UCD CENTRE FOR ECONOMIC RESEARCH

WORKING PAPER SERIES

2023

Online Interactions, Trade And The Gravity Model

Constantin Bürgi   Zuzanna Studnicka

University College Dublin

WP23/18

August 2023

UCD SCHOOL OF ECONOMICS
UNIVERSITY COLLEGE DUBLIN
BELFIELD
DUBLIN 4

Online Interactions, Trade And The Gravity Model*

Constantin Bürgi[†] Zuzanna Studnicka[‡]

August 9, 2023

Abstract

We evaluate the drivers of online interactions and the importance of these interactions for international trade patterns. To this end, we measure the volume of online interactions between countries using a unique individual level data set with over 25 million multiplayer games played between December 2019 and August 2022 in the game Age of Empires II: Definitive Edition. We first show that, in line with the gravity model, distance is a crucial factor in determining the pattern of online interactions when trade costs are low or non-existent. This suggests that the distance effect in this environment captures additional factors such as cultural proximity. Subsequently, we use this interaction data as a new measure of cultural proximity and apply it to bilateral goods trade. We find that our measure is highly significant and reduces the importance of gravity variables, particularly the distance coefficient by around 10 %. It is also more important for diversified goods than homogeneous goods.

JEL: F10, F14, Z10

Keywords: Gravity equation, services trade, online interactions

*We would like to thank Ron Davies and seminar participants at University College Dublin for their valuable comments and support. Thanks also to Yixuan Li for research assistance.

[†] *Email: constantin.burgi@ucd.ie, University College Dublin, Belfield, Dublin 4, Ireland*

[‡] *Email: zuzanna.studnicka@ucd.ie, University College Dublin, Belfield, Dublin 4, Ireland*

1 Introduction

Online interactions across borders offer a vast array of opportunities, including the purchase and use of products and services, participation in online discussions, and the making of new connections. While there are many regional or country-specific restrictions in place with regard to a purchase, online discussions, etc. (e.g., geoblocking, or language) there is one area where such restrictions are much more limited, namely online gaming. Generally, unless the purchase of a game is restricted, there are usually no country-specific limitations on online interactions between players. There are even fewer limitations for slower-paced games like real-time strategy games where proximity to a server or other players is of lesser importance.[1] This feature makes online gaming an ideal environment to test whether the gravity model holds when there are only a few formal cross-country restrictions.

There are additional features of online gaming that reduce barriers present in other services. For example, online discussions and real-time video calls are only possible if participants can understand each other. In contrast, video games do not require players to chat with one another, and therefore, whether they share a common language or not should not make any difference in the likelihood of two people playing against each other. Furthermore, the video games sector has rapidly grown in importance over the past few years. The most important platform Steam has steadily increased the number of concurrent users from around 100 thousand in 2004 to over 30 million in 2022 as shown in Figure 1. The number of unique monthly users is over four times this figure. Similarly, the global video game sector has increased its revenue dramatically over the past decades. In 1980, its size was comparable to just the US box office, which was around $1 billion, and now it has a size similar to the entire global film industry reaching over $130 billion in 2018.[2]

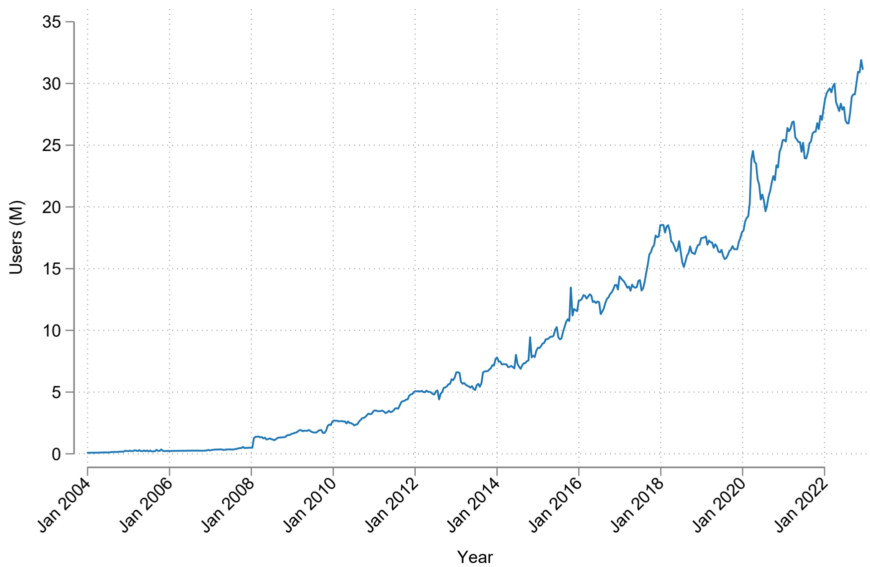
Thanks to their bilateral nature, online interactions can be examined using the gravity model - an analytical framework of significant importance in international economics (see e.g., Anderson 1979, McCallum 1995, Anderson and Van Wincoop 2003 and Yotov et al. 2016). Instead of using masses of and distances between objects to obtain the gravitational force in physics, this model employs factors such as economic size and trade frictions to predict trade patterns between countries. As trade frictions cannot be measured directly, physical distance, common language, and international borders serve as proxies.

Several studies have examined the applicability of the gravity model in a low physical trade cost environ-

[1] However, it is worth noting that some countries may enforce restrictions on the number of hours an individual can spend playing a game (Reuters, 2021).

[2] Source: Newzoo.com and Ibisworld.com

Figure 1: Number of Concurrent Steam Users (2004-2005)



Note: The figure plots the number of players on the video game platform Steam at a given point in time. Source: steamdb.info

ment such as online exchanges. Research on this topic focuses on website views and purchases (Blum and Goldfarb, 2006), the use of online searches (Alaveras and Martens, 2015); as well as entertainment (Hellmanzik and Schmitz 2015 and Broocks and Studnicka 2021). In general, this literature has found strong evidence that people are still more likely to visit websites, watch movies etc., from their home country and from countries nearby, in line with the gravity model. The same is true for services in their language, even after controlling for service quality (see e.g., Broocks and Studnicka 2021) or the stock of migrants.[3]

The literature has proposed two broad explanations driving the relevance of physical distance in a low trade costs environment (even in situations where no physical shipping takes place). First, cultural familiarity/proximity leads to similar tastes (Blum and Goldfarb 2006 or Huang 2007). Two countries that are far apart will have fewer cultural exchanges between them, meaning that their cultures and tastes will be more different. The second explanation is informational familiarity, meaning that without knowledge of potential firms, products or services, there will be no trade (Chaney, 2014). Networks play a key role for informational familiarity and a firm is more likely to trade with a firm in another country if its suppliers and/or customers

[3] These findings are consistent with the high persistence of the distance effects documented by a meta-analysis of Disdier and Head (2008), who examine 1,467 distance effects estimated in 103 papers and find that the negative effect of the distance is highly persistent over time.

trade with firms in that country. However, while directly comparing trade in goods sold on eBay with trade in similar goods sold in a traditional way Lendle et al. (2016) finds that the distance effect is about 65% smaller on eBay suggesting that online technologies reduce this type of frictions.

To our knowledge, whether the distance coefficient is relevant for direct online interactions between people has not yet been explored. Our first hypothesis is that the choice of who interacts with whom in a gaming platform should not be affected by players' locations, languages (or other gravity variables). This is because the transaction/transport costs in this situation are very low, if not non-existent. It is particularly true for the game on which we focus in this paper - Age of Empires II: Definitive Edition.

Several characteristics of this game explain why gravity variables like physical distance should have little to no explanatory power. Firstly, it is a relatively slow-paced game, so input delay caused by the physical distance to the server is not a major factor. Second, players are not required to communicate with each other during the match. While there is a chat function that can be used by the players, this is not mandatory and the only customary exchange is an exchange of "gg" (good game) at the end of the game, initiated by the resigning player.[4]

To explore whether the distance and other trade costs variables matter in this environment, we use various empirical methods and a large array of control variables. Our first finding is that the commonly used proxies for trade costs are still significant despite extremely low/no trade costs. We, therefore, reject our first hypothesis. This suggests that players specifically discriminate due to their preferences.

Once we establish that the gravity model explains the pattern of online interactions even with little to no cost to physical distance, we use our online interactions data to test whether it can be used to predict physical trade between countries. As documented in Blum and Goldfarb (2006), Huang (2007) and Disdier et al. (2010) physical distances might be capturing information frictions and cultural distance between countries. Since online gaming should be affected by these factors as well but not by physical distance, the number of matches played provides a new measure for these factors and we want to test whether this measure is relevant for explaining trade patterns as well. Our second hypothesis then is that this measure can be used to predict trade patterns between countries and reduce the size of the distance coefficient by capturing the cultural/network factors otherwise captured by the distance coefficient and other gravity variables. To test this hypothesis, we use Comtrade data at the product level. Lastly, we hypothesize that

[4]Another, less customary exchange would be "glhf" (good luck, have fun) at the beginning of a match.

the effect of our new variable should be more important for trade in differentiated products than for trade in homogeneous products. Homogeneous goods should be substantially less affected by cultural/network factors, while differentiated products might be specifically targeted at specific groups/cultures. To test this last hypothesis, we classify goods into different categories using the Rauch (1999) classifications. While we find that online interactions have a positive effect on trade in all types of goods, the effect is substantially larger for differentiated goods. Therefore, we confirm both of these hypotheses.

The remainder of the paper is structured as follows. Section 2 describes our data. Section 3 describes our theoretical framework and empirical approach. The following Section 4 provides the empirical results and robustness checks and the final section concludes.

2 Data

We use several data sets coming from multiple sources. Our first data set, on online interactions, is described in Subsection 2.1. We merge it with gravity data described in Subsection 2.2. Finally, we use trade data (described in Subsection 2.3) to see if our online interactions data can be used as an explanatory variable in traditional trade regression.

2.1 Measuring Online Interactions

Our main data sources are aoe2insights.com and aoe2.net that collect data on the universe of online matches played in the game Age of Empires II: Definitive Edition.[5] In this real-time strategy game, players start with a small medieval village and must gather resources to develop and expand their village, as well as strengthen their military forces. The objective is to engage in battles against other players, with the ultimate goal of eliminating the opposing player to secure victory.[6] The game has various settings that can be adjusted such as the size of the initial village.

The focus of this paper lies in a key feature of our data set: the knowledge of each player's location.[7] In addition, the data set also includes data on which civilizations were picked by the players at the beginning

[5]For the former, we are particularly thankful that the owners shared the data with us.

[6]It is fairly common that the losing player resigns well before all units/buildings are eliminated/destroyed.

[7]It is our understanding that these country identifiers are related to the account location at the time when the game was purchased and not where each player was located at the time of a specific online match. However, since player counts are broadly proportional to country populations, we assume that no significant differences between the two locations exist.

of each match. These civilizations possess unique strengths and weaknesses and can be mapped to countries that exist today (e.g., the Britons correspond to the UK).

Our initial data set contains information on 50,753,033 matches played between December 2019 and August 2022.[8] We consider each match played as an online interaction. We focus on matches played between two players as these are the most common and most suitable to analyse in the gravity framework due to their bilateral characteristics. After excluding matches played by more than two players, our final data set contains 24,040,177 matches played by 1,446,571 players located in 189 countries. Table 1 presents the statistics on the number of matches played and the number of players each year.

Table 1: Number of matches and players by year (two players games)

	No of matches	No of players
2019	91,014	63,050
2020	8,409,818	767,578
2021	10,059,715	764,255
2022	5,479,630	430,568
Total	24,040,177	1,446,571

Number of players and matches. Source: aoe2insights.com and aoe2.net (2023), own calculations.

For each match, we have the exact time of the match, the location of each player, and the civilization played. We use this data to calculate the number of matches played and hence online interaction between each country pair (direction specific) by weekday and hour. To do so we consider the country of player one as the country of origin, and the country of player two as the destination. Note that, we do not use the year or the month during which the match took place. This is because most of the variation in the number of matches played depends on the weekday and hour.[9] We first focus on the universe of matches played between two players, but will later differentiate between interactions where players were matched by a server (server-matched) and where player two specifically chooses to play with player one (player-matched). In addition, in our robustness check, we remove the direction of flow, considering each country pair as unidirectional.[10]

[8] 275,806 in 2019, 18,523,857 in 2020, 20,712,048 in 2021 and 11,241,322 in 2022.

[9] With the exception of April 2020 Covid-19 lockdowns during which the number of matches played increased significantly as shown in Figure A.2 in Appendix A.

[10] See Section 4.2 for details.

2.2 Gravity Data

We use standard gravity variables provided by the CEPII: distance between two countries' capitals, common language, contiguity and colony (capturing past colonial ties).[11] We extend the CEPII data set with a set of three countries or territories that were not originally included.[12] We use Wikipedia and Time and Date websites to calculate the time zone difference between the two countries.[13]

In addition, we use the migration data. This comes from the United Nations (UN). For each country pair, we use the stock of nationals from the country of player one (origin) in the country of player two (destination) in 2019 (the idea is that migrants want to play with people from their home country).[14] In our robustness checks, when we treat each country pair as one, without the specific direction of matches, we use the sum of migrants from each of the two countries in each country pair.[15] Finally, for domestic matches we use the population size.

Table 2 presents summary statistics for our sample of country pairs with a number of matches greater than zero. The average number of matches played between the two countries was 35.72 with a minimum of one and a maximum of 20,718 (US-US). This number is slightly lower, 31.25 when we exclude server-matched games and focus only on player-matched games. Note that even if these numbers seem small, this is the average number of games played in each hour, including during the night.[16] The average number of players playing during each hour was around 3,000 in both the origin and the destination. In addition, Figure 2 shows the number of matches played by hour in the destination country. Unsurprisingly, the majority of games occur in the evening (between 7PM and 8PM). Conversely, there are fewer matches played overnight (between 3 AM and 8 AM).

[11] Mayer and Zignago (2011)

[12] These include: Guam, Lichtenstein and Monaco

[13] For countries with multiple time zones we use the time zone of the capital.

[14] We have also included the number of migrants from the country of player two (destination) in the country of player one (origin). These results are only marginally different and do not change the rest of our results. They are available upon request.

[15] For example, we use the sum of the German nationals in France and of the French nationals in Germany for games played between Germany and France.

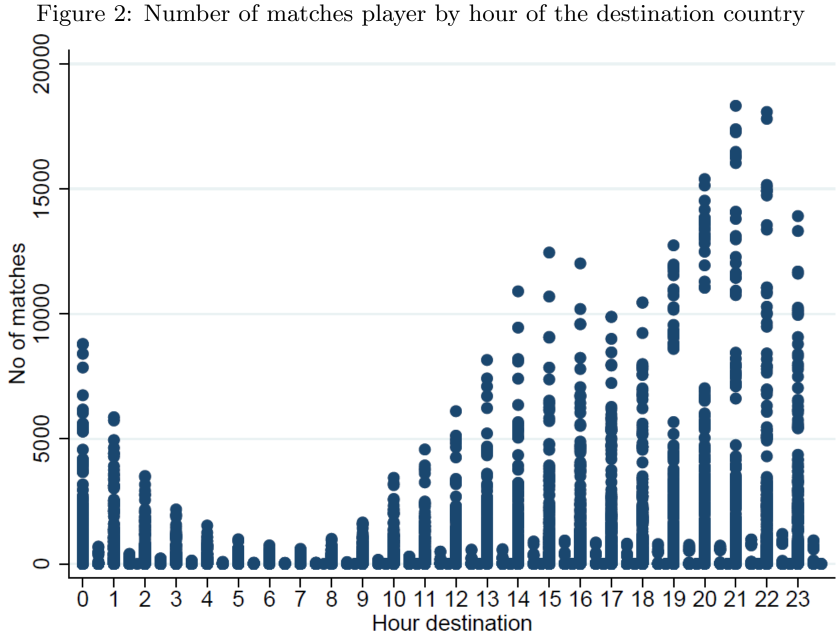
[16] Note that the average number of concurrent users in this game is around 20,000. Source: https://steamdb.info/app/813780/charts/

Table 2: Summary statistics (countries with the number of matches greater than zeros)

Variable	Obs	Mean	Std. Dev.	Min	Max
No of matches$_{ijwt}$ (all)	673,049	35.72	309.21	1	20,718
No of matches$_{ijwt}$ (player-matched)	259,134	31.25	346.49	1	18,327
No of players origin $_{iwt}$	673,049	3,043	5,219	1	34,936
No of players dest.$_{jwt}$	673,049	3,052	5,271	1	35,326
Hour origin$_{it}$	673,049	12.76	6.86	0	23.75
Hour dest.$_{jt}$	673,049	12.76	6.86	0	23.75
Time difference$_{ijw}$	673,049	4.01	3.38	0	12
Migrants$_{ij}$	673,049	1,527,721	3.27e+07	0	1.43e+09
Distance$_{ij}$	673,049	6,660	4,931	.53	19,951
Domestic$_{ij}$	673,049	0.02	0.16	0	1
Colony$_{ij}$	673,049	0.04	0.19	0	1
Contiguity$_{ij}$	673,049	0.05	0.21	0	1

Summary statistics for the sample of country pairs with the number of matches greater than zero. Variables are not in logs. Source: various data sets, own calculations (2023).

Figure 2: Number of matches player by hour of the destination country



This figure plots the number of matches by hour. Source: aoe2insights.com and aoe2.net (2023), own calculations.

2.3 Trade Data

In order to see whether online interactions can explain trade in physical goods between countries, we look at two measures. The first measure is the total number of player-matched games for each country pair. The second measure is the total number of server-matched games for each country pair.

To test our hypothesis, we use Comtrade export data for the year 2017 at the HS 6-digit level. We use conversion tables to convert the HS2007 classification to the HS2002 classification. This allows us to match the HS codes with the corresponding SITC Rev. 2 codes to classify commodities into goods traded on organised exchanges (homogeneous goods), reference priced (with reference prices displayed in specialized publications) and differentiated products based on the Rauch (1999) classification. If the number of matches is a measure for cultural proximity and network effects, we would expect that it matters more for differentiated goods than for the two other types of goods. As a robustness check, we use the "Micro-D" classification proposed by Bernini et al. (2018) to classify goods into differentiated and undifferentiated products. Micro-D classification is defined at a lower level of aggregation (HS 6-digit level in our case) and therefore is more accurate than the Rauch classification in the identification of differentiated products. In addition, we match our trade data with CEPII and migration data described above.[17] Table 3 presents the summary statistics for the sample in our trade regressions.

Table 3: Summary statistics (trade regressions)

	Obs.	Mean	Std. Dev.	Min	Max
ln Exports$_{ijk}$	6,106,200	9.91	3.27	0	24.70
ln Distance$_{ij}$	6,106,200	8.04	1.10	4.09	9.89
Contiguity$_{ij}$	6,106,200	0.12	0.30	0	1
Common language$_{ij}$	6,106,200	0.15	0.36	0	1
Colony$_{ij}$	6,106,200	0.06	0.24	0	1
ln Migrants$_{ij}$	6,106,200	6.46	4.29	0	16.26
ln No Matches$_{ij}$ (player matched)	6,106,200	3.76	2.48	0	10.31
ln No Matches$_{ij}$ (server matched)	6,106,200	5.49	3.01	0	12.13
Liberal$_k$	6,106,200	1.29	0.53	1	3
Conservative$_k$	6,106,200	1.24	0.48	1	3
Micro-D$_k$	6,106,200	1.27	0.44	1	2

Summary statistics for the sample in Table 7. Liberal and conservative are the classifications described in Rauch (1999), Micro-D is the classification described in Bernini et al. (2018). Source: various data sets, own calculations (2023).

3 Theoretical Framework and Empirical Approach

In this section we explain our theoretical framework for online interactions (Subsection 3.1) and for physical trade (Subsection 3.2).

[17]Similarly to the online interactions, we use the number of migrants from the origin in the destination of the export flow.

3.1 Online Interactions and Trade Costs

Starting with the structural gravity model (see e.g., Yotov et al., 2016), in order to adapt it to online interactions, we use the separability property of the model and focus on a single variety (game). Second, once a representative player purchased the game, the price of each match is the same across countries. Hence, $p_i = p_j$ and $p_{ij} = \tau_{ij}$ where τ_{ij} are trade costs. Our model is therefore:

$$Y_{ijwt} = \frac{N_{iwt}N_{jwt}}{N}\tau_{ij} \tag{1}$$

where Y_{ijwt} is the number of matches played between players in countries i and j on weekday w and hour t. N_{iwt} is the number of players in the country i on weekday w and hour t, N_{jwt} is the number of players in the country j on weekday w and hour t, N is the total number of players, and τ_{ij} are trade costs between the two countries, that can be approximated by a set of gravity variables. In our case, we can safely assume that physical transport costs are zero, so these variables approximate cultural and informational familiarity.

After log linearizing, our model becomes:

$$lnY_{ijwt} = lnN_{iwt} + lnN_{jwt} - lnN + ln\tau_{ij} \tag{2}$$

Which yields the following gravity equation:

$$Y_{ijwt} = \exp\left[\alpha + \beta\tau_{ij} + \nu_{iwt} + \gamma_{jwt}\right] \times \varepsilon_{ijwt} \tag{3}$$

where Y_{ijwt} is the number of matches played between players in countries i and j on weekday w and hour t. τ_{ij} are the gravity variables, ν_{iwt} and γ_{jwt}, are different sets of fixed effects (with the country of player one (i) $\times$ weekday (w)$\times$ hour (t); and the country of player two (j) $\times$ weekday (w)$\times$ hour (t) forming our preferred specification which ultimately absorbs the number of players in each country).[18] Finally, ε_{ijwt} is the error term that we cluster by country pair (direction specific).

We first estimate our model using all matches played between two players. In the second step, we exclude server-matched games and estimate our model for player-matched games only. In Appendix C in Table C.1

[18]Note that in the absence of prices and expenditures, our two sets of fixed effects control for unobserved characteristics that vary over time for each of the countries, rather than the multilateral resistance terms derived from a standard structural gravity model.

we examine whether the civilization choices show the same pattern as found for matches played.

When estimating the model in equation 3 we encounter one of the common issues in the trade literature: zero matches between countries. For this reason, we use the Poisson pseudo-maximum likelihood estimation method (PPML) as proposed by Silva and Tenreyro (2006) and recently reviewed by Santos Silva and Tenreyro (2022). To do so, we construct a full matrix of all possible origin-destination-weekday-hour combinations. The full data set contains 6,001,128 observations corresponding to 189 countries of origin $\times$ 189 countries of destination $\times$ 7 days $\times$ 24 hours.[19] Out of these, only 673,049 combinations are different from zero. In addition, for completeness, we estimate our model without zeros (excluding country pairs that did not play together on weekday w and hour t), using highly-dimensional fixed effects OLS, to check whether our results still hold. We present our results in Appendix E.

3.2 Online Interactions and Physical Trade

After establishing that online interactions respond to usual trade costs proxies, we use the total number of player-matched games server-matched as a measure of countries' cultural proximity in regular trade flows. We begin with a standard structural gravity model:

$$Trade_{ijk} = \frac{Y_{ik}E_{jk}}{Y_k}\left(\frac{t_{ij}}{P_{jk}\pi_{ik}}\right)^{1-\sigma_k} \tag{4}$$

Where $Trade_{ijk}$ is the value of exports between i (origin) and j (destination) in good k in 2017, which depends on the hypothetical level of trade between the two countries $\frac{Y_{ik}E_{jk}}{Yk}$ based on the relative expenditure E and production Y in the two countries when trade costs are zero (and therefore the second term $\left(\frac{t_{ij}}{P_{jk}\pi_{ik}}\right)^{1-\sigma_k}$ is equal to one).

The second term captures the trade cost in terms of factors related to the physical distance between countries i and j. In our setting, $t_{ij} = c_{ij} * d_{ij}$, where c_{ij} captures factors related to cultural and informational familiarity (in our case the number of online interactions) and d_{ij} captures traditional gravity variables. P_{jk} and π_{ik} capture multilateral resistance terms, and σ_k is the elasticity of substitution.

We estimate the following gravity equation:

$$Exports_{ijk} = \exp\left[\alpha + \beta d_{ij} + \delta c_{ij} + \nu_{ik} + \gamma_{jk}\right] \times \varepsilon_{ijk} \tag{5}$$

[19]We allow for domestic matches.

Where d_{ij} are the traditional gravity variables, c_{ij} is the natural log (plus one) of the number of games played between players in countries i and j, ν_{ik} is a set of origin-product fixed effects, γ_j is a set of destination-product fixed effects and ε_{ijk} is the error term (clustered at country-pair level). Using these effects allows us to control for multilateral resistances, and potentially for any other observable and unobservable characteristics that vary at the ik or jk level (Yotov et al. 2016). In subsequent specifications, we add a set of interaction effects that allow us to test whether differentiated products are more affected by our measure than homogeneous goods or reference-priced goods. Due to the size of our dataset, we estimate it using a highly-dimensional fixed effects OLS rather than PPML.

4 Results

We begin this section by discussing the results of our online interactions regressions (Subsection 4.1). We then discuss the robustness checks (Subsection 4.2). In the last Subsection (4.3) we discuss the results of the online interactions and trade.

4.1 Online Interactions

Table 4 presents the results for all matches between two players. We begin with our baseline specification in column (1). Column (2) adds a quadratic term of the time difference between the countries in order to control for possible non-linearities. In column (3) we check if our results are robust to a different set of fixed effects. In the last column (4) we use a sample of foreign countries only.[20] Note that the number of observations varies slightly across specifications using different fixed effects. This is due to the specificity of the command we use; *ppmlhdfe* for PPML which drops observations when the set of fixed effects explains all the variation (singletons). This also explains why the number of observations corresponds roughly to half of our initial matrix.

Across all columns, we find a significant negative effect of the distance, ranging from -0.257 in column (2) to -0.452 in column (4). The magnitude of this effect is somewhat smaller than in traditional gravity estimates (see e.g., Disdier and Head 2008), with around half of the usual effect in column (4), which is the most comparable to typical foreign trade estimates.[21] It is also smaller in absolute terms than the negative

[20] As in a traditional gravity model estimations.

[21] Note that this is not surprising as the size of the distance effects is usually smaller when intra-national flows are taken into

effect found by Blum and Goldfarb (2006) (larger than unity across categories) and Hellmanzik and Schmitz (2015) around -0.7. The general significance, however, is somewhat surprising as we did not expect physical distances to matter for this type of interactions.

Moving on to other variables, we find a strong positive coefficient on the domestic dummy suggesting that players choose to play against other players from the same country (home bias). This effect is in line with previous literature on online exchanges (see e.g., Broocks and Studnicka 2021). The magnitude of the domestic dummy is very large around 185%. The effects of other gravity variables such as common language or the number of migrants are all of the expected signs but their effects are rather small. The effect of the time difference is very small indicating that it does not play an important role in determining the number of matches played between two countries. In column (3) we include a different set of fixed effects (origin-weekday and destination-weekday) and additional control variables. Since our results are largely unchanged, and we do not find an important non-linear effect of the time difference in column (2), columns (2) and (4) form our baseline specifications.

Table 5 presents the same set of results as Table 4 but excludes games matched automatically by the server.[22] Here player one hosts a game, and player two picks the game to join from a list of games (lobby). The information a potential player gets is the distance to the server (in milliseconds), the language setting of the other player, and the player name (see Figure D.1 in Appendix D). Based on this information, a potential player can discriminate with regard to these characteristics. In Figure 3, we show the colour-coded pings from the game with corresponding physical airline distances from the server. Based on this, there is no specific reason, why the distance to the server should matter, particularly for shorter distances.[23]

The coefficients on the gravity variables (including the distance) are of a larger magnitude. The effect of the domestic dummy is more than 700%, suggesting even stronger home bias when players are not matched by the server. This could be due to friends playing with each other, rather than online acquaintances. The effect of a common language is also much stronger. This suggests that this is indeed a factor that players use as a choice variable. The effect of the stock of migrants is larger in column (4) suggesting that some players may choose to play with people from their country of origin.

account. See e.g., Yotov et al. (2016) and Yotov (2022).

[22]The algorithm used to automatically match players might include the distance to improve the experience for players.

[23]A comparison to faster games can be found in Appendix B.

Table 4: Baseline results (all matches)

	(1)	(2)	(3)	(4)
ln Distance$_{ij}$	-0.317***	-0.257***	-0.322***	-0.452***
	(0.037)	(0.037)	(0.036)	(0.030)
Domestic$_{ij}$	1.048***	1.030***	1.049***	
	(0.112)	(0.103)	(0.109)	
Common language$_{ij}$	0.113*	0.105*	0.111*	-0.056
	(0.063)	(0.061)	(0.063)	(0.048)
Colony$_{ij}$	-0.020	0.063	-0.027	0.012
	(0.072)	(0.080)	(0.073)	(0.055)
Contiguity$_{ij}$	0.093	0.049	0.093	0.011
	(0.075)	(0.070)	(0.074)	(0.051)
ln Migrants$_{ij}$	0.047***	0.046***	0.045***	0.044***
	(0.010)	(0.009)	(0.010)	(0.007)
Time difference$_{ij}$	-0.072***	-0.195***	-0.075***	-0.060***
	(0.010)	(0.027)	(0.010)	(0.009)
Time difference sq.$_{ij}$		0.012***		
		(0.002)		
Hour origin$_{it}$			-0.002***	
			(0.0005)	
Hour destination$_{jt}$			-0.003***	
			(0.0006)	
ln Players origin$_{iwt}$			0.641***	
			(0.011)	
ln Players destination$_{iwt}$			0.650***	
			(0.011)	
Constant	7.508***	7.175***	-3.58***	8.294***
	(0.339)	(0.310)	(0.376)	(0.256)
Observations	3,093,251	3,093,251	3,876,756	2,948,495
Origin×weekeday×hour FE	yes	yes		yes
Dest.×weekeday×hour FE	yes	yes		yes
Origin×weekeday FE			yes	
Dest.×weekeday FE			yes	
Pseudo R2	0.964	0.965	0.965	0.952
Log-likelihood	-3.598e+06	-3.523e+06	-3.700e+06	-2.484e+06

Dependent variable: number of matches between a country pair by weekday-hour. The country of player one is the country of origin, country of player two is the country of destination. Estimation method: PPML. Standard errors are clustered by country pair. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Baseline results (excluding server-matched games)

	(1)	(2)	(3)	(4)
ln Distance$_{ij}$	-0.417***	-0.378***	-0.417***	-0.485***
	(0.052)	(0.042)	(0.051)	(0.054)
Domestic$_{ij}$	3.862***	3.792***	3.868***	
	(0.138)	(0.114)	(0.136)	
Common language$_{ij}$	0.918***	0.897***	0.906***	0.639***
	(0.068)	(0.069)	(0.068)	(0.057)
Colony$_{ij}$	0.109	0.214**	0.108	0.152*
	(0.075)	(0.088)	(0.076)	(0.078)
Contiguity$_{ij}$	0.204**	0.126	0.208**	0.109*
	(0.101)	(0.094)	(0.100)	(0.066)
ln Migrants$_{ij}$	0.042**	0.039**	0.039**	0.067***
	(0.017)	(0.016)	(0.016)	(0.013)
Time difference$_{ij}$	0.072***	-0.065*	0.069***	0.055***
	(0.021)	(0.037)	(0.021)	(0.018)
Time difference sq.$_{ij}$		0.012***		
		(0.005)		
Hour origin$_{it}$			-0.0005	
			(0.002)	
Hour destination$_{jt}$			-0.0001	
			(0.002)	
ln Players origin$_{iwt}$			0.572***	
			(0.011)	
ln Players destination$_{iwt}$			0.533***	
			(0.011)	
Constant	5.519***	5.410***	-3.100***	5.453***
	(0.502)	(0.440)	(0.487)	(0.474)
Observations	2,256,392	2,256,392	2,976,966	1,885,706
Origin×weekeday×hour FE	yes	yes		yes
Dest.×weekeday×hour FE	yes	yes		yes
Origin×weekeday FE			yes	
Dest.×weekeday FE			yes	
Pseudo R2	0.973	0.973	0.974	0.814
Log-likelihood	-1.237e+06	-1.225e+06	-1.239e+06	-925354

Dependent variable: number of matches (player-matched) between a country pair by weekday-hour. The country of player one is the country of origin, country of player two is the country of destination. Estimation method: PPML. Standard errors are clustered by country pair. *** p<0.01, ** p<0.05, * p<0.1.

Figure 3: Age of Empires II: Definitive Edition, ping

This figure shows the ping (in milliseconds) from Dublin to various servers across the globe.
Source: Age of Empires II: Definitive Edition (2023), own calculations.

4.2 Robustness Checks: Online Interactions

In this section, we address two potential issues arising in our analysis so far. The first issue is the fact that players could be seen as randomly assigned to their player slot. For example, if two friends have agreed to play a game there is no difference whether player one or two is joining the game. So far, we have ignored this issue and assumed that the player's one country is the country of origin and the player's two country is the country of destination. We now remove this assumption using two approaches. First, we assume no specific direction of matches. We sum the number of matches between each country pair and estimate our model with country dummies (this approach is not direction specific: "Remove direction"). Note that the inclusion of dummies leads to multicollinearity with the domestic dummy and hence it is omitted. Second, we split the number of matches equally between each country pair (direction specific: "50/502). In both specifications, we exclude served-matched games. Our results are in line with our previous findings, confirming that the physical distance matters even in a low trade costs environment with and without considering the direction of flows. We present them in the first two columns of Table 6.

Another issue is the question of distance. While Figure 3 shows that the ping indicator in the game suggests that distance shouldn't matter, we have not tested this explicitly in the previous section. In the third column of Table 6, we show the results of restricting the sample of countries to those with (log)

Table 6: Robustness checks (excluding server-matched games)

	(1) Remove direction	(2) 50/50	(3) Short Distance
ln Distance$_{ij}$	-0.156***	-0.487***	-0.265***
	(0.038)	(0.054)	(0.043)
Domestic$_{ij}$		2.985***	3.806***
		(0.141)	(0.156)
Common language$_{ij}$	0.798***	0.849***	1.049***
	(0.085)	(0.058)	(0.095)
Colony$_{ij}$	0.061	0.145**	-0.022
	(0.100)	(0.065)	(0.159)
Contiguity$_{ij}$	0.361***	0.146*	0.143
	(0.120)	(0.085)	(0.101)
ln Migrants sum $_{ij}$	0.014		
	(0.013)		
ln Migrants$_{ij}$		0.053***	0.087***
		(0.019)	(0.023)
ln Players sum$_{ij}$	0.850***		
	(0.037)		
Time difference$_{ij}$	-0.026	0.065***	0.024
	(0.026)	(0.016)	(0.034)
Constant	4.907***	6.655***	3.928***
	(0.455)	(0.550)	(0.503)
Observations	3,016,440	1,170,663	735,023
Origin×weekeday×hourFE	yes	yes	yes
Dest.×weekeday×hour FE	yes	yes	yes
Pseudo R2	0.973	0.9797	0.986
Log-likelihood	-1328128.67	-844244.62	
-522821.125			

Dependent variable: column (1) total number of player-matched games between each country pair, column (2) splits the number of matches equally between each country pair, column (3) number of matches. Estimation method: PPML. Standard errors are clustered by country pair. *** p<0.01, ** p<0.05, * p<0.1.

distances below 8.5, roughly 5,000km. As before, we use the number of player-matched games. The distance coefficient stays negative but has a somewhat lower magnitude when compared to the regressions in Table 5. This continued significance provides us with confidence that while the significance of physical distance between players may be partly driven by player experience (and hence physical distance), there are other factors like cultural proximity and network effects at play as well, contributing to the overall relevance of the distance variable. Based on these results in the next section, we consider the number of matches played as a new measure of the aforementioned cultural and network factors.

4.3 Results: Online Interactions and Trade

In this section, we explore whether the number of online interactions (games) can be used as a predictor of regular trade flows. Our intuition is that the number of matches played between two countries can act as a measure of cultural proximity and network effects in the gravity equation. In addition, contrary to Disdier et al. (2010) who use trade in cultural goods as a measure of cultural proximity, the number of matches is completely exogenous as it is unrelated to trade in goods. Our measure is also similar to the measure used by Guiso et al. (2009) who use trust as an additional measure to capture cultural differences across countries and Huang (2007) who use historic and systematic differences of uncertainty-aversion across countries in their gravity estimates.

As a first step, we calculate two measures of the number of online interactions between a country pair. The first measure uses the total number of player-matched games, the second one is the total number of server-matched games. As before, since player two decides to play with player one for player-matched games, we choose player two as the destination (country j) and player one as the origin (country i). We replace missing values with zero (no matches played) and use the natural log plus one of each of the variables. Our hypothesis is that player-matched games can help explain physical trade patterns as they capture cultural familiarity and network effects which are not directly captured by traditional gravity variables. Our second trade-related hypothesis is that differentiated goods are more prone to be affected by cultural familiarity and network effects than homogeneous goods and reference-priced goods.

Table 7 presents our results. Column (1) contains the traditional gravity results, that are in line with the literature. The number of player-matched games (in columns (2)-(5)) is positive and significant, with an elasticity of around 0.28. In addition, including this variable reduces the distance effect in absolute terms by around 10% and decreases the magnitude of other gravity variables. These findings are in line with the assumption that our variable captures cultural proximity and network effects. The significance of our measure of online interactions also confirms our hypothesis that they predict physical trade. The effect of this variable is substantially reduced for homogeneous goods and reference-priced goods as well as for undifferentiated goods (columns (3)-(5)). This is in line with our hypothesis that trade in differentiated goods should be more positively related to the number of online interactions than homogeneous goods. It is of note that there is still an effect even for homogeneous goods.

Table 7: Online interactions (player-matched) and international trade

	(1) baseline	(2) p-matched	(3) liberal	(4) conservative	(5) micro-D
ln Distance$_{ij}$	-1.053***	-0.914***	-0.913***	-0.913***	-0.913***
	(0.022)	(0.023)	(0.023)	(0.023)	(0.023)
Common language$_{ij}$	0.429***	0.317***	0.319***	0.319***	0.319***
	(0.048)	(0.045)	(0.045)	(0.045)	(0.045)
Colony$_{ij}$	0.446***	0.389***	0.389***	0.389***	0.389***
	(0.068)	(0.066)	(0.066)	(0.066)	(0.066)
Contiguity$_{ij}$	0.713***	0.676***	0.679***	0.679***	0.682***
	(0.068)	(0.065)	(0.065)	(0.065)	(0.065)
ln Migrants$_{ij}$	0.073***	0.053***	0.053***	0.053***	0.053***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
ln P-Matches$_{ij}$		0.268***	0.282***	0.281***	0.291***
		(0.016)	(0.016)	(0.016)	(0.016)
Reference$_k \times$ln P-Matches$_{ij}$			-0.050***		
			(0.007)		
Homogeneous$_k \times$ln P-Matches$_{ij}$			-0.151***		
			(0.015)		
Reference$_k \times$ln P-Matches$_{ij}$				-0.063***	
				(0.008)	
Homogeneous$_k \times$*ln P-Matches$_{ij}$				-0.147***	
				(0.016)	
Undiff.$_k \times$ln P-Matches$_{ij}$					-0.101***
					(0.009)
Constant	17.746***	15.770***	15.781***	15.782***	15.793***
	(0.196)	(0.229)	(0.229)	(0.229)	(0.229)
Observations	6,106,200	6,106,200	6,106,200	6,106,200	6,106,200
Exporter×product FE	yes	yes	yes	yes	yes
Importer×product FE	yes	yes	yes	yes	yes
R2	0.608	0.611	0.612	0.612	0.612

Dependent variable: ln(exports). Number of matches is the total number of player-matched games during our sample. Column (3) uses the liberal classification as described in Rauch (1999), column (4) uses the conservative classification. Column (5) uses the Micro-D classification described in Bernini et al. (2018). Estimation method: OLS. Standard errors are clustered by country pair. *** p<0.01, ** p<0.05

Table 8 presents the results of estimating the same model but using server-matched games as an additional measure. Columns (1) and (2) only include the number of server-matched games (S-matches). In this table, we use the liberal Rauch classification of goods. Columns (3) and (4) show that when we include both types of interaction measures, server-matched games (S-matches) become insignificant while player-matched games (P-matches) are still significant. This suggests that player-matched games are a more relevant measure in the gravity model of trade. We thus consider the player-matched games as a better measure of online interactions as they better capture cultural familiarity and network effects.

Table 8: Online interactions (player-matched and server-matched) and international trade

	(1) s-matched	(2) liberal	(3) both	(4) both liberal
ln Distance$_{ij}$	-0.942***	-0.943***	-0.901***	-0.900***
	(0.027)	(0.027)	(0.026)	(0.026)
Common language$_{ij}$	0.451***	0.451***	0.324***	0.325***
	(0.048)	(0.048)	(0.045)	(0.045)
Colony$_{ij}$	0.487***	0.486***	0.397***	0.397***
	(0.069)	(0.069)	(0.067)	(0.067)
Contiguity$_{ij}$	0.773***	0.774***	0.686***	0.689***
	(0.069)	(0.069)	(0.066)	(0.066)
ln Migrants$_{ij}$	0.068***	0.068***	0.052***	0.052***
	(0.005)	(0.005)	(0.005)	(0.005)
ln S-Matches$_{ij}$	0.167***	0.175***	0.026	0.019
	(0.019)	(0.019)	(0.021)	(0.022)
ln P-Matches$_{ij}$			0.260***	0.277***
			(0.018)	(0.018)
Reference$_k$ $\times$ln S-Matches$_{ij}$		-0.018**		0.039***
		(0.009)		(0.011)
Homogeneous$_k$ $\times$ln S-Matches$_{ij}$		-0.153***		-0.055***
		(0.016)		(0.021)
Reference$_k$ $\times$ln P-Matches$_{ij}$				-0.065***
				(0.009)
Homogeneous$_k$ $\times$ln P-Matches$_{ij}$				-0.128***
				(0.019)
Constant	15.956***	15.973***	15.550***	15.560***
	(0.305)	(0.305)	(0.292)	(0.292)
Observations	6,106,200	6,106,200	6,106,200	6,106,200
Exporter$\times$product FE	yes	yes	yes	yes
Importer$\times$product FE	yes	yes	yes	yes
R2	0.609	0.609	0.611	0.612

Dependent variable: ln(exports). Columns (1) and (2) include only server-matched games (S-Matches). Columns (3) and (4) include server-marched and player-marched games (P-Matches). Columns (2) and (4) use the liberal classification as described in Rauch (1999). Estimation method: OLS. Standard errors are clustered by country pair. *** p<0.01, ** p<0.05

In Appendix C, we show that similar results can be obtained when using the civilization choices by players instead of their choices of opponents.

5 Conclusion

Gravity model estimates have consistently shown that bilateral exchanges are negatively affected by the distance between countries, making it one of the most robust empirical findings in economics. It is widely accepted that geographic distances and other gravity variables are not only a proxy for physical transport costs but also for all sorts of informal and cultural barriers. In this paper, we use the gravity approach to test the validity of the gravity model in an extremely low trade costs environment of online interactions. Our initial hypothesis, that gravity variables should not matter as there is no real distance, nor language barrier between countries is strongly rejected.

In the second step, we check if the number of matches can be used as an additional variable explaining trade under the assumption that it is a measure of cultural affinity and network effects. We find that the effect of this variable is positive and significant. In addition, when we include this variable in our model, the effect of the distance and other gravity variables is reduced. This is in line with the assumption that it captures cultural proximity and network effects. The confirmation of our hypothesis that differentiated goods should be more impacted by online transactions than homogeneous goods is also in line with this assumption.

Taken together, our findings suggest that the gravity equation is a good predictor of human interactions that go beyond trade in goods and services. Furthermore, online interactions are clearly related to physical trade and can provide a new measure to capture network effects and all sorts of cultural differences between people.

References

Alaveras, G. and Martens, B. (2015). International trade in online services. *Available at SSRN 2670614*.

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *American Economic Review*, 69(1):106–116.

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

Bernini, F., González, J., Hallak, J. C., and Vicondoa, A. (2018). The micro-D classification: A new approach to identifying differentiated exports. *Economia*, 18(2):59–85.

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

Broocks, A. and Studnicka, Z. (2021). Gravity and trade in video on demand services. *University College Dublin, WP21/27*.

Chaney, T. (2014). The network structure of international trade. *American Economic Review*, 104(11):3600–3634.

Disdier, A.-C. and Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and Statistics*, 90(1):37–48.

Disdier, A.-C., Tai, S. H., Fontagné, L., and Mayer, T. (2010). Bilateral trade of cultural goods. *Review of World Economics*, 145:575–595.

Guiso, L., Sapienza, P., and Zingales, L. (2009). Cultural biases in economic exchange? *The Quarterly Journal of Economics*, 124(3):1095–1131.

Hellmanzik, C. and Schmitz, M. (2015). Virtual proximity and audiovisual services trade. *European Economic Review*, 77:82–101.

Huang, R. R. (2007). Distance and trade: Disentangling unfamiliarity effects and transport cost effects. *European Economic Review*, 51(1):161–181.

Lendle, A., Olarreaga, M., Schropp, S., and Vézina, P.-L. (2016). There goes gravity: eBay and the death of distance. *The Economic Journal*, 126(591):406–441.

Mayer, T. and Zignago, S. (2011). Notes on CEPII's distances measures: The GeoDist database.

McCallum, J. (1995). National borders matter: Canada-US regional trade patterns. *American Economic Review*, 85(3):615–623.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics*, 48(1):7–35.

Reuters (2021). Three hours a week: Playtime's over for China's young videogamers.

Santos Silva, J. and Tenreyro, S. (2022). The log of gravity at 15. *Portuguese Economic Journal*, 21(3):423–437.

Silva, J. S. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658.

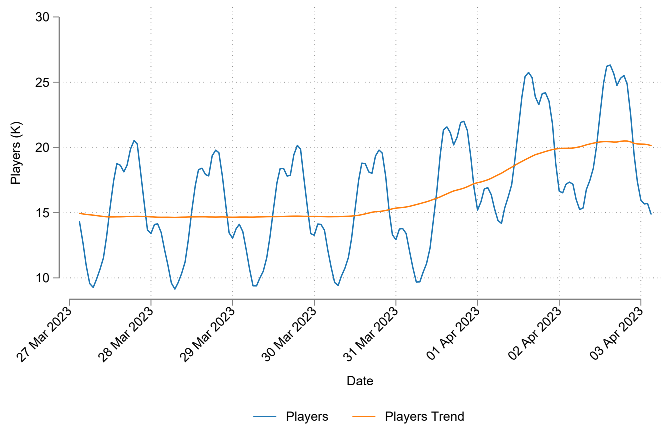
Yotov, Y. (2022). Gravity at Sixty: The Bijou of Trade. Technical report, LeBow College of Business, Drexel University.

Yotov, Y. V., Piermartini, R., Larch, M., and Monteiro, J.-A. (2016). *An advanced guide to trade policy analysis: The structural gravity model*. WTO iLibrary.

Appendix

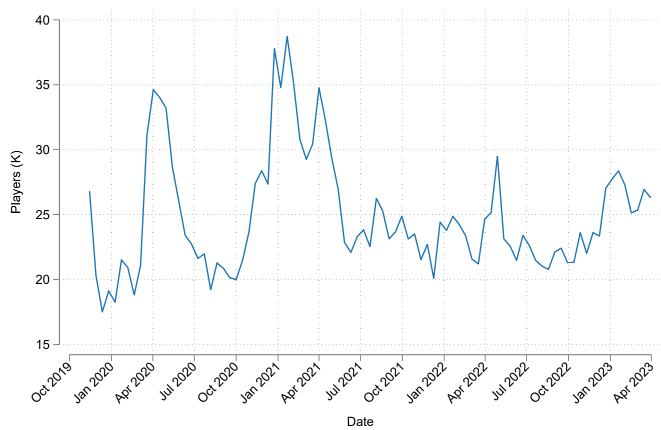
A. Additional Statistics

Figure A.1: Hourly Players (26 March - 4 April 2023)



This figure plots the number of hourly players. Source: steamdb (2023), own calculations

Figure A.2: Monthly Players (14 November 2019 - 14 March 2023)

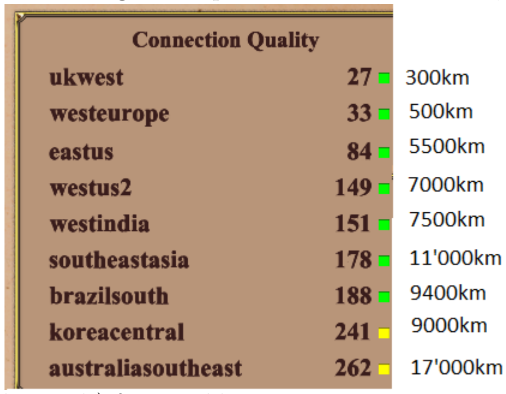


This figure plots the number of monthly players. Source: steamdb (2023), own calculations

B. Comparison of ping times

Figure B.1 presents the connection indicator from Dublin to various servers across the globe.

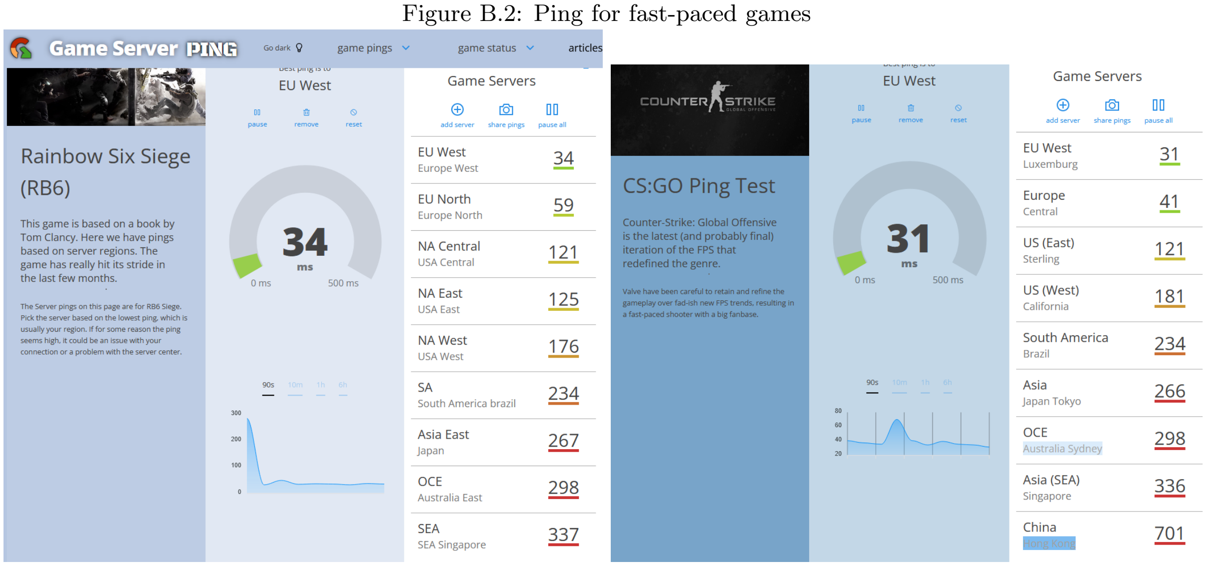
Figure B.1: Age of Empires II: Definitive Edition, ping



This figure shows the ping (in milliseconds) from Dublin to various servers across the globe.
Source: Age of Empires II: Definitive Edition (2023)

Even if relatively short distances have a much faster connection, the ping indicator remains green for considerable distances like Brazil or India. This implies that for shorter distances (e.g., below 5,000km), distance does not make the game experience worse for players. These contrast with the pings for faster-paced games such as first-person shooters, where high pings would generally be regarded as bad as shown in Figure B.2.

Figure B.2: Ping for fast-paced games



This figure shows the ping (in milliseconds) from Dublin to various servers across the globe. Source: https://gameserverping.com/csgo and https://gameserverping.com/rb6 (2023)

C. Civilizations

Each player can pick among 42 civilizations.[24] Of these 42 civilizations, we were able to map all except for the Goths and Huns to 36 existing countries.[25]

Each civilization possesses specific strengths and weaknesses, accompanied by culturally themed military units and artwork. For example, the Britons have particularly powerful archers including a longbow unit, while Franks have particularly strong cavalry. Both civilizations have a Western European art set, meaning their buildings look similar. With very few exceptions, all civilizations have access to the same buildings, albeit with variation in their appearance, and for most strategies, the first 10-15 minutes of the game are very similar. This means that even if a player starts with the civilization they most identify with, they can easily learn to play with other civilizations as well.

In addition to picking specific civilizations from an alphabetical menu, players can also select a random civilization and a mirror civilization. The former option randomly picks one of the 42 civilizations while the latter option picks the same civilization as the opponent. Unfortunately, we only know the actual civilizations played and cannot infer in which way a civilization was picked. Anecdotally, the random option is picked quite frequently (for some players more often than 80% of games) which might be related to the similarities between civilizations. A consequence of players picking random civilizations is that the distance coefficient in our regressions likely becomes smaller in magnitude. Due to this, finding even small but significant coefficients constitute strong evidence in favour of players discriminating while choosing a civilization.

One additional caveat to note is that some civilizations were introduced for purchase during our sample period.[26] This means that not all players had access to all civilizations at the beginning of our sample or purchased additional civilizations later on.

Despite these difficulties, we run regressions pooled across the entire sample period as well as pooled across players one and two to check whether distance matters for players while picking civilizations. In addition, in order to ensure that our results are not driven by the specific countries selected, we run a regression for games that were not matched by the server (player-matched games) and include only the country pairs also

[24]These are: Aztecs, Bengalis, Berbers, Bohemians, Britons, Bulgarians, Burgundians, Burmese, Byzantines, Celts, Chinese, Cumans, Dravidians, Ethiopians, Franks, Goths, Gurjaras, Hindustanis, Huns, Incas, Italians, Japanese, Khmer, Koreans, Lithuanians, Magyars, Malay, Malians, Mayans, Mongols, Persians, Poles, Portuguese, Saracens, Sicilians, Slavs, Spanish, Tatars, Teutons, Turks, Vietnamese, and Vikings.

[25]A list of correspondences is available upon request.

[26]Namely, in January 2021, Burgundians and Sicilians were added, in August 2021, Poland and Bohemians were added, and in April 2022, Indians were split into Hindustanis, Gurjaras, Bengalis and Dravidians.

Table C.1: Civilization regressions (player-matched games)

	(1) baseline	(2) baseline civ.	(3) foreign	(4) no MX and PE	(5) no FR
ln Distance$_{ij}$	-0.549***	-0.037**	-0.052***	-0.037**	-0.039**
	(0.084)	(0.018)	(0.014)	(0.018)	(0.018)
Domestic$_{ij}$	3.322***	0.552***		0.559***	0.613***
	(0.160)	(0.090)		(0.095)	(0.099)
Common language$_{ij}$	0.820***	0.034	0.041	0.003	0.026
	(0.117)	(0.029)	(0.028)	(0.029)	(0.031)
Colony$_{ij}$	-0.002	0.062**	0.059*	0.072**	0.083**
	(0.115)	(0.031)	(0.032)	(0.032)	(0.035)
Contiguity$_{ij}$	0.070	-0.027	-0.023	-0.026	-0.050
	(0.097)	(0.030)	(0.027)	(0.033)	(0.031)
ln Migrants$_{ij}$	0.078***	-0.002	-0.007**	-0.002	-0.002
	(0.010)	(0.004)	(0.003)	(0.004)	(0.004)
Time difference$_{ij}$	0.030	-0.005	-0.002	-0.004	-0.005
	(0.029)	(0.004)	(0.003)	(0.005)	(0.004)
Constant	11.027***	9.412***	9.565***	9.400***	9.342***
	(0.605)	(0.152)	(0.119)	(0.153)	(0.154)
Observations	6,084	6,552	6,517	6,154	6,335
Origin FE	yes	yes	yes	yes	yes
Dest. FE	yes	yes	yes	yes	yes
Pseudo R2	0.994	0.994	0.994	0.993	0.993
Log-likelihood	-141016	-96762	-84505	-92047	-91474

Dependent variable: number of matches. Column (1) presents the results for matches between players from the same countries as available civilisations, columns (2)-(5) present the results for the selection of civilizations. All regressions exclude server-matched games. Estimation method: PPML. Standard errors are clustered by country pair. *** p<0.01, ** p<0.05, * p<0.1.

included in the civilization regressions (a subsample of the specification in column (2) in Table 5).

Table C.1 presents the results. Column (1) shows that distance is significant for games played between two people from the same country pair as available civilizations. The coefficients are broadly in line with the results in Table 5. The remaining four columns show the results for the civilization choices. Column (2) has the same control variables and sample country pairs as column (1) with the distance effect still negative and significant. As expected, probably due to the random choice option, the coefficient is much smaller in absolute terms than the coefficients Tables 4-5 in the main text.

Columns (3)-(5) restrict the sample for additional robustness checks. Column (3) removes the home country as a choice, column (4) excludes civilizations that do not have cavalry and hence have a slightly different play style, and column (5) excludes the most popular civilization (Franks/France) from the regression. In all three cases, physical distance matters for the civilization choice. We run additional robustness checks such as adding a variable capturing the alphabetical ordering in the civilizations menu and running placebo regressions. Our results still hold and are available upon request. Unfortunately, the sample would get too small if we focused on specific strategies (e.g., the best archer civilizations). Nevertheless, our results suggest that cultural preferences play a significant role in the selection of civilizations, and geographical distance can serve as a useful proxy for capturing these distinctions.

Next, we use civilization choices as a predictor of bilateral trade. We test whether the number of players from one country picking specific civilizations can predict trade between the two countries. Here the country of the origin of exports (i) is the country of the civilization picked, and the country of destination (j) is the country of the player.[27]

There are two caveats to this analysis. First, this dramatically reduces our sample as there are only around 40 countries to which the civilizations map. Second, players could choose a random civilization as well, which increases our standard errors. Despite these caveats, we anticipate a significant effect. We present the results in Table C.2. They are largely in line with the results in Table 7.

One exception to this are the interaction effects. It does not appear that civilization choices have any different effect for homogeneous goods or differentiated goods. Compared with the player-matched regressions, it is also the case that the number of times specific civilizations were picked do not substantially impact any of the gravity variables.

[27]We believe that this corresponds to the direction of export flows.

Table C.2: Civilization choices and international trade

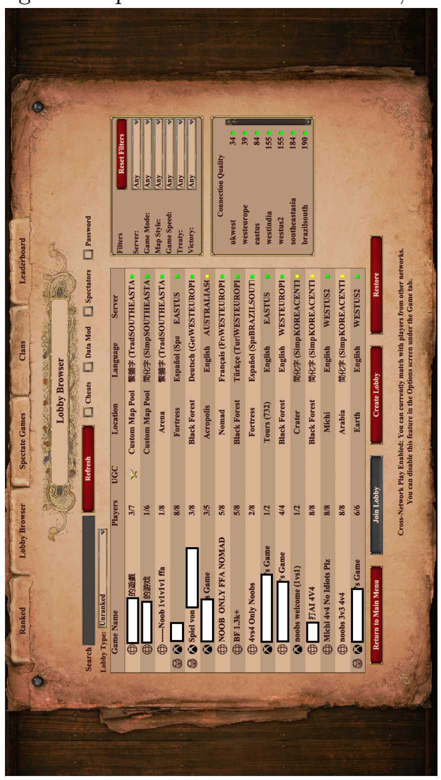
	(1) baseline	(2) matches	(3) liberal	(4) conservative	(5) micro-D
ln Distance$_{ij}$	-1.069***	-1.066***	-1.066***	-1.066***	-1.066***
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
Common language$_{ij}$	0.650***	0.630***	0.630***	0.629***	0.630***
	(0.092)	(0.092)	(0.092)	(0.092)	(0.092)
Colony$_{ij}$	0.542***	0.554***	0.554***	0.554***	0.554***
	(0.100)	(0.099)	(0.099)	(0.099)	(0.099)
Contiguity$_{ij}$	0.662***	0.650***	0.650***	0.649***	0.650***
	(0.096)	(0.094)	(0.094)	(0.094)	(0.094)
ln Migrants$_{ij}$	0.069***	0.068***	0.068***	0.068***	0.068***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
ln Matches$_{ij}$		0.197***	0.191***	0.189***	0.191***
		(0.043)	(0.043)	(0.043)	(0.043)
Reference$_k$×ln Matches$_{ij}$			0.024		
			(0.023)		
Homogeneous$_k$×ln Matches$_{ij}$			0.044		
			(0.056)		
Reference$_k$×ln Matches$_{ij}$				0.036	
				(0.029)	
Homogeneous$_k$×ln Matches$_{ij}$				0.097*	
				(0.059)	
Undiff.$_k$×ln Matches$_{ij}$					0.026
					(0.029)
Constant	18.252***	17.180***	17.169***	17.163***	17.167***
	(0.313)	(0.403)	(0.404)	(0.404)	(0.404)
Observations	3,421,908	3,421,908	3,421,908	3,421,908	3,421,908
Exporter×product FE	yes	yes	yes	yes	yes
Importer×product FE	yes	yes	yes	yes	yes
R2	0.640	0.640	0.640	0.640	0.640

Dependent variable: ln(exports). Number of matches is the total number of civilization choices (i) by players in j. Column (3) uses the liberal classification as described in Rauch (1999), column (4) uses the conservative classification. Column (5) uses the Micro-D classification described in Bernini et al. (2018). Estimation method: OLS. Standard errors are clustered by country pair. *** p<0.01, ** p<0.05

D. Lobby

In the main text, we mentioned that players can discriminate with regard to the server location, the match name and the language setting of the other player in games that are not server-matched. Figure D.1 shows a screenshot of the lobby browser where players select which game they want to join.

Figure D.1: Age of Empires II: Definitive Edition, Lobby Browser



This figure shows the lobby of the game. Source: Age of Empires II: Definitive Edition (2023)

E. OLS Results

Table E.1 presents the results of the OLS regressions. The distance coefficient remains highly significant and negative and the other control variables have the same signs as the PPML results in Table 4.

Table E.1: Baseline results OLS (all matches)

	(1)	(2)	(3)	(4)
ln Distance$_{ij}$	-0.269***	-0.212***	-0.263***	-0.286***
	(0.016)	(0.018)	(0.015)	(0.016)
Domestic$_{ij}$	2.120***	2.215***	2.111***	
	(0.069)	(0.070)	(0.066)	
Common language$_{ij}$	0.124***	0.132***	0.121***	0.109***
	(0.028)	(0.027)	(0.027)	(0.027)
Colony$_{ij}$	-0.186***	-0.176***	-0.188***	-0.177***
	(0.040)	(0.040)	(0.038)	(0.039)
Contiguity$_{ij}$	0.091**	0.103**	0.101**	0.080*
	(0.042)	(0.042)	(0.041)	(0.042)
ln Migrants$_{ij}$	0.021***	0.020***	0.021***	0.019***
	(0.003)	(0.003)	(0.003)	(0.003)
Time difference$_{ijt}$	-0.058***	-0.121***	-0.056***	-0.058***
	(0.004)	(0.009)	(0.004)	(0.004)
Time difference sq.$_{ij}$		0.005***		
		(0.001)		
Hour origin$_{it}$			-0.004***	
			(0.000)	
Hour destination$_{jt}$			-0.004***	
			(0.000)	
ln Players origin$_{iwt}$			0.485***	
			(0.005)	
ln Players dest$_{jwt}$			0.490***	
			(0.005)	
Constant	3.814***	3.457***	-2.380***	4.016***
	(0.133)	(0.139)	(0.131)	(0.133)
Observations	668,739	668,739	672,958	652,216
Origin×weekeday×hour FE	yes	yes		yes
Dest.×weekeday×hour FE	yes	yes		yes
Origin×weekeday FE			yes	
Dest.×weekeday* FE			yes	
Adjusted R2	0.827	0.828	0.809	0.817

Dependent variable: ln(number of matches) between a country pair by weekday-hour. The country of player one is the country of origin, country of player two is the country of destination. Estimation method: OLS. Standard errors are clustered by country pair. *** p<0.01, ** p<0.05, * p<0.1.

UCD CENTRE FOR ECONOMIC RESEARCH – RECENT WORKING PAPERS

WP22/21 Yota Deli, Manthos D. Delis, Iftekhar Hasan, Panagiotis N. Politsidis, Anthony Saunders: 'Corporate tax changes and credit costs' July2022

WP22/23 Karl Whelan: 'US Taxation of GamblingWinnings and Incentives to Bet' August 2022

WP22/23 Karl Whelan:'Risk Aversion and Favorite-Longshot Bias in a Competitive Fixed-Odds Betting Market' August 2022

WP22/24 Diane Pelly, Orla Doyle: 'Nudging in the workplace: increasing participation in employee EDI wellness events' October 2022

WP22/25 Margaret Samahita, Kevin Devereux: 'Are economics conferences gender-neutral? Evidence from Ireland' October 2022

WP22/26 Michele Gubello, Nora M. Strecker: 'Weakly Progressive: Disproportionate Higher Education Attendance and the Structure of Income Taxes' November 2022

WP22/27 David Madden: 'Mental Health in Ireland During the Covid Pandemic: Evidence from Two Longitudinal Surveys' December 2022

WP22/28 Ronald B. Davies, Zuzanna Studnicka: 'Tariff Evasion, the Trade Gap, and Structural Trade' December 2022

WP23/01 Xidong Guo: 'An Analysis of a Rural Hospital’s Investment Decision under Different Payment Systems' January 2023

WP23/02 Emanuele Albarosa, Benjamin Elsner: 'Forced Migration and Social Cohesion: Evidence from the 2015/16 Mass Inflow in Germany' January 2023

WP23/03 Xidong Guo: 'Hospital Efficiency and Consultants’ Private Practices: Analysing the Impact of a Voluntary Reform' February 2023

WP23/04 Tadgh Hegarty, Karl Whelan: 'Calculating The Bookmaker’s Margin: Why Bets Lose More On Average Than You Are Warned' February 2023

WP23/05 Tadgh Hegarty, Karl Whelan: 'Forecasting Soccer MatchesWith Betting Odds: A Tale of Two Markets' February 2023

WP23/06 Karl Whelan: 'How Do Prediction Market Fees Affect Prices and Participants?' March 2023

WP23/07 Kevin Devereux, Zuzanna Studnicka: 'Local Labour Market Concentration and Wages in Ireland' March 2023

WP23/08 Kevin Devereux, Zuzanna Studnicka: 'When and Where do Minimum Wage Hikes Increase Hours? Evidence from Ireland' April 2023

WP23/09 Karl Whelan 'Fortune’s Formula or the Road to Ruin?The Generalized Kelly CriterionWith Multiple Outcomes' April 2023

WP23/10 Manuel E. Lago: '#MeToo...or not? Do salient shocks affect gender social norms?' April 2023

WP23/11 Cormac Ó Gráda, Tyler Anbinder, Simone A. Wegge: 'Gaming theSystem: The Not-So-Poor and Savings Banks in Antebellum New York' April 2023

WP23/12 Tadgh Hegarty, Karl Whelan: 'Disagreement and Market Structure in Betting Markets: Theory and Evidence from European Soccer' May 2023

WP23/13 Tadgh Hegarty, Karl Whelan: 'Do Gamblers Understand Complex Bets? Evidence From Asian Handicap Betting on Soccer' May 2023

WP23/14 Matthew Amalitinga Abagna: ‘Special Economic Zones and Local Economic Activities in Ethiopia.’ May 2023

WP23/15 Gregory Corcos, Stefanie Haller: 'Importer Dynamics: Do Peers Matter?' July 2023

WP23/16 Karl Whelan: 'How Does Inside Information Affect Sports Betting Odds?' July 2023

WP23/17 Michele Gubello: 'Social Trust and the Support for Universal Basic Income' July 2023

UCD Centre for Economic Research
Email economics@ucd.ie