignificant. To sum, for international trade of PC games on the digital platform, physical distances play more important role than other geographical variables. The sign of the home bias depends on the specification.

3.1 Game-level gravity

In this section, we fully utilize our cross-country game-level sales information to perform gravity regressions at product level. This allows us to include game fixed effects and language adoption variables, which is defined as whether a game supports the official language in the destination market. Our specification is as follows:

$$
\begin{aligned}
lnX_{i(j)n}(j) = {} & \beta_{dist} ln(\mathit{Distance}_{i(j)n}) + \beta_{cont}\mathit{Contiguous}_{i(j)n} + \beta_{colony}\mathit{Colony}_{i(j)n} \\
& + \beta_{commonlang}\mathit{Common\ language}_{i(j)n} + \beta_{HMK}\mathit{Home\ market}_{i(j)n} \\
& + \beta_{langadoption}\mathit{Language\ adoption}(j)_{jn} + \xi_j + \zeta_n + \varepsilon_{jn},
\end{aligned}
$$

where $X_{i(j)n}(j)$ is the total sales in n of game j developed in i. In addition to country-level variables, we include the dummy variable Language adoption$(j)_n$, which takes one if game j supports the language used in country n. This represents the role of language in a finer level than the common language indicator. The last two variables, ξ_j and ζ_n, are game-j and destination-n specific fixed effects.

The results are shown in Table 5. Each column follows the specifications shown in the previous table. The coefficients on distance are statistically significant at the 5% level for all specifications. The magnitude of the coefficients of OLS is much smaller than that of country-level estimation, which is likely due to numerous zero sales for games. When zero sales are included, the coefficients are even larger than the country-level estimates, at -0.19 for both with and without domestic sales. A novel fact we found is the role of language adoption. When the game adopts the local language, the sales increases about 19% to 40%. While this is not a causal estimate, the regression indicates language's

Table 5: Gravity: Game level sales

	(1) OLS	(2) OLS	(3) PPML	(4) PPML
Distance	-0.043***	-0.041***	-0.19**	-0.19**
	(0.0099)	(0.010)	(0.086)	(0.080)
Contiguity	0.028	0.018	-0.39*	-0.38**
	(0.030)	(0.029)	(0.21)	(0.19)
Colonial relationship	0.091***	0.093***	0.17**	0.16***
	(0.033)	(0.031)	(0.069)	(0.051)
Common official languages	0.082***	0.11***	0.038	0.095
	(0.030)	(0.030)	(0.083)	(0.074)
Local language adopted	0.19***	0.19***	0.35***	0.40***
	(0.048)	(0.048)	(0.11)	(0.11)
Home market		0.16***		-0.25
		(0.045)		(0.19)
Observations	106,060	111,694	609,538	661200

Standard errors in parentheses

Game / Destination fixed effects are included and standard errors are clustered

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

significant role in trade at the product level. Other geographical variables and the home bias remain unstable across specifications.

4 Discussion

International transactions of digital goods are facilitated by digital means in that no physical shipping and transportation costs are involved in product delivery. However, we still observe a significant and sizable effect of distance on bilateral sales. This disparity leads us to ask Why distances may still be relevant in a purely digital environment. We rule out two explanations discussed in the literature of E-Commerce. We first rule out the role of physical geographical barriers. Games sold on Steam do not come in physical medium; hence, we believe that physical trade costs do not explain the observed gravity in game sales. The second potential mechanism for the distance effect is the geographically correlated transaction risks. Cheating is a common concern for e-commerce. For example, sellers may not deliver goods or ship low-quality products. This dishonest behavior may become more prevalent when sellers and consumers are more distant from each other. [6] In usual trade, for consumer to resolve for these cheating behaviors, it would require buyers to negotiate with dishonest sellers. The costs of potential negotiation are higher when sellers are farther away. As a result, the buyer may be more hesitant to buy products from distant locations. However, the transaction risks are small on Steam because game purchases are simple outright purchases, and Steam also has an established uniform returning policy (the platform issue a refund for any game title that is requested within fourteen days of purchase and has been played for less than 2 hours). Even if the quality of a game is much lower than the developer promised, consumers can easily receive refund without interacting with the developer. Therefore, based on the robust consumer protection policies, we rule out geographically correlated transaction risk as an explanation to observed distance effects.

We investigate various potential mechanisms for the observed gravity, including trade costs, information channels, and taste frictions. We first discuss the trade costs channel. Conventional theory suggests that trade costs increase with distance, which would indicate that the prices are higher when games are delivered to places farther from the developers, leading to lower sales. To investigate this channel, we decompose sales to price and quantity, and we examine the effect of each force separately. Our regression specification follows:

$$\begin{aligned} lnZ_{i(j)n}(j) &= \beta_{dist} ln(\mathit{Distance}_{i(j)n}) + \beta_{cont}\mathit{Contiguous}_{i(j)n} + \beta_{colony}\mathit{Colony}_{i(j)n} \\ &+ \beta_{langadoption}\mathit{Language\ adoption}(j)_n + \beta_{HMK}\mathit{Home\ market}_{i(j)n} \\ &+ +\beta_{commonlang}\mathit{Common\ language}_{i(j)n} + \xi_j + \zeta_n + \varepsilon_{jn}, \end{aligned}$$

where $Z_{i(j)n}(j)$ is the price or quantity sold in n for game j developed in i.

The results are shown in Table 6 and Table 7. Each column on the table displays the coefficients from our regressions. The coefficients of distance on prices are very small. Even though the coefficients in OLS are statistically significant, they are much smaller when compared to the coefficients on total sales. In contrast, the distance coefficients on quantity are close to that of sales. The coefficients are statistically significant at the 10% level, and the magnitude is in line with that of sales regression. These results suggest that the decline in total sales due to distances is mainly driven by changes in quantity sold, not through changes in prices. Hence, we rule out the trade-cost explanation; if

[6] This is close to the argument in Lendle et al. (2016). They found that sales of Power seller – a seller with a higher rating – is less affected by distance.

Table 6: Gravity: Game level price

	(1) OLS	(2) OLS	(3) PPML	(4) PPML
Distance	0.0029**	0.0030**	-0.00061	-0.0013
	(0.0012)	(0.0012)	(0.0021)	(0.0020)
Contiguity	0.0079	0.0078	-0.013	-0.013
	(0.0052)	(0.0051)	(0.0086)	(0.0083)
Colonial relationship	0.0038	0.0039	0.0060*	0.0062**
	(0.0028)	(0.0027)	(0.0032)	(0.0031)
Common official languages	0.0033	0.0037	0.00039	0.00047
	(0.0036)	(0.0036)	(0.0019)	(0.0021)
Local language adopted	0.00070	0.0016	0.0019	0.0026
	(0.0039)	(0.0038)	(0.0059)	(0.0058)
Home market		0.013		0.0055
		(0.0077)		(0.0089)
Observations	733,439	743,468	733,439	743,468

Standard errors in parentheses

Game / Destination fixed effects are included and standard errors are clustered

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

trade costs are responsible for the distance effect, prices should also respond to geographical distances. The result on quantity suggests there are other possible channels: information frictions and taste differences, which we discuss below.

We move on to the information channel. Steam users can directly look up information about games within the platform[7] or they can also learn about these games from external sources. We argue that the information, especially the information outside of the Steam platform, plays an important role, because games developed in farther away

[7] We believe friction to search games within Steam, should not be the explanation. Consumers from different countries face the exact same storefront. Unlike Amazon, where the storefronts are different across countries (e.g., amazon.com, amazon.co.uk, etc.), consumers can easily search the games from different countries. This does not mean that Steam suggests the exact same games. While differential suggestions of the game could be possible and likely to happen, we believe this is not the reason for gravity. Valve, the company operating Steam, has no incentive to recommend games by the origin of the developer.

Table 7: Gravity: Game level quantity

	(1) OLS	(2) OLS	(3) PPML	(4) PPML
Distance	-0.042***	-0.040***	-0.17*	-0.15*
	(0.0097)	(0.0099)	(0.085)	(0.085)
Contiguity	0.022	0.012	-0.25	-0.23
	(0.026)	(0.026)	(0.16)	(0.16)
Colonial relationship	0.088***	0.090***	0.22***	0.19***
	(0.032)	(0.031)	(0.077)	(0.067)
Common official languages	0.079***	0.10***	0.10	0.19***
	(0.028)	(0.028)	(0.077)	(0.071)
Local language adopted	0.19***	0.20***	0.29**	0.34***
	(0.048)	(0.047)	(0.12)	(0.12)
Home market		0.15***		-0.11
		(0.042)		(0.17)
Observations	106,060	111,694	609,538	661,200

Standard errors in parentheses

Game / Destination fixed effects are included and standard errors are clustered

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

places may be less likely to be noticed by consumers. To examine this plausible channel, we split our sample into new games and old games by their release dates. News about releases of new games may take longer to travel if the developers are far away from gamers, which may explain the effect of geographical distance in the observed gravity. Furthermore, since older games had more time to establish the reputation, they are less likely to be influenced by the presence of physical distances. Games in our sample are split into new games and old games according to the release dates. A game is considered a new game if it is released after January 1, 2017, and is considered an old game if it is released before January 1, 2017.[8] Our results are shown in Table 8, and we observe that distance has much weaker impact on old games than the on new games. Our results corroborate the information channel story. Newer games are more affected by geographical

[8]We also tried to split the sample with the median release date, around October 2015. We obtained a similar result.

distances, because time is needed for information about new games to travel to gamers that are farther away.[9] Nevertheless, when we only focus on older games, distances still have significant impact on sales, which indicates that thee are other potential forces beyond information channels driving the distance effect.

We next discuss the information channel from a different perspective. We split our sample into games that are labelled as indie games and that are not labelled as indie games. Indie games usually are developed by smaller developers with more limited resources, while non-indie games are usually developed by larger studios with more abundant resources. Furthermore, indie games developers usually advertise less intensively, and they often focus more on online advertising if any. In contrast, non-indie game developers utilize both online and offline advertising as their marketing strategies.[10] The results are shown in table 9. Columns (1)-(4) show the regression results for only indie games, and columns (5)-(8) show the results for only non-indie games. While distance coefficients are statistically significant for both indie and non-indie games in the PPML specification, the magnitudes of the coefficients vary vastly. The distance coefficients on indie games are much smaller compared to non-indie games. These results highlight the possible role of information channel through advertisement. Since non-indie games rely more on offline ad campaigns, and may advertise less intensively in more distant countries, they often result in lower sales in a remote market.

We finally argue about the role of taste friction. The taste of games can be geographically correlated. Hence people favor games that are developed in closer countries. As seen in the results, our proposed information channel does not fully account for the distance effect in digital trade. We conjecture this unexplained component may be an indirect evidence of taste friction.

[9] This could be a quality concern (consumers may hesitate to buy goods without enough reviews), but we believe this is less likely. Steam has a return policy that allows testing games before returning. Furthermore, given the nature of the goods, the cost of returning goods is uniform and does not depend on the distance.

[10] Online advertising utilizes Social Network Services, YouTube and other online mediums, while offline advertising utilizes trade shows, and physical ad campaigns.

Table 8: Gravity: Games before and after 2017

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Old games	Old games	Old games	Old games	New games	New games	New games	New games
	OLS	OLS	PPML	PPML	OLS	OLS	PPML	PPML
Distance	-0.0340***	-0.0362***	-0.0571**	-0.0608**	-0.0943***	-0.0890***	-0.343***	-0.334***
	(-3.71)	(-4.10)	(-2.74)	(-3.10)	(-3.82)	(-3.90)	(-3.75)	(-4.00)
Contiguity	0.0323	0.0187	-0.108	-0.111*	-0.0426	-0.0611	-0.698*	-0.661**
	(1.40)	(0.80)	(-1.80)	(-2.30)	(-0.83)	(-1.34)	(-2.51)	(-2.66)
Colonial relationship	0.0643*	0.0698*	0.0740	0.0689	0.134**	0.136***	0.221*	0.200*
	(2.31)	(2.62)	(1.34)	(1.56)	(3.27)	(3.71)	(2.41)	(2.00)
Common official language	0.0840**	0.101***	0.0512	0.0970	0.0850	0.122**	0.103	0.155
	(3.14)	(3.96)	(0.84)	(1.61)	(1.99)	(2.80)	(0.70)	(1.16)
Local language adopted	0.164***	0.169***	0.260**	0.304**	0.200*	0.212*	0.491***	0.560***
	(3.68)	(3.86)	(2.90)	(3.12)	(2.45)	(2.58)	(4.50)	(4.85)
Home market		0.128***		-0.0342		0.129*		-0.367
		(4.05)		(-0.56)		(2.61)		(-1.53)
N	86,215	90,602	451,619	481,804	10,945	11,750	109,944	126,317

Standard errors in parentheses

Origin (Game) / Destination fixed effects are included and standard errors are clustered

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Gravity: Comparing indie games and non-indie games

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Indie	Indie	Indie	Indie	Non-Indie	Non-Indie	Non-Indie	Non-Indie
	OLS	OLS	PPML	PPML	OLS	OLS	PPML	PPML
Distance	-0.019	-0.016	-0.032**	-0.033*	-0.070***	-0.071***	-0.23**	-0.23***
	(0.012)	(0.014)	(0.016)	(0.017)	(0.014)	(0.012)	(0.096)	(0.089)
Contiguity	0.080*	0.066*	0.026	0.019	-0.050	-0.057*	-0.54**	-0.52**
	(0.041)	(0.039)	(0.052)	(0.039)	(0.031)	(0.030)	(0.27)	(0.24)
Colonial relationship	0.075**	0.079**	0.13***	0.12***	0.095***	0.095***	0.16**	0.14**
	(0.037)	(0.034)	(0.047)	(0.040)	(0.032)	(0.030)	(0.075)	(0.059)
Common official languages	0.080**	0.11***	0.071	0.14**	0.095***	0.11***	0.036	0.090
	(0.030)	(0.033)	(0.054)	(0.057)	(0.035)	(0.033)	(0.11)	(0.10)
Local language adopted	0.15***	0.17***	0.22***	0.27***	0.20***	0.21***	0.41***	0.44***
	(0.048)	(0.048)	(0.077)	(0.084)	(0.049)	(0.048)	(0.12)	(0.11)
Home market		0.21***		0.19***		0.091**		-0.42*
		(0.061)		(0.056)		(0.042)		(0.24)
Observations	57,482	60,834	381,819	417,706	48,578	50,860	227,719	243,494

Standard errors in parentheses

Game / Destination fixed effects are included and standard errors are clustered

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To summarize, we argue that, because of the institutional features of the Steam platform, the observed distance effects are unlikely to be explained by trade costs and contract issues, in contrast to more conventional understanding of trade. Our price regressions solidify the fact that prices are not affected by distance, hence refute the trade costs explanation. We estimate game-level gravity by splitting games in our sample into indie games and non-indie games, and new games and old games. We find that information channels can be potential drivers for the observed distance effects. We also believe that geographically correlated tastes can partially explain the results.

4.1 Robustness

We use additional empirical specifications for robustness. The first concern is that geographical distance may be a simple proxy for cultural and language proximity. To address this concern, we incorporate other variables, such as language and cultural proximity between buyers' and developers' locations. We adopt the language proximity index constructed by Melitz and Toubal (2014).[11], and cultural distances as in Hofstede's Cultural Dimensions constructed by Geert Hofstede. The language proximity capture the closeness and similarity for languages used by each country pair, while Hofstede's cultural dimensions measures cultural aspects across countries, including individualism, masculinity, and uncertainty avoidance. We calculate the Mahalanobis' distance of these cultural variables to construct a single-dimensional cultural distance for each country pair. The results are shown in Table 11. Columns (1)-(2) are results from data aggregated to country-level, and columns (3)-(4) are results from game-level data. For all of our specifications, coefficients on distance all remain statistically significant and are close to that of our previous estimates. Furthermore, the coefficients on language proximity are positive, implying that the closer the languages in origin country and destination country are, the larger the sales. In contrast, the coefficients on cultural distance share the same positive sign, but are statistically insignificant in most of our specifications. We show that physical distance on bilateral sales is robust even when other possible cultural distances are controlled for.

In addition, we consider the same-region effect by examining whether trading partners are situated in the same region. Using regions defined in the United Nations Geoscheme, we include a dummy variable that takes a value one if the origin country and the destination country are located in the same region, and takes a value of zero otherwise. By using this classification, we investigate whether the inclusion of regional

[11] We use the unadjusted value of linguistic proximity (Tree). The result remains unchanged when we use other variables constructed in the paper.

Table 10: Gravity: Regional dummies

	(1)	(2)	(3)	(4)
	Country sales	Country sales	Game sales	Game sales
	OLS	PPML	OLS	PPML
Distance	-0.083**	-0.15**	-0.020**	-0.14***
	(0.040)	(0.060)	(0.0094)	(0.052)
Contiguity	0.056	-0.45**	0.0037	-0.43**
	(0.082)	(0.22)	(0.031)	(0.20)
Colonial relationship	0.054	0.18**	0.079***	0.17**
	(0.088)	(0.070)	(0.024)	(0.068)
Common official languages	0.055	0.14	0.083**	0.13
	(0.11)	(0.14)	(0.036)	(0.11)
Language proximity	0.056**	0.070*	0.039***	0.063*
	(0.026)	(0.039)	(0.0086)	(0.033)
Cultural distance	-0.022	-0.023	-0.046***	-0.017
	(0.039)	(0.026)	(0.0098)	(0.020)
Common region	0.025	0.15**	0.030	0.13*
	(0.057)	(0.075)	(0.022)	(0.070)
Local language adopted			0.17***	0.37***
			(0.051)	(0.11)
Observations	1,996	2,593	90,568	405,345

Standard errors in parentheses

Origin (Game) / Destination fixed effects are included and standard errors are clustered

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Gravity: Including various distances

	(1) Country sales OLS	(2) Country sales PPML	(3) Game sales OLS	(4) Game sales PPML
Distance	-0.087**	-0.17**	-0.024**	-0.16**
	(0.040)	(0.073)	(0.0092)	(0.065)
Contiguity	0.059	-0.39*	0.011	-0.37**
	(0.082)	(0.21)	(0.029)	(0.18)
Colonial relationship	0.051	0.18***	0.077***	0.17***
	(0.089)	(0.068)	(0.024)	(0.066)
Common official languages	0.059	0.16	0.090**	0.15
	(0.12)	(0.14)	(0.036)	(0.12)
Language proximity	0.057**	0.085*	0.041***	0.076**
	(0.026)	(0.045)	(0.0083)	(0.038)
Cultural distance	-0.024	-0.031	-0.047***	-0.023
	(0.039)	(0.030)	(0.0099)	(0.024)
Local language adopted			0.17***	0.37***
			(0.051)	(0.11)
Observations	1,996	2,593	90,568	405,345

Standard errors in parentheses

Game (Origin) / Destination fixed effects are included and standard errors are clustered

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

commonality (e.g., common cultural references in the region) affects the distance coefficients in our gravity estimation. We find that distance coefficients are still significant even controlling for the common region variable. The results are shown in Table 10. Columns (1)-(2) are results from data aggregated to country-level, and columns (3)-(4) are results from game-level data. In all our specifications, the estimated distance coefficients are very close to original estimates, showing that the physical distance effect on digital trade are robust when common-region effects are controlled for.[12]

5 Quantitative analysis

Our empirical investigation indicates enduring trade frictions associated with physical distance in the pure digital distribution platform, which resembles an economy with free trade because the need for physical transportation and shipping is completely eliminated. Even if all international trade becomes completely digital and goods can be transported freely across the globe, certain trade frictions may linger. We construct a simple quantitative model that incorporates the enduring trade friction to answer the following question: How much are we overestimating the gains from free trade if the enduring frictions are neglected?

We adapt the standard Armington model (Armington, 1969) to incorporate the enduring trade frictions that cannot be eliminated by free trade.[13] There are two types of trade frictions in the model: iceberg trade costs t_{in}, and the residual trade frictions τ_{in}. Our model yields the following gravity equation:

$$X_{in} = \xi_i \zeta_n t_{in} \tau_{in}^{1-\sigma}$$

where X_{in} is the total export from i to n, ζ_i and ξ_n are functions of multilateral resistances. And σ is the elasticity of substitution. Trade frictions t_{in} and τ_{in} are inseparable when using conventional trade data on physical goods.[14] We adopt a two-step procedure to estimate and calibrate these two types of trade frictions. First, we run a gravity regression using conventional commodity trade data, which gives $t_{in}\tau_{in}^{1-\sigma}$. They are

[12]On top of this specification, we tried including an EU dummy, where the variable takes one if both countries are in the EU. The distance coefficients are similar.

[13]Details of the model are explained in the appendix. It is possible to develop a model with product heterogeneity, but since our interest is to quantify the role of unexplained frictions, we start with the Armington model.

[14]In the model, we are agnostic about the mechanism of the friction and simply formulate t_{in} as a demand shifter.

specified as:

$$t_{in}\tau_{in}^{1-\sigma} = (Distance_{in})^{\delta_{dist}} * exp\Big(\delta_{cont}Contiguity_{in} + \delta_{colony}Colony_{in} + \delta_{commonlang}Common\ language_{in}\Big).$$

Next, we utilize the gravity regression coefficients estimated using the Steam data to back out the enduring trade friction t_{in}. Since physical shipping and transportation are completely absent in game sales on Steam ($\tau_{in} = 1$), we use the coefficients as proxies of t_{in}. Specifically, we assume that:

$$t_{in} = Distance_{in}^{\beta_{dist}}.$$

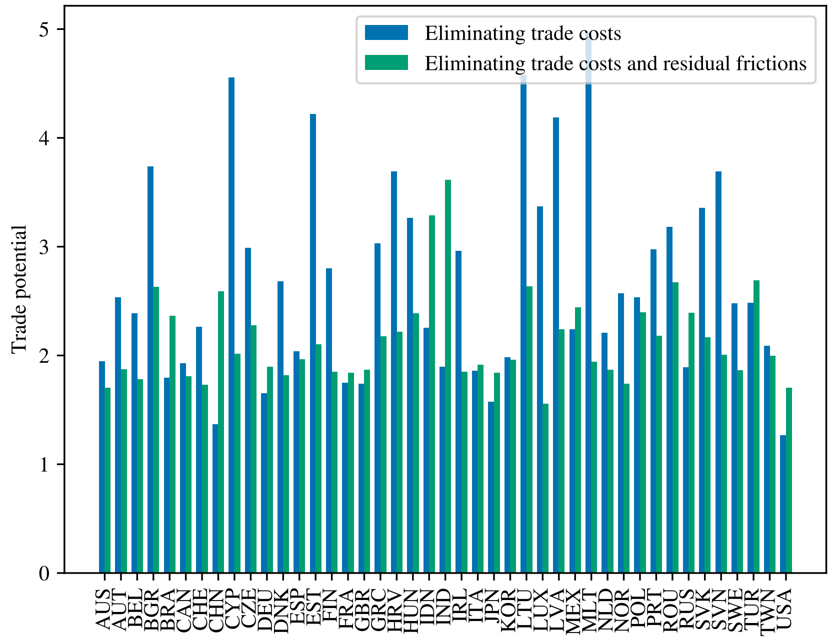
We take value of β_{dist} from the gravity regression using the digital trade data from Steam. In particular, β_{dist} is set to 0.1.[15] Combining our estimated coefficients from trade in physical goods and trade in digital goods on Steam, we back out enduring trade frictions t_{in} and iceberg trade frictions τ_{in}.

We calibrate the parameters of our quantitative model for 43 countries to year 2014 using the data from the World Input-output Database (Timmer et al., 2015)[16]. The elasticity parameter σ is set to 5, following Simonovska and Waugh (2014). Using the calibrated model, we perform following two counterfactual experiments: (1) eliminate trade costs completely while keeping residual frictions intact, and (2) eliminate both trade costs and residual frictions completely. We denote the original parameter as x and counterfactual parameters as x'. For the first counterfactual experiment, we set $\tau'_{in} = 1$, and $t'_{in} = t_{in}$, while for the second counterfactual, we set: $\tau'_{in} = 1$ and $t'_{in} = 1$.

We evaluate and compare the welfare gains from these two counterfactual experiments. The welfare gains from the benchmark economy to free trade economy echoes the concept of trade potential in the literature, see Waugh and Ravikumar (2016) for example. Specifically, trade potential calculates the utility ratio between the current and the autarky equilibrium. The first quantitative experiment represents the true trade potential (welfare gains from free trade) when all the trade costs and enduring trade frictions are correctly considered, while the second experiment represents the mis-specified trade potential when residual frictions are treated the same as iceberg trade costs. Our interest is in the relative welfare gains of the two counterfactual scenarios. In the correctly specified free trade economy, only the iceberg trade costs are removed, and the enduring trade friction persist even in the free trade economy; failing to acknowledge

[15]Our assumption is that these frictions observed in Steam can be extrapolated to aggregate economy.
[16]Details of the calibration is shown in the appendix[:unedited]

Figure 1: Trade potentials

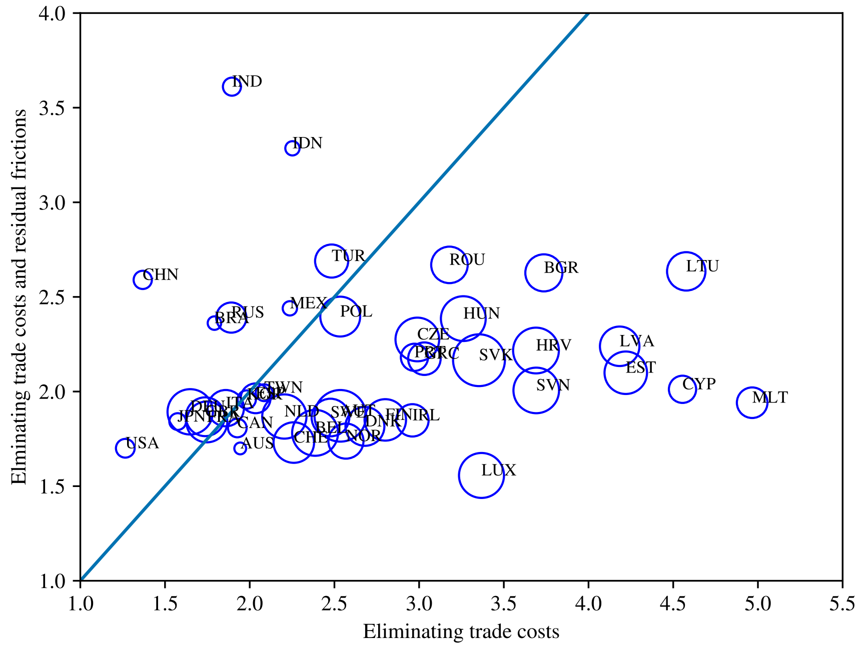


the enduring friction would lead to biased quantitative evaluation of the trade potential. Hence, we compare the welfare implications of these two experiments to highlight the quantitative importance of residual trade friction.

The result is shown in Figure 1. Trade potentials for both specifications, range from 1.2 to 5.0, where the magnitude is in line with Waugh and Ravikumar (2016). These number indicates that countries have 20% to 400% higher welfare from being in free trade. However, there are significant differences between these two correctly specified and mis-specified trade potentials. For example, the trade potential for India is around 3.6 for the incorrectly specified case, but is only 1.9 for the correctly specified case. In that case, trade potential in India is overstated when free trade fails to consider the role of enduring friction. However, the presence of the residual frictions may also lead to understatement of trade potential for some countries. Countries may have better terms of trade (relative wage) from the existence of residual frictions. For example, the trade potential for Malta is around 5.0 for the correctly specified case, but is only around 2.0 for the incorrectly specified case.

Which countries would show larger discrepancies in trade potential when residual frictions are correctly incorporated? In figure 2, we compare these two trade potentials in a scatter-plot. The vertical axis shows the trade potential of the incorrect specification, and the horizontal axis displays the trade potential of the correct specification. The size of the circle represent the centrality measure for each country, which is calculated

Figure 2: Trade potentials: Comparison



by the sum of the reciprocal of bilateral distances between countries[17]. A larger centrality value indicates that a country is more geographically centered. We can roughly divide countries in our sample into two two groups: (1) large but more geographically isolated countries, such as China, India and Indonesia (2) small but geographically centered countries, such as Malta, Cyprus and Lithuania. The first group exhibits higher trade potential for the incorrectly specified case, while the latter group exhibits higher trade potentials for the correctly specified case. The intuition is the following. Since residual frictions are driven by distance, eliminating them benefits the isolated countries while harms the well-connected countries (by the deterioration of terms of trade). Overall, residual frictions are quantitatively important in measuring trade potentials, and have unequal effects on different countries.

6 Conclusion

Digital distribution is drastically reshaping the landscape in consumption and delivery of goods. Using a purely digital distribution platform, consumers can enjoy content from far away places without paying any additional costs. In this paper, we ask if a purely digital distribution platform is able to eliminate the role of distance in the gravity model. We answer this question by using the data from Steam, the world's largest

[17]Specifically, geographical centrality for country i is calculated by a following formula: $\text{centrality}_i = \sum_{1=n}^{N} \frac{1}{dist_{in}}$.

distribution platform of computer games. Using the sales and the developers' location of each game, we show that, even in the entirely digital economic environment, geographical resistances persist and remain strong. By estimating gravity regression using data on trade in digital goods, we find that sales still decline as the developer and the consumer are more distant. Nevertheless, the distance elasticity on sales ranges between -0.1 to -0.2, which is about 10% to 20% of the conventional estimates on using trade in physical goods. Our result indicates that although the digital trade is successful to reduce the cost from distances, there are tenacious frictions that persist even in this pure digital environment. Our additional analysis shows that these frictions work mainly through changes in quantity sold, not through changes prices, suggesting that the distance mostly affects digital trade via the demand channel, not via the changes in trade costs. We further show that the information friction may be a plausible explanation on this distance effect. Our result is robust to various estimation specifications and choices made on sample selection.

We argue that enduring trade frictions are quantitatively important. By adopting a standard Armington model, we compare the welfare implications for complete trade cost reduction with and without considering the enduring trade frictions. Specifically, we compare the two counterfactuals, which eliminate only the trade cost but not enduring trade friction, and the another eliminates both the trade cost and the trade friction. Without acknowledging the existence of non-erasable trade friction, the quantitative analysis would lead to a very biased evaluation of trade potentials. Ignoring enduring trade friction and eliminating it tends to overstate trade potentials for some countries and understate trade potentials for some other countries.

References

Akerman, Anders, Edwin Leuven, and Magne Mogstad. 2022. "Information Frictions, Internet, and the Relationship between Distance and Trade." *American Economic Journal: Applied Economics*, 14(1): 133–63.

Allen, Treb. 2014. "Information Frictions in Trade." *Econometrica*, 82(6): 2041–2083.

Armington, Paul S. 1969. "A Theory of Demand for Products Distinguished by Place of Production (Une théorie de la demande de produits différenciés d'après leur origine) (Una teoría de la demanda de productos distinguiéndolos según el lugar de producción)." *Staff Papers (International Monetary Fund)*, 16(1): 159–178.

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

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.

Chen, Maggie X., and Min Wu. 2021. "The Value of Reputation in Trade: Evidence from Alibaba." *The Review of Economics and Statistics*, 103(5): 857–873.

Disdier, Anne-Célia, and Keith Head. 2008. "The Puzzling Persistence of the Distance Effect on Bilateral Trade." *The Review of Economics and Statistics*, 90(1): 37–48.

Disdier, Anne-Célia, Silvio H. T. Tai, Lionel Fontagné, and Thierry Mayer. 2010. "Bilateral trade of cultural goods." *Review of world economics*, 145(4): 575–595.

Fan, Jingting, Lixin Tang, Weiming Zhu, and Ben Zou. 2018. "The Alibaba effect: Spatial consumption inequality and the welfare gains from e-commerce." *Journal of international economics*, 114: 203–220.

Ferreira, Fernando, and Joel Waldfogel. 2013. "Pop Internationalism: Has Half a Century of World Music Trade Displaced Local Culture?" *The Economic Journal*, 123(569): 634–664.

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

Friedman, Thomas L. 2007. *The world is flat.* Penguin Books.

Goldfarb, Avi, and Catherine Tucker. 2019. "Digital Economics." *Journal of economic literature*, 57(1): 3–43.

Hortaçsu, Ali, F. Asís Martínez-Jerez, and Jason Douglas. 2009. "The Geography of Trade in Online Transactions: Evidence from eBay and MercadoLibre." *American Economic Journal: Microeconomics*, 1(1): 53–74.

Lendle, Andreas, Marcelo Olarreaga, Simon Schropp, and Pierre-Louis Vézina. 2016. "There Goes Gravity: eBay and the Death of Distance." *The Economic journal (London)*, 126(591): 406–441.

Mayer, Thierry, and Soledad Zignago. 2011. "Notes on CEPII's distances measures: The GeoDist database." CEPII Working Papers 2011-25.

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

Silva, J. M. C. Santos, and Silvana Tenreyro. 2006. "The Log of Gravity." *The Review of Economics and Statistics*, 88(4): 641–658.

Simonovska, Ina, and Michael E Waugh. 2014. "The elasticity of trade: Estimates and evidence." *Journal of International Economics*, 92(1): 34–50.

Timmer, Marcel P., Erik Dietzenbacher, Bart Los, Robert Stehrer, and Gaaitzen J. de Vries. 2015. "An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production." *Review of International Economics*, 23(3): 575–605.

Waugh, Michael E., and B. Ravikumar. 2016. "Measuring openness to trade." *Journal of Economic Dynamics and Control*, 72: 29–41.

A Quantitative model

Here we illustrate the details of our quantitative model. The model extends the Armington model to include the enduring frictions. The residual frictions are modeled as demand shifters, in the sense that consumers derive less utility from goods that are developed in more distant locations. There are N countries in the economy, and in each country, there is a representative consumer and a representative firm.

The representative firm in each country produce goods using labor, and sell the goods at marginal cost. Since the market is perfectly competitive, the representative firm in i sells the product to location n at marginal costs, which equals the cost s of labor and iceberg trade costs. The price of the goods produced in i and sold in n is

$$p_{in} = A_i w_i \tau_{in},$$

where A_i is the labor requirement for producing goods in i, w_i is the labor wage in i, and τ_{in} is the iceberg trade cost of delivering goods from i to n.

Country i supplies L_i amount of labor inelastically, and use entirety of the earning to consumer goods. A representative consumer in i derives utility from consuming goods from countries $n \in N$:

$$U_i = \left(\sum_{n \in N} t_{in}^{1/\sigma} q_{in}^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)},$$

where q_{in} is the consumption of goods from j by a representative consumer i, σ is the elasticity of substitution between goods, and t_{in} is a bilateral demand shifter. This demand shifter corresponds to the residual frictions we estimate in Steam data. The demand function of the representative consumer i becomes

$$X_{in} = t_{in} p_{in}^{1-\sigma} P_n^{\sigma-1} I_n$$

where P_{in} is a price index of n and I_n is a total expenditure of n.

The market clearing condition is:

$$w_i L_i = \sum_{n \in N} X_{in} + D_i.$$

Where D_i is a trade deficit of country i. An equilibrium of this economy is $\{w_i, p_i, X_i, D_i\}$ that satisfies consumer optimization, producer optimization and market clearing conditions.

A.1 Calibrating the model

We use the baseline parameters taken from the World Input Output tables (WIOD) and calibrate further parameters using the regression result from Steam. Our model yields the following gravity equation:

$$X_{in} = \xi_i \zeta_n t_{in} \tau_{in}^{1-\sigma}$$

where $\zeta_i = A_i w_i$ and $\xi_n = P_n X_n$. Denote $d_{in} = t_{in} \tau_{in}^{\sigma-1}$ as a combination of trade costs and the residual frictions.

We retrieve $\boldsymbol{w}$, $\boldsymbol{X}$ and $\boldsymbol{I}$ from the WIOD, and we calibrate $\boldsymbol{A}$ and $\boldsymbol{d}$ using gravity equation. More precisely, we perform a two-way fixed effect gravity equation with the following specification for $\boldsymbol{d}$:

$$d_{in} = (\text{Distance}_{in})^{\delta_{dist}} * exp\Big(\delta_{cont} Contiguity_{in} + \delta_{colony} Colony_{in} + \delta_{commonlang} Colony_{in}\Big).$$

The productivity parameter $\boldsymbol{A}$ is backed out from the estimate of $\xi_i = A_i w_i$ and the wage data from WIOD. To separate out $\boldsymbol{t}$ and $\boldsymbol{\tau}$, we adopt a following specification:

$$t_{in} = \text{Distance}_{in}^{\beta_{dist}}.$$

and the coefficient β_{dist} is taken from the gravity regression using Steam data. [18] We can now calibrate $\boldsymbol{\tau}$, since we have a calibrated both $\boldsymbol{d}$ and $\boldsymbol{t}$.

In summary, $\boldsymbol{w}$, $\boldsymbol{X}$ and $\boldsymbol{I}$ are taking from WIOD, $\boldsymbol{A}$ and $\boldsymbol{d}$ are calibrated with gravity equations using WIOD, and Steam data is used to separate $\boldsymbol{d}$ to $\boldsymbol{t}$ and $\boldsymbol{\tau}$.

B List of countries

The list of exporting countries are Argentina, Australia, Austria, Brazil, Bulgaria, Canada, Chile, China, China, Hong Kong Special Administrative Region, Colombia, Costa Rica, Croatia, Cyprus, Czech Republic, Denmark, Ecuador, Egypt, Estonia, Finland, France, Georgia, Germany, Hungary, India, Indonesia, Ireland, Israel, Italy, Japan, Kazakhstan, Latvia, Lithuania, Malaysia, Malta, Mexico, Morocco, Netherlands, New Zealand, Norway, Pakistan, Peru, Philippines, Poland, Portugal, Republic of Korea, Russian Federation, Saudi Arabia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sweden, Switzerland, Taiwan, Thailand, Turkey, Ukraine, United Arab Emirates, United Kingdom of

[18] The assumption is that the residual frictions in the aggregate economy is common with that of Steam. While this may not be the exactly true, this is a good starting point for a quantitative exploration.

Great Britain and Northern Ireland, United States of America, Uruguay, Venezuela (Bolivarian Republic of), Viet Nam.

The importing countries are Argentina, Australia, Austria, Azerbaijan, Bangladesh, Bolivia (Plurinational State of), Bosnia and Herzegovina, Brazil, Bulgaria, Canada, Chile, China, China, Hong Kong Special Administrative Region, Colombia, Costa Rica, Croatia, Cyprus, Czech Republic, Denmark, Ecuador, Egypt, Estonia, Finland, France, Georgia, Germany, Hungary, India, Indonesia, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Latvia, Lithuania, Malaysia, Malta, Mexico, Morocco, Netherlands, New Zealand, Norway, Pakistan, Peru, Philippines, Poland, Portugal, Qatar, Republic of Korea, Russian Federation, Saudi Arabia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sweden, Switzerland, Taiwan, Thailand, Turkey, Uganda, Ukraine, United Arab Emirates, United Kingdom of Great Britain and Northern Ireland, United States of America, Uruguay, Venezuela (Bolivarian Republic of), Viet Nam.