# Asymmetric Cultural Proximity and Greenfield FDI*

Matteo Fiorini† Giorgia Giovannetti‡ Mauro Lanati§ Filippo Santi¶

March 12, 2018

**Abstract**

This paper investigates the role of asymmetric cultural proximity (CP) on greenfield foreign direct investment (FDI) from an origin to a destination country. We build a conceptual framework that explicitly accounts for cultural attractiveness as an asymmetric dimension within a broad notion of CP. We revisit the existing origin-side theories of bilateral FDI to derive a gravity equation suited for testing the impact of (i) the attractiveness of destination's culture for citizens in the origin country, and (ii) the attractiveness of origin's culture for individuals in the destination economy. While the role of the former direction of CP is well understood in the literature, we propose new mechanisms to rationalize that of the latter. We use exports and imports of cultural goods to proxy for the two directions of asymmetric and time-dependent CP in the same empirical specification. The econometric analysis confirms a positive role of asymmetric CP as a determinant of Greenfield FDI. Moreover, it suggests a stronger investment effect of the origin's culture attractiveness for the destination country. Finally, it provides support for the mechanisms proposed in the theoretical discussion.

**Keywords:** cultural proximity; greenfield FDI; cultural trade; gravity model
**JEL Classification:** F14; F21; F23; Z10

*The authors are grateful to Thomas Zylkin, João Santos Silva, Davide Castellani and Marco Gardini for helpful comments and conversations. We also thank participants at the 19th ETSG conference in Florence, the 12th ISGEP workshop in Reading and the Geneva Trade and Development Workshop at the Graduate Institute in Geneva. Finally, we wish to thank Alessandra Venturini and Bernard Hoekman for supporting this research.

†RSCAS (European University Institute). email: Matteo.Fiorini@EUI.eu
‡University of Florence and RSCAS (European University Institute). email: giorgia.giovannetti@unifi.it
§RSCAS (European University Institute). email: Mauro.Lanati@EUI.eu
¶University of Florence. email: filippo.santi@unifi.it (Corresponding author)

# 1 Introduction

The role of foreign direct investment (FDI) in generating net gains for both origin and destination countries is well documented. The growth-enhancing potential of FDI has spurred an in-depth analysis of its determinants. One of the most robust findings pertains to the cultural relationships between the investing and the receiving country: investment from origin to destination is relatively higher if the two countries share similar cultural traits, such as those embedded in language, religion, ethnicity or genetics (see for instance Blonigen and Piger, 2014). However, economically relevant dimensions of cultural relationships go well beyond the symmetric (and largely time-invariant) nature of proxies capturing the extent to which individuals in two countries speak the same language or share similar genetic traits (Shenkar, 2001; Felbermayr and Toubal, 2010; Tung and Verbeke, 2010). This leads to the question of whether and how asymmetric (and time-dependent) cultural variables, such as preferences for cultural systems or bilateral trust, play out as determinants of investment patterns. The literature here offers only half of the answer. While the seminal contribution by Guiso et al. (2009) has shown that investment increases if individuals in the investing country trust the citizens of the receiving economy, the potential role of the opposite direction of trust is left unexplored. More generally, we lack a comprehensive assessment of the asymmetric dimensions in bilateral cultural relationships as determinants of FDI. Given the premise that the cultural relationship between two countries, say Kenya and the UK, features an asymmetric element such as the appreciation of each other's cultural systems, it is obvious to anyone that the way individuals in Kenya appreciate British culture might be very different from how much Kenyan culture is attractive for the UK. It is equally straightforward that these patterns are likely to change over time. How do these two different and evolving forces affect British FDI in Kenya? Is one more relevant than the other? These are questions that motivate this paper, which represents a first attempt to assess the effect of cultural proximity (CP) on FDI, explicitly accounting for the asymmetric and time-dependent dimensions of CP.

To this end we first provide a simple conceptual framework for the notion of CP. By encompassing contributions from international business scholars and economists, we present a workable definition of CP accounting for multiple dimensions of the cultural relationship between two countries. These include symmetric sharing of common cultural traits as well as asymmetric cultural attractiveness. The latter component is allowed to vary over time. In line with Disdier et al. (2010), we use bilateral trade in cultural goods as a proxy for asymmetric and time-dependent CP. Indeed, the value of imports of cultural goods reflects the attractiveness of the exporter's culture for the importer. Moreover, bilateral cultural trade is correlated with standard, symmetric and time-invariant measures of CP, showing the capacity of this proxy to capture all dimensions of CP. We provide some suggestive evidence of the asymmetry embedded in bilateral cultural relationships with a descriptive exercise, conducted on a broad sample of countries. The perspective on cultural asymmetry embedded in cultural trade data differs from and complements the seminal work by Guiso et al. (2009), where data on bilateral trust are analyzed on a sample of European countries. The variation in cultural relationships that can be captured with trade in cultural goods covers both developed and developing countries, an advantage with respect to other asymmetric

measures which tend to be confined to EU countries. This is particularly relevant when greenfield FDI is the object of interest, as the scale and scope of South-South greenfield FDI is growing at fast pace (UNCTAD, 2017) and North-South and South-North greenfield has increased their size and relevance.

Equipped with a definition and an empirical measure of CP that account for asymmetry and time variation, we investigate the linkages between CP and greenfield FDI. The paper revisits the theories used in the literature to derive gravity equations of greenfield FDI. These are partial-equilibrium, supply-side models that subsume all gravity forces into monitoring and transaction costs which ultimately determine the investment decisions of the multi national enterprise (MNE). In this context we discuss the role played as determinants of investment decisions of both directions of asymmetric CP, i.e. the attractiveness of the culture in the origin country for individuals in the destination and the attractiveness of destination's culture for the origin. On the one hand, we argue that the cultural attractiveness of the destination country plausibly (and exhaustively) operates via the monitoring-transaction cost channel. On the other hand, the cultural attractiveness of the origin country for the destination is likely to play a role also through other channels. If the FDI project is conducted to serve consumers demand in the destination country (i.e. horizontal FDI), the attractiveness of the origin country's culture for (destination) consumers positively affects the value they put on the output of the origin's MNE and therefore increases the payoff of the FDI project. We denote this mechanisms as 'destination consumers demand' channel. Moreover, the realization of an FDI project can be facilitated (or opposed) by political pressures in the destination country. Under the assumption of political accountability, politicians in the destination country will allocate pressures to facilitate FDI projects also according to the degree by which the culture of the origin countries are attractive for the individuals (voters) in the destination (we call this the 'destination political economy' channel). All in all, the monitoring-transaction costs channels and the 'destination-side' mechanisms unambiguously imply a positive role of both directions of asymmetric CP in determining greenfield FDI from the origin to the destination country. However, the assessment of the relative importance of one direction over the other is an empirical matter.

A structural gravity equation, fully consistent with our theoretical discussion, is brought to the data. The primary source of information on bilateral greenfield FDI is the fDIMarket Database, collected by the FDI Intelligence Unit of the Financial Times ltd. The database contains detailed information on all the greenfield investment projects across more than 150 origin/destination countries for the period 2003-2014. Relying on the Poisson pseudo-maximum likelihood (PPML) estimation technique, our baseline results show a positive and significant effect of asymmetric CP on greenfield FDI. As for the relative importance of each direction of asymmetric CP, our findings suggest that investment projects from an origin to a destination country tend to increase more with the attractiveness of the origin for the destination. More precisely, the elasticities of (the number of) greenfield investment projects amount to 0.30 and 0.07 for (origin to destination) cultural exports and (origin from destination) imports, respectively. This baseline pattern holds across a number of alternative specifications, including the addition of source-destination dyadic fixed effects and instrumentation of cultural trade. Moreover, results are robust to the use of total and average value of greenfield FDI as dependent variables and to different approaches in

the definition of cultural trade.

Our findings shed new light on the mechanisms linking asymmetric CP and greenfield investment. In particular they suggest a stronger role of the 'destination-side' mechanisms. We extend the core analysis of the paper by conducting an empirical test of 'destination consumers demand' and the 'destination political economy' channels and find supportive evidence. We also investigate whether and how the effect of the asymmetric and time-dependent dimension of CP varies at different levels of its symmetric and time-invariant components. We find that time-contingent positive shocks in the asymmetric component of CP increase greenfield FDI only at low levels of the time-invariant, symmetric dimension of CP. This is consistent with a relationship of substitutability between (i) time-contingent, asymmetric and (ii) time-invariant, symmetric dimensions of CP in triggering FDI, with the former operating as a bridgehead between otherwise culturally distant countries.

### 1.1 Related literature

Our paper speaks to the growing literature that considers culture as an important determinant of economic outcomes (see among others Guiso et al., 2006; Fernández, 2008, 2011; Alesina and Giuliano, 2015). We contribute in particular to the debate on whether and how the relationship between cultures affects exchanges and investment patterns across countries (see for instance Head and Mayer, 2014; Giuliano et al., 2014).

To the best of our knowledge this is the first analysis that explores the relationship between CP and FDI fully accounting for the asymmetric nature of CP.[1] This complements the seminal contribution by Guiso et al. (2009) that focus on the impact on international transactions of a related cultural variable: trust. While trust is inherently asymmetric these authors only focus in their FDI gravity regression on one direction of the cultural relationship: i.e. how much individuals in the FDI origin country trust on average individuals in the destination country. While CP and trust are two different cultural variables, their positive correlation (empirically assessed by the these authors in the same paper) and our results suggest that FDI could also positively respond to the trust of citizens in the destination country for those in the country where FDI is coming from.

Our paper is closely related to the two existing studies on the relationship between asymmetric CP and international trade: Disdier et al. (2010) and Felbermayr and Toubal (2010). The former introduces for the first time cultural trade as a proxy for asymmetric and time-dependent CP, the latter uses instead the Eurovision Song Contest voting results. They both find a positive role CP as determinant of trade patterns. Beside the focus on FDI, we contribute to this literature by providing a unifying conceptual framework for CP. In doing that we establish a connection with a related strand in the international business literature, where scholars have started to criticize

[1] There exist empirical studies of bilateral FDI that, while not centering their research question on the link between CP and FDI, include a symmetric (and often time-invariant) regressor to capture CP in an FDI gravity equation. These include Javorcik et al. (2011) and Blonigen and Piger (2014). They all find a positive relationship between CP and FDI. Similar symmetric and often time-invariant measures of CP have been used extensively in gravity equations for trade (see among others Anderson and Van Wincoop, 2003; Head and Mayer, 2014; Feenstra, 2015) as well as migration flows (Bertoli and Moraga, 2013; Beine et al., 2016).

the symmetric and time-invariant concept and measures of CP well before economists. We draw from the seminal work of Shenkar (2001) and propose a definition of CP which accounts for many of the critiques emerging from that literature. From the same strand in international business we acknowledge the recent contribution by Li et al. (2017). These authors focus on role of cultural attractiveness for FDI related outcomes. Differently from our approach, they construct a measure of cultural attractiveness using survey data from the GLOBE project covering 62 societies (House et al., 2004) and do not rely on a structural gravity econometric framework. Moreover, similarly to Guiso et al. (2009), while both directions of cultural attractiveness can potentially affect the same direction of the economic relationship, these authors only focus on the attractiveness of the destination's culture for the origin country, showing a positive role of attractiveness for FDI. Our finding of a strong role of the the origin's culture attractiveness for the destination country extends and complements their investigation.

Our conceptual framework speaks to the theoretical literature that provides micro-foundations to a structural gravity equation for FDI, notably Head and Ries (2008) and de Sousa and Lochard (2011). The 'destination-side' channels that explain the role of the origin's culture attractiveness for the destination country bring novel forces in the existing supply/origin-side gravity models, providing a rationale for the introduction of an additional term in the gravity equation to capture multilateral resistance from the side of the destination country. Our empirical results suggest that these forces are actually at work.

The rest of the paper is organized as follows. Section 2 builds a conceptual framework that explicitly accounts for the asymmetric dimension of CP and presents our proxy based on cultural trade. Section 3 discusses the various elements of the econometric framework proposed to assess the empirical role of CP as a determinant of Greenfield FDI. Baseline estimation results and robustness checks are discussed in Section 4 while Section 5 presents our extensions to the main analysis. Section 6 concludes.

## 2 Asymmetric cultural proximity

Economists and international business scholars have successfully used the concept of culture to identify factors that - in their cross-country variation - (i) explain international economic interactions and (ii) are not captured by relevant parameters such geographic distance or other forms of transaction costs.[2] The definition of culture used in this paper is willingly broad and it accounts for the ideas (values, beliefs, norms) and practices (behavioral patterns) prevailing among respective groups of agents (Leung et al., 2005).

The characterisation of CP between two countries - $i$ and $n$ - as the degree by which the shared ideas and practices of one country tend to be similar to the ones of the other suffers from important limitations which have been highlighted in both the international business and the

[2]While not departing from this approach, we acknowledge that it is not uniformly adopted across social sciences. Indeed, many anthropologists tend to refuse the notion of cultures as bounded, essentialized and internally homogenous entities that can be used to classify, differentiate and compare groups of individuals (see for instance Abu-Lughod, 1996; Appadurai, 1996).

economic literature. Numerous studies including Shenkar (2001), Tung and Verbeke (2010) and Li et al. (2017) demonstrate how cultural relationships which are relevant in the context of international investment are far from being symmetric. For instance Shenkar (2001) relabels the assumption of symmetry in CP as the "illusion of symmetry". One key element is that "symmetry between (1) the distance perceived by country $n$ economic actors vis-à-vis country $i$ and (2) the distance perceived by country $i$ economic actors vis-à-vis country $n$, is often not warranted" (Tung and Verbeke, 2010). Ultimately, the behaviour of economic agents will be affected by their perceptions and therefore needs to be taken as a function of an asymmetric construct of CP. The analysis conducted by these papers provides empirical ground to support this critique. Using data from the GLOBE Project survey Li et al. (2017) find evidence of asymmetry in CP once cultural practices of a target country are mapped with values of an observer country. Practices records represent how a number of cultural elements (such as assertiveness, future orientation, gender egalitarianism) "are" according to the respondents in target while perceptions reflect how the same elements "should be" according to respondents in the observer country. Similar conclusions have been reached by economists. Felbermayr and Toubal (2010) state that "[a] country's citizens can display respect and sympathy for the cultural, societal, and technological achievements of another country without this feeling necessarily being reciprocal". They argue that such asymmetric assessment is relevant in determining bilateral economic interactions among countries and therefore call for a broad notion of CP capable of reflecting asymmetric affinity between two countries. Similar considerations can be found in Guiso et al. (2009) and Disdier et al. (2010) even though, because the empirical exercise in these papers involve only one focal country, the asymmetric aspect of CP is reduced to imply symmetry.

Consistently with these approaches, we assume cultural relationships to be asymmetric and we propose a notion of CP that accounts for that. We explicitly introduce cultural attractiveness as an element of CP. Indeed, individuals in country $i$ can attribute desirable properties to the culture of country $n$ independently on actual similarity between the two cultures.[3] Overall, attractiveness is asymmetric and varies over time. For instance, certain historical events happening in a country could alter the degree by which foreigners find that country's culture attractive. The election of a new president in the United States is likely to change the way countries around the world find American culture attractive as a function of the ideas and practices which are more represented by the elected candidate as well as the specific perceptions of each observer country. This alters the distribution of the US culture's attractiveness across foreign countries, not necessarily having any effect on the way Americans find foreign cultures attractive.

The implication of this discussion is that the asymmetric dimension in the relationship between two cultures can potentially affect economic interactions, and therefore needs to be taken into account when investigating the role of CP for international trade or investment. Formally, we define CP between two countries $i$ and $n$ as

$$CP_{ni,t} = f(\mathrm{S}_{ni}; \mathrm{A}_{ni,t}) \tag{1}$$

[3]Li et al. (2017) derive the construct of cultural attractiveness from the interpersonal attraction framework introduced by the social psychology and sociology literature. The analysis in the present paper does not depart from that conceptualisation.

where $f$ is an increasing function on the unspecified support between minimum and maximum $CP$. $S_{ni}$ denotes the actual similarity between $i$'s culture and $n$'s culture, with $S_{ni} = S_{in}$, while $A_{ni,t}$ is the attractiveness of the $n$'s culture for individuals in $i$. A is asymmetric as the identity $A_{ni,t} = A_{in,t}$ is potentially not verified. Finally, we allow $A_{ni,t}$ to vary over time.[4] In practice $S_{ni}$ can also be subject to time variation. Patterns of migration or geo-political design of national entities are two potential time dependent factors shaping religious, ethnic, linguistic similarity between two countries. We neglect this dimension for three reasons. First, its inclusion does not alter in any way the key results of our study. Second, changes in $S_{ni}$ tend to take place in the long run while variations in the asymmetric component of CP can be relatively quick. This is because attractiveness might respond to a much broader set of events: from the changes of political representation (as in the case of the election example above), to the adoption of new communication technologies capable of better transmitting/accessing cultural contents across countries (for instance the development of machine learning translation algorithms), to the effectiveness of governments to promote the visibility of national cultures abroad, to the international diffusion of pop music from one particular country (e.g. the big success of pop music from South Korea in South America in 2016 and 2017). Third, a symmetric component of CP which is also time invariant represents the exact conceptual counterpart of the standard symmetric and time invariant empirical measures of CP and therefore will allow us for a more direct mapping between the theoretical constructs and the empirical measures (see Section 5.2).[5]

## 2.1 Bilateral cultural trade as a proxy for CP

We argue that bilateral trade flows in cultural goods can be used as meaningful proxies for CP. In particular, the value of $i$'s imports of cultural goods exported by $n$ at time $t$ - $\text{CulIMP}_{ni,t}$ - is an accurate proxy for $CP_{ni,t}$. As discussed by Disdier et al. (2010), $\text{CulIMP}_{ni,t}$ directly and intuitively accounts for $n$'s culture attractiveness for individuals in $i$. Similarly, the value of $i$'s exports of cultural goods imported by $n$ - $\text{CulEXP}_{ni,t}$ - is an accurate proxy for $CP_{in,t}$. As for the capacity of cultural trade to capture the symmetric component of CP, our data shows that there exists a statistically significant empirical relationship between the two, indicating that attractiveness is positively correlated with similarity.[6]

Bilateral cultural trade flows are constructed from the BACI dataset by CEPII[7] and cultural goods identified through the classification of proposed by UNCTAD (UNCTAD, 2010).[8] Table 1

[4]This definition and the subsequent analysis do not rest on the assumption that cultures and perceptions are fixed over time and therefore avoid the "illusion of stability" (Shenkar, 2001).

[5]The definition given in (1) is silent on the potential relationships between $S_{ni}$ and $A_{ni,t}$ or $A_{ni,t}$. The theoretical discussion of these links remain to a large extent outside the scope of the current paper. However, on an empirical ground there exists a positive correlation between $S_{ni}$ and $A_{ni,t}$ (see Appendix B). Moreover, the subsequent empirical exercise allows us to assess the qualitative nature of the relationship between $S_{ni}$ and $A_{ni,t}$ (whether they are complements or substitutes) as determinants of patterns of FDI.

[6]See Appendix B.

[7]See `http://www.cepii.fr/cepii/en/bdd_modele/presentation.asp?id=1` and Appendix A for a detailed discussion of the data.

[8]The choice of the UNCTAD classification to define the relevant set of cultural goods serves the purpose of maximizing the country coverage of the resulting estimation sample. We depart from Disdier et al. (2010) that define cultural goods following a different scheme. The implications due to the adoption of a different classification scheme are discussed in Appendix A.

reports the products which are classified as cultural goods. The UNCTAD classification divides them into two categories, 'core' and 'optional' cultural goods, listed in the first and second column of Table 1 respectively. Each category has two headings, arts and media within the 'core' category and heritage and functional creation within the optional one. Core cultural goods generally embed a higher cultural content and they are listed across other available classification schemes such as the one developed by UNESCO.

Table 1: Categories of Goods with Cultural Content (UNCTAD, 2010)

| **Core Cultural Goods** | **Optional Cultural Goods** |
|---|---|
| *Arts (Performing and Visual)* | *Heritage (Arts Crafts)* |
| Music (CD, Tapes), Printed Music, Painting, Photography, Sculpture and Antiques | Carpets, Celebration, Paperware, Wickerware, Yarn and Other |
| *Media (Publishing and Audio-Visual)* | *Functional Creations (Design and New-Media)* |
| Books, Newspaper, Other Printed Matter, Film | Architecture, Fashion, Interior, Glassware, Jewellery, Toys, Recorded Media and Video Games |

Notes: Further information on the classification can be found in UNCTAD (2010). This table replicates Table 4.2, p. 112 of UNCTAD (2010).

Before the merging with FDI and other data the cultural trade database has a coverage of 176 countries on the period 2003-2014. On average across countries and over time trade in cultural goods accounts for 2.7% of total trade in this sample. As noted in Disdier et al. (2010), cultural trade is highly concentrated. Summing cultural trade flows across importers and over time, the top five exporters - China, Germany, USA, Italy and France - account for 55% of total cultural trade. When looking at all trade instead, the top 5 exporters - China, Germany, USA, Japan and France - account for 37% of the total.

## 2.2 A detour on asymmetry

Before turning to the main research question in the paper, we provide some descriptive evidence of the asymmetry embedded in the bilateral flows of cultural goods.

We start by constructing an empirical measure of asymmetry in CP. This is done in two steps. First, we estimate a simple linear model where cultural trade $\text{CulIMP}_{ni,t}$ is regressed on importer-time fixed effects $\delta_{i,t}$; country pair fixed effects $\gamma_{ni}$; and an error term $\varepsilon_{ni,t}$. The empirical estimate $\hat{\gamma}_{ni}$ has a useful economic interpretation: it captures, on average over time, how much individuals in (importing) country $i$ consider the culture of (exporting) country $n$ attractive above or below the attractiveness of the average country.[9] Second, for each (undirected) pair of different countries we compute the absolute value of the difference between $\hat{\gamma}_{ni}$ and $\hat{\gamma}_{in}$. We interpret the result as a proxy for the degree of asymmetry in the CP between two countries.

While the data - covering bilateral cultural trade for 176 countries - would in principle allow to estimate this measure for 15400 country pairs, due to the high number of zeros we are able to derive both $\hat{\gamma}_{ni}$ and $\hat{\gamma}_{in}$ only for 4137 pairs. While they account for just less than one third

[9]This regression adapts to the context of our data an empirical exercise proposed in Felbermayr and Toubal (2010).

of all potential combinations, these 4137 pairs account for 49.1% and 55.8% of total trade and total trade in cultural goods respectively. To illustrate the scope of the asymmetry embedded in cultural trade, Table 2 reports the country pairs with the highest and the lowest value of the asymmetry measure. For these two pairs we report the directed attractiveness premia and the resulting value of asymmetry implied by cultural trade.

Table 2: Max and Min Asymmetry

| Country $n$ | Country $i$ | Attractiveness premium of $i$ for $n$ ($\hat{\gamma}_{ni}$) | Attractiveness premium of $n$ for $i$ ($\hat{\gamma}_{in}$) | Asymmetry ($\lvert\hat{\gamma}_{ni} - \hat{\gamma}_{in}\rvert$) |
|---|---|---|---|---|
| China | Paraguay | 7.211 | -3.686 | 10.897 |
| Morocco | Singapore | 0.047 | 0.046 | 0.001 |

Notes: The table lists the two pairs showing respectively the higher (lower) asymmetry in attractiveness premia awarded to each other, according to the full sample of countries for which the estimated measure of asymmetry is available.

Table 2 shows the maximum and minimum values taken by the measure of asymmetry described above. The highest asymmetry estimated from our sample is between Paraguay ($i$) and China ($n$). In particular, China appears much more attractive for Paraguay relative to the average country ($\hat{\gamma}_{ni}$ = 7.211). On the contrary the attractiveness of Paraguay's culture for China is lower than the average country's attractiveness ($\hat{\gamma}_{in}$ = −3.686). In other words, individuals in Paraguay tend to put a positive attractiveness premium on Chinese culture while Chinese individuals tend to find Paraguay's culture less attractive than others. In order to get a more concrete understanding of this maximum asymmetry one can look at the actual value of the relevant cultural trade flows in the whole sample of bilateral cultural trade. In particular, the average value - across years and exporters - of Paraguay's imports of cultural goods is USD 2,087,000 while on average across years Paraguay imports from China USD 273,137,000 (almost 131 times the cross country average). On the other hand, the average Chinese imports of cultural good (across years and exporting countries) is USD 29,563,000 while its average yearly imports from Paraguay is just USD 23,000 (0.08% of the average value across exporters).[10] Minimum asymmetry is found between Morocco and Singapore. In this case there exists a very balanced neutrality, with each country awarding the other with a very low attractiveness premium.

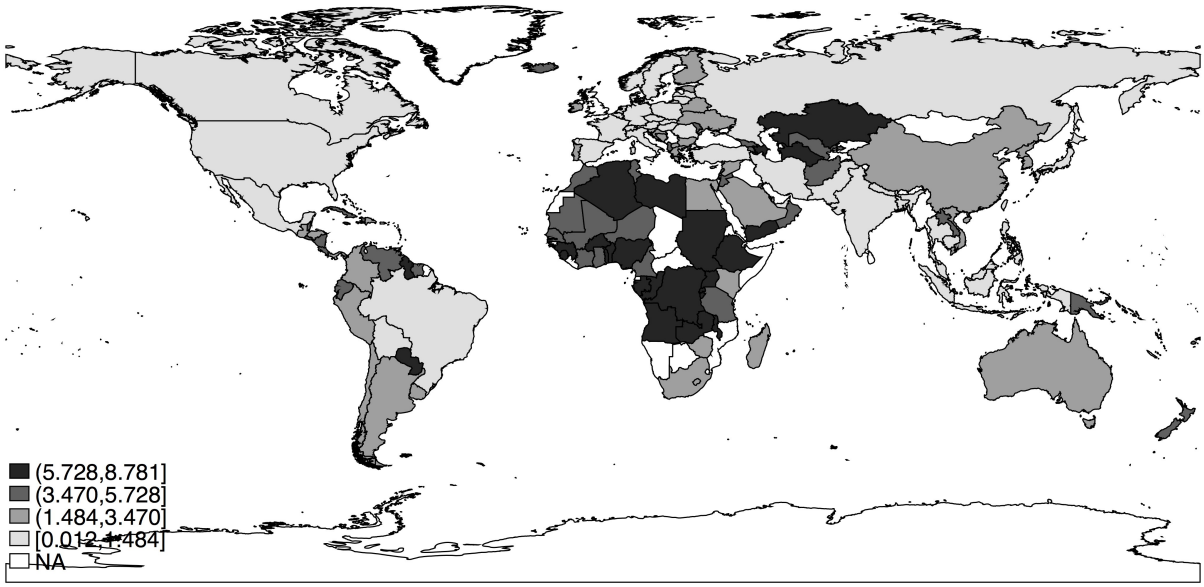
We complement the discussion of the extreme values of asymmetry by exploring the case of the UK and its bilateral cultural relationships with the other countries. The UK is the sixth biggest exporter and the second importer of cultural goods.[11] Because of the British Empire the legal, linguistic and cultural connections of the UK are many and relatively well known. For these reasons the UK represents a useful reference point for this exercise. Figure 1 provides a graphical representation of the distribution of asymmetry in the 156 available country pairs involving the UK. The colors denote the four quartiles of the distribution over these 156 observations: darker tones indicate higher asymmetry.

A low degree of asymmetry in the cultural relationship reflected in cultural trade involving the

[10] This case seems to be suggestive of a potential correlation between asymmetry in export capacity and high asymmetry in cultural relationships: indeed, even if the table only shows the upper bound, this pattern finds support in the data. See Appendix C for a simple assessment of this correlation. A comprehensive investigation of the determinants of asymmetry in CP goes beyond the scope of the preset paper.

[11] This ranking is based on total trade flows for the period 2003-2014 across 176 countries.

Figure 1: Asymmetry in CP Between the UK and the Rest of the World




UK is apparent for many European countries (with the notable exception of Ireland); for many economies in the South-East Asia region; for Russia; for the North American countries; and for some Latin American ones. High asymmetry emerges between the UK and countries in the African continent (with few exceptions below the median level of asymmetry including Madagascar and South Africa); countries in the Central Asia region; and few countries in Latin America. Relatively low asymmetry in the cultural relationships with European countries highlights the capacity of our empirical framework and of its wide country coverage to complement previous studies on the role of asymmetric cultural variables for economic transactions with a focus on European countries. Indeed, both Guiso et al. (2009) and Felbermayr and Toubal (2010) document the existence of a significant degree of asymmetry in patterns of trust and of affinity by using data on a relatively narrow and homogeneous set of countries. The case of the UK presented in Figure 1 suggests that intra Europe bilateral cultural relationships appear relatively more symmetric when extrapolated from a global empirical framework. The issue of overestimated asymmetry in settings with narrow and homogeneous country coverage is further explored in Appendix C.

Finally, while the exercise in Figure 1 provides suggestive evidence for the distribution of the asymmetric component in cultural relationships, it remains largely uninformative regarding the type of asymmetry in each country pair. For instance, the relatively high asymmetry between the UK and Ireland (2.700) originates from a very high affinity premium placed by Ireland on the UK ($\hat{\gamma}_{GBR,IRL} = 8.677$) and only partly reciprocated by the still high affinity premium of the UK for Ireland ($\hat{\gamma}_{IRL,GBR} = 5.977$). On the contrary, the almost identical asymmetry score between the UK and Honduras reflects a low affinity premium of Honduras for the UK ($\hat{\gamma}_{GBR,HND} = 0.175$) to which the UK corresponds a negative one ($\hat{\gamma}_{HND,GBR} = -2.525$).

The descriptive detour proposed in this section served the purpose of illustrating the existence and scope of asymmetry in CP as an empirical phenomenon captured by bilateral cultural trade.

A focus on such asymmetry is central to our main research question, which we now turn to address.

# 3 Econometric framework

The econometric framework used to assess the empirical relationship between CP proxied by cultural trade and Greenfield FDI is constructed in several steps. First, we introduce a gravity model of bilateral FDI building on Head and Ries (2008) and de Sousa and Lochard (2011). Then, equipped with the definition of CP given in Section 2, we discuss theoretical mechanisms linking CP and greenfield FDI. Finally, the estimation strategy and data are presented.

## 3.1 Asymmetric CP and FDI gravity models

To assess how bilateral, asymmetric and time-varying CP affects bilateral patterns of greenfield FDI, we follow the theoretical model of greenfield FDI proposed by de Sousa and Lochard (2011) which is rooted in the seminal theory by Head and Ries (2008). Both models are characterized by a partial equilibrium, supply side perspective. Moreover, their gravity nature accounts for multilateral frictions, i.e. decisions made by MNEs to invest in a particular destination are not independent on their investment decisions into other countries.[12]

The theory is simple. Greenfield FDI projects are modelled as inspection games between the manager of a MNE (MM) and that of its foreign subsidiary (Sub). The payoff of the MM denoted by $\nu$ is a negative function of an inspection cost $c$ and a transaction cost $\tau$. The former reflects the standard costs of monitoring which can be implemented by the MM in order to detect a shirking behavior of Sub. The latter materializes whenever Sub exerts effort and adds value to the investment project. $\tau$ encompasses all types of costs associated with greenfield FDI beyond inspection costs. Examples includes the costs of dealing with "currency risks, exchange-rate transaction costs, trading- and liquidity-related costs as well as differentials of taxation, accounting, and legal standards in a broader interpretation" (de Sousa and Lochard, 2011, p. 554). Both $c$ and $\tau$ are functions of a vector of formal investment policies, geographic and cultural proximity.

In a multi country framework with stochastic MNE's payoff functions, MM chooses to invest in a country where the highest value of a project is higher than the highest value of projects in all other countries. The model allows to represent the number (or value) of greenfield FDI projects from origin country $i$ into destination country $n$ with a structural gravity equation of the kind

$$FDI_{ni} = K_i A_i^{-1} M_n T_{ni} \tag{2}$$

The term $K_i$ is a function of the origin/parent country specific parameters, such as the total number of investment projects that can be financed (the total capital stock). $A_i^{-1}$ is a multilat-

[12]This approach differentiates these models from the knowledge-capital model of MNEs.

eral resistance component, capturing the attractiveness of alternative locations for investors in country $i$. $M_n$ is a function of the destination/host country specific parameters, which include the total number of potential investment projects and the average contribution of Sup across projects. Finally, $T_{ni}$ is the bilateral component, a function of both monitoring and transaction costs, but also of the vector of formal investment policies, geographic proximity and CP. Intuitively, the model specifies $T_{ni}$ as a decreasing function of $c$ and $\tau$. The qualitative relationship between these costs and formal investment policies as well as geographical distance parameters is taken from Head and Ries (2008) and de Sousa and Lochard (2011). The existence of FTAs (Free Trade Agreements) or BITs (Bilateral Investment Treaties) between $i$ and $n$ can potentially reduce both monitoring and transaction costs, which are also assumed to decrease with geographical proximity.

The way $c$ and $\tau$ depend upon the symmetric component of CP is not new to the FDI gravity literature in economics: higher similarity between the two cultures implies lower monitoring as well as lower transaction costs. What has not been discussed is how monitoring and transaction costs react to the asymmetric component of CP. In what follows we address this in a broader discussion on how greenfield FDI from origin $i$ to destination $n$ depends upon both $CP_{ni,t}$ and $CP_{in,t}$.

Higher $CP_{ni,t}$ reduces the costs that the parent MNE has to pay to monitor the activities of its foreign subsidiary. This is intuitive if higher $CP_{ni,t}$ reflects higher $\mathrm{S}_{ni}$. Indeed, for many symmetric dimensions of CP (common language, similar legal practices and contracting behaviour) clearly facilitate monitoring activities. However, $\mathrm{A}_{ni,t}$, the degree of attractiveness for individuals in the origin country $i$ of the ideas and practices which are prevalent among individuals in destination $n$, is also a determinant of lower monitoring costs. It minimizes assessment errors and facilitate the assessment processes themselves by making easier for $i$ individuals (that have to evaluate the effort exerted by the subsidiary located in $i$) to establish an effective interaction with $n$ agents, beyond a common language framework. By effective interaction we mean an interaction that favours a quicker and more precise understanding of what the other is saying as well as of what she is hiding. As for transactions costs, both $\mathrm{S}_{ni}$ and $\mathrm{A}_{ni,t}$ minimize the costs to cope with different accounting/legal standards and in general with all corporate standards that might differ across the parent and the host country. Finally, from the point of view of country $i$ parent personnel, if an inspection activity or the work needed to harmonize different corporate-related standards involves interaction with $n$'s individuals and/or business trips to country $n$, higher appreciation by country $i$ individuals of the culture of country $n$ reduces the costs associated with these activities.[13] These mechanisms altogether unambiguously predict a positive effect of $\mathrm{CulIMP}_{ni,t}$ on greenfield investment from $i$ to $n$.

Let us now consider the role of $CP_{in,t}$ in explaining greenfield FDI from origin country $i$ to destination $n$. Notice that our arguments on the role of $\mathrm{S}_{ni}$ apply to $\mathrm{S}_{in}$ as well due to the symmetric nature of S. Discussing the role of $CP_{in,t}$ therefore amounts to consider the role of $\mathrm{A}_{in,t}$, i.e. of the attractiveness of the $i$'s culture for individuals in $n$. From the point of

[13] For a detailed review of the mechanisms that make destination's cultural attractiveness for the origin country a relevant driver of origin's MNEs' FDI decisions see Li et al. (2017).

view of the subsidiary personnel in the destination country $n$, the attractiveness of $i$'s culture for them results in a good attitude toward interactions with the parent's personnel. Smoother interactions reduce inspection as well as transaction costs for the MNE. But $A_{in,t}$ can be relevant for $i$'s investment in $n$ beyond its effect on $i$'s MNE monitoring and transaction costs. First, in so far as the $n$ subsidiary is intended to serve the $n$ market, the value that consumers in $n$ put on the output of $i$'s MNE increases the average payoff from a greenfield investment in country $n$. This preference value is likely to be a positive function of how much individuals (consumers) in $n$ are attracted by $i$'s culture ($A_{in,t}$), also relatively to the cultures of other potential investors. This 'destination consumers demand' channel is likely to be particularly relevant (i) when the outcome of the FDI project is a final consumption good and (ii) in sectors where FDI is the prevailing mode of international provision, as it is still the case for many services sectors. Second, the realization of an FDI project by $i$ can be facilitated or opposed by political pressures in the host country $n$. A plausible assumption is that political pressures to facilitate inward foreign investment will be allocated to $i$'s projects, also according to the degree by which individuals (voters) in $n$ appreciate $i$'s culture with respect to those of other potential investors. We expect this 'destination political economy' channel to be more pronounced for destination countries with higher political accountability, i.e. where politicians tend to be less independent from voters preferences in their political and economic decisions.

These 'destination-side' mechanisms are not accounted for in the classical theoretical framework of de Sousa and Lochard (2011) and they call for an additional term in the gravity equation to capture multilateral resistance from the side of the destination country $n$. We rewrite (2) as

$$FDI_{ni} = K_i A_i^{-1} M_n B_n^{-1} T_{ni} \tag{3}$$

where $B_n^{-1}$ is a function of the attractiveness of alternative investors for $n$'s consumers and/or voters.

The micro-foundation of the destination-side mechanisms by extending the theory of de Sousa and Lochard (2011) is a task that goes beyond the scope of the current paper: in fact they do not suggest any theoretical ambiguity about the sign of the relationship between $CP_{in,t}$ and $i$'s investment into $n$. All in all, the discussed mechanisms unambiguously imply a positive effect of $CP_{in,t}$ on greenfield investment from $i$ to $n$.

### 3.2 Baseline estimation, identification strategy and data

The structural gravity model (3) augmented with the time dimension can be brought to the data. Following Santos Silva and Tenreyro (2006) we rely on the PPML estimation method.

The dependent variable used in the baseline estimation exercise is $C_{ni,t}$, the number of Greenfield FDI project from an origin country $i$ to a destination country $n$ at time $t$. The origin and destination specific components $K_{i,t}$ and $M_{n,t}$, as well as the multilateral resistances $A_{i,t}^{-1}$ and $B_{n,t}^{-1}$ are accounted for through origin-time and destination-time fixed effects. The elements of the bilateral component $T_{ni,t}$ are captured through (i) the log of the distance between origin

and destination ($\ln \text{dist}_{ni}$); (ii) a dummy for geographical contiguity ($\text{contig}_{ni}$) as proxies for transportation costs; (iii) the number of FTAs and BITs involving $i$ and $n$ which are in force at time $t$ ($\text{FTA}_{ni,t}$ and $\text{BIT}_{ni,t}$) as measures of formal investment policy. Finally, the elements of $T_{ni,t}$ which pertain to CP are proxied with both directions of cultural trade between $i$ and $n$, ($\text{CulIMP}_{ni,t}$ and $\text{CulEXP}_{ni,t}$). In order to identify the specific role of the asymmetric component of CP ($\text{A}_{ni,t}$ and $\text{A}_{in,t}$) we control for its symmetric component ($\text{S}_{ni} = \text{S}_{in}$) by adding to our specification the standard symmetric and time-invariant measures of CP (a former colony dummy $\text{colony}_{ni}$, linguistic $\text{lang}_{ni}$, religious $\text{comrelig}_{ni}$, and institutional proximity $\text{comleg}_{ni}$). We acknowledge from the outset that our identification can be potentially undermined by endogeneity arising from omitted variable or reverse causality issues. We address this concern in Section 4.2.1.

The fDiMarket Database we use, collects information on greenfield FDI from January 2003 onward, and it is constantly updated. To the best of our knowledge, it constitutes the most reliable and complete existing source of greenfield investment data.[14]

In addition to Greenfield FDI information for the dependent variables and the data on cultural trade flows which constitute the main regressors of interest (see Section 2.1 above), we include in the gravity specification measures of linguistic proximity from Melitz and Toubal (2014) and Adsera and Pytlikova (2015). These indices integrate the standard bilateral linguistic measures adopted in the majority of gravity models that do not focus on CP. Data on bilateral investment treaties come from the UNCTAD Investment Policy Hub. All remaining gravity and distance related variables used throughout the empirical analysis come from the CEPII's *geodist* and *gravdata* datasets. See Appendix A for a more thorough description of data sources and how the dataset is created.

The dataset used for the baseline estimation consists of an unbalanced panel of 87,448 observations. It features 144 origin and 178 destination countries over the 12 years period from 2003 to 2014. Summary statistics for the variables used in the baseline estimation are given in Table 3.

[14]Completeness does not exclude misreporting or missing data, but such missing data are likely to be very limited and continuously revised by the dataset provider (`http://www.fdiintelligence.com/fDi-Tools/fDi-Markets`).

Table 3: Summary Statistics from Baseline Estimation Sample

| Variable | Mean | Median | sd | Min | Max |
|---|---|---|---|---|---|
| $C_{ni,t}$ | 1.551 | 0 | 8.897 | 0 | 400 |
| $\ln \text{dist}_{ni}$ | 8.482 | 8.747 | 0.910 | 4.107 | 9.892 |
| $\text{colony}_{ni}$ | 0.032 | 0 | 0.177 | 0 | 1 |
| $\text{lang}_{ni}$ | 0.157 | 0 | 0.364 | 0 | 1 |
| $\text{comrelig}_{ni}$ | 0.173 | 0.033 | 0.266 | 0 | 0.989 |
| $\text{contig}_{ni}$ | 0.038 | 0 | 0.190 | 0 | 1 |
| $\text{comleg}_{ni}$ | 0.293 | 0 | 0.455 | 0 | 1 |
| $\text{FTA}_{ni,t}$ | 0.269 | 0 | 0.444 | 0 | 1 |
| $\text{BIT}_{ni,t}$ | 0.393 | 0 | 0.488 | 0 | 1 |
| $\ln \text{CultIMP}_{ni,t}$ | -0.454 | -0.429 | 3.273 | -6.908 | 10.644 |
| $\ln \text{CultEXP}_{ni,t}$ | -0.145 | -0.086 | 3.114 | -6.908 | 10.644 |

Notes: This table reports summary statistics for the variables used in the baseline estimation exercise (see Table 4). The related estimation sample consists of 87,448 observations.

# 4 Results

In this section we present the results of the empirical analysis. We discuss the baseline estimation results in Section 4.1 and then the main robustness tests in Section 4.2. Further extensions to the core analysis of the paper are discussed separately in Section 5.

## 4.1 Baseline results

Table 4 below presents the main results of our empirical exercise. The positive and statistically significant coefficient of $\ln \text{CultIMP}_{ni,t}$ in column (1) shows that the attractiveness of the $n$'s culture for individuals in country $i$ ($\text{A}_{ni,t}$) is a determinant of the number of greenfield FDI projects from $i$ to $n$. In particular, the number of investments from an origin country to a destination economy increases with $\text{A}_{ni,t}$ as captured by the value of $i$'s cultural imports from $n$. Analogously, the estimated coefficient of $\ln \text{CultEXP}_{ni,t}$ in column (2) is positive and statistically significant, showing that the number of greenfield FDI projects from origin $i$ to destination $n$ is higher for stronger attractiveness of the $i$'s culture for individuals in the in $n$ ($\text{A}_{in,t}$). Finally, both bilateral flows of cultural goods between the origin $i$ and the destination $n$ are included in the specification reported in column (3) of Table 4. Their estimated coefficients remain positive and highly significant but the magnitude of the point estimate for $\ln \text{CultIMP}_{ni,t}$ is more than halved. The impact of trade in cultural goods on the number of greenfield FDI projects is identified beyond the role of the other gravity variables and of the standard proxies for CP. This shows that the asymmetric component of CP plays a role above and beyond its symmetric elements.

These results suggest that investment projects from $i$ to $n$ tend to increase more with the attractiveness of the origin's culture for individuals in the destination - $\text{A}_{in,t}$ - rather than with

Table 4: Impact of CP on Greenfield FDI (Number of Projects)

| Dep. Var. | Count $C_{ni,t}$ | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| $\ln \text{CultIMP}_{ni,t}$ | 0.165*** | | 0.0690*** |
| | (11.87) | | (5.90) |
| $\ln \text{CultEXP}_{ni,t}$ | | 0.330*** | 0.305*** |
| | | (23.71) | (21.91) |
| $\ln \text{dist}_{ni}$ | -0.407*** | -0.214*** | -0.179*** |
| | (-11.60) | (-6.19) | (-5.13) |
| $\text{colony}_{ni}$ | 0.478*** | 0.387*** | 0.366*** |
| | (7.89) | (6.95) | (6.85) |
| $\text{lang}_{ni}$ | 0.254*** | 0.189*** | 0.181** |
| | (4.20) | (3.73) | (3.53) |
| $\text{comrelig}_{ni}$ | 1.002*** | 0.893*** | 0.883*** |
| | (9.47) | (9.51) | (9.21) |
| $\text{contig}_{ni}$ | -0.114 | 0.0752 | -0.0977 |
| | (-1.71) | (-1.21) | (-1.61) |
| $\text{comleg}_{ni}$ | 0.253*** | 0.170*** | 0.153*** |
| | (6.01) | (4.59) | (4.06) |
| $\text{FTA}_{ni,t}$ | 0.172** | 0.135* | 0.118* |
| | (3.02) | (2.49) | (2.19) |
| $\text{BIT}_{ni,t}$ | 0.0398 | 0.0119 | 0.0115 |
| | (0.93) | (0.29) | (0.29) |
| Imp×Year FE | √ | √ | √ |
| Exp×Year FE | √ | √ | √ |
| Obs | 87448 | 87448 | 87448 |
| % Zeros | 0.749 | 0.749 | 0.749 |
| $R^2$ | 0.9056 | 0.9216 | 0.9221 |
| Estimator | PPML | PPML | PPML |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z-statistics in parentheses. Standard errors are clustered by trading-pair. The dependent variable "Count" $C_{ni,t}$ is the bilateral number of Greenfield FDI projects from country $i$ to country $n$. It includes the zero flows. The estimates are obtained with PPML using the *PPML panel sg* command written by Thomas Zylkin which simultaneously allows to absorb pair-wise as well as origin-by-time and destination-by-time FEs (see Larch et al., 2017). The model includes origin×time and destination×time FEs. The sample size in this table is invariant to the number of covariates included and refers to the regression which features both imports and exports of cultural goods. The information which belong to groups with all zeros or missing values are automatically dropped by the estimator as FEs cannot be computed.

$A_{ni,t}$. Relying on the point estimates in column (3) of Table 4, the elasticities of cultural trade on the number of greenfield investment projects amount to 0.30 and 0.07 for (source to destination) exports and (source from destination) imports respectively. This finding sheds some light on the relative importance of the theoretical mechanisms linking asymmetric CP and greenfield investment. In particular it points to a relatively stronger role of those mechanisms discussed in Section 3.1 that explain greenfield FDI of $i$ into $n$ with the attractiveness of the culture of the origin country $i$ for individuals in the destination country $n$. Our results confirm that it certainly matters how much the manager of the $i$ MNE appreciates the culture in the country where the company invests, as this would imply expectations of lower monitoring and transaction costs. However, it matters more how much individuals in the destination economy appreciate the culture represented by the affiliate of the MNE in their country. Our conceptual framework

(see Section 3.1) suggests that this too can be due to the MNE manager's expectations of lower monitoring and transaction costs (because of smoother interaction with agents that appreciate the culture represented by the MNE) but also to destination-specific channels. These are a higher propensity of the individuals in the destination country to buy the output of the MNE affiliate in their country ('destination consumers demand' channel) as well as to approve political (and economic) support toward the FDI project by their government ('destination political economy' channel). Both channels increase the profitability of the FDI project and therefore stimulate greenfield investment.[15]

## 4.2 Robustness checks

In this section we test the robustness of our results. The main econometric concern in our benchmark estimates is the potential endogeneity of our proxy for CP - i.e. trade in cultural goods - which may derive from multiple sources: for instance because of the omission of dyadic specific unobserved factors that might be correlated both with the error term (hence with FDI) and with CP. In particular, as noted by Felbermayr and Toubal (2010) and Disdier et al. (2010) these unobserved elements are often related to initial conditions, since the mutual learning due to strong pre-existing ties may favor convergence of cultural characteristics which in turn can trigger even more intense FDI flows. Furthermore, the link between CP and FDI may be subject to reverse causality as there might be determinants of FDI that drive both economic outcomes as well as cultural attractiveness, making it difficult to establish a clear direction of causation (see Felbermayr and Toubal, 2010; Guiso et al., 2009). Indeed, positive FDI shocks may increase the interactions with foreign partners which in turn could lead to mutual learning and further cultural convergence and appreciation. Finally, measurement error can bias the estimated impact of our parameters of interest. In particular, the data on Greenfield FDI from the FT dataset include estimates for capital investment (derived from algorithms) when a company does not release the information (see Desbordes and Wei, 2017; Lee and Ries, 2016). As for CP, the cultural content embodied in different categories of cultural goods may reflect different degrees of bilateral CP. We deal with the first two sources of endogeneity - namely omitted variable bias and reverse causality - through the inclusion of dyadic fixed effects and by adopting an instrumental variable (IV) approach, respectively.[16] We address the measurement error concerns by first testing our benchmark specification on different measures of the bilateral volume of FDI and then replicating the same specification separately on different sub-categories of the baseline set of cultural goods.

### 4.2.1 Controlling for time-invariant unobserved factors and reverse causality

We start discussing the inclusion of dyadic fixed effects. Table 5 compares our benchmark results with the fully specified model. The inclusion of dyadic fixed effects absorbs all the cross section

[15] In Section 5 we present a more detailed test of the 'destination consumers demand' and the 'destination political economy' channels.

[16] In Appendix D we further test the consistency of our benchmark results by augmenting the specification with the inclusion of observable variables of dimension *nit* that might capture (part) of the unobserved time-varying dyadic factors.

variability in our sample, so that the impact of CP depends solely upon time contingent cultural factors. To allow for comparison of the results, the sample size is identical in all columns as we maintain the same sample for the fully specified model across all specifications. The models with country×year fixed effects (columns 1-3) deliver roughly the same results as Table4, so the reduction of the sample size does not significantly alter our benchmark estimates. On the other hand, similarly to Felbermayr and Toubal (2010) and Disdier et al. (2010), the inclusion of dyadic fixed effects in column (4) substantially affects our parameters of interest. Trade in cultural goods retains a positive impact on FDI, but the magnitude of both the elasticities of cultural imports and exports is much lower with respect to the benchmark equation, indicating that CP is largely captured by an unobservable time invariant component. In addition, only the impact of exports remain statistically significant: this finding suggests that only the time variation of attractiveness of the origin's culture for the individuals in the destination economy plays a role in the MNE decision to invest.

We now move to the issue of reverse causality. In the literature the simultaneity problem has been commonly addressed with an IV strategy where current levels of CP are instrumented with their past values (see for instance Felbermayr and Toubal (2010)). This strategy hinges on the assumptions that (i) lagged bilateral values CP predict their current levels sufficiently well and that (ii) current shocks in the gravity equation are uncorrelated to past cultural relationships. While we find the first validity condition plausible, the latter which refers to the exogeneity of the instrument is neither obvious, nor easy to demonstrate. For instance, it could be argued that part of the current variation of FDI is associated to the evolution of cross-country cultural relationships and therefore depends on past shocks of CP. Indeed, FDI normally requires a long-term focus and the MNEs decision to invest is likely to depend even more on past than current levels of CP. In our conceptual framework an alternative way to address the issue of reverse causality is to adopt a completely different approach by replacing current levels of cultural trade with their lagged values as the main variable of interest. The advantage of this strategy is that trade flows are predetermined with respect to FDI which is likely to attenuate the issue of reverse causality, without being a binding/necessary condition for the consistency of the estimator. Although they are based on somewhat contrasting assumptions, in our robustness analysis we propose both strategies - the IV and the lagged approach - to address the simultaneity problem. In the first two columns of Table 6 we estimate our baseline specification with the predetermined values of cultural trade at t-2 and t-5 in columns 1 and 2, respectively. The point estimates of our parameters of interest in both regressions are very close to the baseline results, which we find as reassuring. In addition, the very limited variation over time of the impact of trade in cultural goods suggests a persistence in bilateral cultural tastes or, alternatively, a very similar variation in CP over time for all country pairs. This finding is in line with the relatively low impact of the time variation of CP on FDI obtained by introducing country pair fixed effects in Table 5.

The IV strategy reported in the remaining two columns of Table 6 builds on Combes et al. (2005), Briant et al. (2014) and Felbermayr and Toubal (2010) and exploits the longitudinal nature of the BACI dataset by instrumenting current levels of cultural trade flows with lagged values of the same variables $(t-12)$.[17] Columns 3 and 4 compare the PPML estimates with

[17]The earliest year available from BACI dataset is 1995: this forces us to reduce the time span (2007-2014)

Table 5: Impact of Cultural Proximity on Greenfield FDI: Adding Country Pair FE

| Dep. Var. | Count $C_{ni,t}$ | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $\ln \text{CultIMP}_{ni,t}$ | 0.145*** | | 0.0522*** | 0.00677 |
| | (10.35) | | (4.43) | (0.78) |
| $\ln \text{CultEXP}_{ni,t}$ | | 0.314*** | 0.295*** | 0.0499*** |
| | | (22.57) | (21.04) | (3.72) |
| $\ln \text{dist}_{ni}$ | -0.404*** | -0.208*** | -0.181*** | |
| | (-11.94) | (-6.27) | (-5.42) | |
| $\text{colony}_{ni}$ | 0.481*** | 0.388*** | 0.372*** | |
| | (8.04) | (7.14) | (7.08) | |
| $\text{lang}_{ni}$ | 0.244*** | 0.180*** | 0.173*** | |
| | (4.06) | (3.58) | (3.43) | |
| $\text{comrelig}_{ni}$ | 0.957*** | 0.855*** | 0.847*** | |
| | (9.04) | (9.06) | (8.84) | |
| $\text{contig}_{ni}$ | -0.0905 | -0.0578 | -0.0754 | |
| | (-1.40) | (-0.96) | (-1.28) | |
| $\text{comleg}_{ni}$ | 0.246*** | 0.164*** | 0.151*** | |
| | (5.90) | (4.43) | (4.03) | |
| $\text{FTA}_{ni,t}$ | 0.147** | 0.109* | 0.0976 | 0.0499 |
| | (2.62) | (2.09) | (1.87) | (1.12) |
| $\text{BIT}_{ni,t}$ | -0.0145 | -0.0368 | -0.0358 | 0.117 |
| | (-0.34) | (-0.93) | (-0.92) | (1.41) |
| Imp×Year FE | √ | √ | √ | √ |
| Exp×Year FE | √ | √ | √ | √ |
| Country Pair FE | | | | √ |
| Obs | 49702 | 49702 | 49702 | 49027 |
| % Zeros | 55.99 | 55.99 | 55.99 | 55.99 |
| $R^2$ | 0.9053 | 0.9222 | 0.9224 | 0.9686 |
| Estimator | PPML | PPML | PPML | PPML |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z-statistics in parentheses. Standard errors are clustered by trading-pair. The dependent variable "Count" $C_{ni,t}$ is the value of the aggregated bilateral flow of greenfield investments from country $i$ to country $n$, including zero flows. The estimates are obtained with PPML using the *PPML panel sg* command written by Thomas Zylkin which simultaneously allows to absorb pair-wise as well as origin-by-time and destination-by-time FEs. The columns (1) to (3) replicate table 4 results, and include origin×time and destination×time FEs only. Column (4) includes Country Pair FE, to address multilateral resistance, Baldwin and Taglioni (2006), Baier and Bergstrand (2007), Head and Mayer (2014) and Piermartini and Yotov (2016) among the others, suggest to include country×time dummy and trading pair dummies.

the correspondent coefficients obtained with IVPPML using the reduced sample of Felbermayr and Toubal (2010). Concerning our parameters of interest, controlling for endogeneity leads to results that are in line with the literature and consistent with the estimates of the fully specified model. The elasticity of imports of cultural goods roughly maintain the same magnitude as in the PPML model, but becomes statistically not significant. As for exports, when instrumented their coefficient remains statistically significant at the 1% confidence level, and substantially in-

in our IV analysis. The time varying lagged instrument is relevant as it is strongly correlated to the endogenous variable as showed in Appendix E. The IV strategy is performed with the Stata command IVPOISSON which doesn't allow for the inclusion of high dimensional fixed effects. In order to include a comprehensive set of fixed effects which account for time varying importer and exporter heterogeneity, our strategy is to reduce the sample size to ensure convergence in the estimation.

creases in magnitude. Hence, once we control for reverse causality, we find that only the cultural attractiveness of the origin country for potential destinations have an impact on greenfield investment. Furthermore, the instrumented exports' elasticity is more than twice as large, suggesting a downward bias in the impact of exports of cultural goods. However, the resulting downward bias is substantially smaller compared to the estimates emerging from previous studies on the impact of CP on economic exchanges, suggesting that our gravity specification suffers relatively less from endogeneity compared to other proxies used so far in the literature.[18]

### 4.2.2 FDI count versus value

Table 7 replicates the same specification of Table 4 using the total value of bilateral investments ($V_{ni,t}$) rather than their number.

The focus on the number of projects (count) as opposed to their total or average value has the advantage of minimizing the potential distortions induced by the imputation techniques used in the construction of the value-related variables,[19] but has its own limitations: for instance it is equivalent to imposing to all projects the same weight in terms of economic relevance, without discriminating them for their actual size. For instance, an investment in a legal consultant office (the business sector with the lowest average capital investment in our sample) is implicitly evaluated as an investment in a plant for oil refinery, which is roughly 257 times larger (5.344 millions US$ against more than 1.372 billions US$ on average for the two types of investments respectively). Beyond these measurement related considerations, the size of bilateral FDI and the number of investments may (or may not) react differently to variation in CP as they capture different aspects of internationalization. This is ultimately an empirical question. The reported results show that the impact of CP is still positive but generally lower when considering the value ($V_{ni,t}$) as dependent variable. Moreover, when both directions of trade are included simultaneously (third column), imports of cultural goods become statistically not significant as in the fully specified model. These combined findings suggest that the destination side mechanisms are relevant across different measures of bilateral volume of FDI, and that the decision on whether or not to invest is more sensitive to the asymmetric components of CP than the actual size of bilateral FDI.

Similar conclusions apply when we investigate the impact of asymmetric CP on the intensive margin of investment as captured by the average value of investment ($\bar{V}_{ni,t}$). The estimates reported in Table 8 indicate that, despite being halved in their magnitude, the coefficients of both $\ln \text{CultIMP}_{ni,t}$ and $\ln \text{CultEXP}_{ni,t}$ remain statistically significant at least at the 5% confidence level.

[18] In Felbermayr and Toubal (2010) the impact of cultural proximity on trade is more than ten times higher when instrumented. The gap between OLS and 2SLS estimates is even higher in the analysis of Guiso et al. (2009) when the dependent variable is FDI.

[19] See Table A-3 in Appendix A for a more precise assessment of the scope of imputation.

Table 6: Impact of Instrumented Cultural Proximity on Greenfield FDI

| Dep. Var. | Count $C_{ni,t}$ | | | |
|---|---|---|---|---|
| | 2 year lag | 5 year lag | Baseline | IV |
| | (1) | (2) | (3) | (4) |
| $\ln \text{CultIMP}_{ni,t}$ | | | 0.0658** | 0.0736 |
| | | | (2.96) | (1.35) |
| $\ln \text{CultEXP}_{ni,t}$ | | | 0.247*** | 0.619*** |
| | | | (9.43) | (6.54) |
| $\ln$ lagged $\text{CultIMP}_{ni,t-2}$ | 0.0740*** | | | |
| | (6.32 ) | | | |
| $\ln$ lagged $\text{CultEXP}_{ni,t-2}$ | 0.296*** | | | |
| | (21.27) | | | |
| $\ln$ lagged $\text{CultIMP}_{ni,t-5}$ | | 0.0784*** | | |
| | | (6.59 ) | | |
| $\ln$ lagged $\text{CultEXP}_{ni,t-5}$ | | 0.286*** | | |
| | | (19.51) | | |
| $\ln \text{dist}_{ni}$ | 0.179*** | 0.182*** | 0.806*** | 0.350** |
| | ( 5.08) | ( 5.17) | ( 11.26) | ( 2.70) |
| $\text{colony}_{ni}$ | 0.380*** | 0.385*** | 0.0193 | 0.0177 |
| | (7.14) | (7.23) | ( 0.23) | ( 0.18) |
| $\text{lang}_{ni}$ | 0.167** | 0.152** | 0.0723 | 0.0436 |
| | (3.26) | (2.99) | (0.70) | ( 0.30) |
| $\text{comrelig}_{ni}$ | 0.877*** | 0.872*** | 0.118 | 0.206 |
| | (9.02) | (8.99) | ( 0.95) | ( 1.49) |
| $\text{contig}_{ni}$ | 0.106 | 0.117 | 0.147* | 0.283*** |
| | ( 1.75) | ( 1.92) | ( 2.36) | ( 3.93) |
| $\text{comleg}_{ni}$ | 0.155*** | 0.157*** | 0.330*** | 0.219** |
| | (4.07) | (4.17) | (5.89) | (3.20) |
| $\text{FTA}_{ni,t}$ | 0.127* | 0.133* | 0.394*** | 0.0725 |
| | (2.34) | (2.45) | (3.49) | (0.48) |
| $\text{BIT}_{ni,t}$ | 0.00909 | 0.0311 | 0.172* | 0.0757 |
| | ( 0.23) | ( 0.78) | (2.23) | (0.83) |
| Imp×Year FE | √ | √ | √ | √ |
| Exp×Year FE | √ | √ | √ | √ |
| Obs | 84568 | 80057 | 10596 | 10040 |
| Estimator | PPML | PPML | PPML | IV PPML |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z statistics in parentheses. Standard errors are clustered by trading pair. The dependent variable "Count" $C_{ni,t}$ is the bilateral number of Greenfield FDI projects from country $i$ to country $n$. It includes the zero flows. The estimates in columns (1) to (3) are obtained by PPML using the *PPML panel sg* command written by Thomas Zylkin which simultaneously allows to absorb pair wise as well as origin by time and destination by time FEs. The model includes origin×time and destination×time FEs. Estimates in column (4) are computed via IVPPML using the *ivpoisson* command built in *STATA 13*. Due to convergence reasons, in column (3) and (4) the sample is reduced to the subset of importing and exporting countries as in Felbermayr and Toubal (2010). A drawback of IVPOISSON command is that it cannot handle high dimensional FE. Nonetheless, the estimates are consistent to a broader sample estimated with a reduced set of fixed effects (available upon request to the authors), suggesting that they are robust to different specifications.

### 4.2.3 Core versus optional cultural goods

Table 9 compares our benchmark results of Table 4 with the estimates obtained with only core and optional cultural goods, respectively. The distinction between core and optional hinges on

Table 7: Impact of Cultural Proximity on the Total Value of Greenfield FDI

| Dep. Var. | Value $V_{ni,t}$ | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| $\ln \text{CultIMP}_{ni,t}$ | 0.0984*** | | 0.0221 |
| | (4.82) | | (1.07) |
| $\ln \text{CultEXP}_{ni,t}$ | | 0.277*** | 0.269*** |
| | | (13.28) | (11.44) |
| $\ln \text{dist}_{ni}$ | -0.469*** | -0.248*** | -0.237*** |
| | (-9.64) | (-4.58) | (-4.44) |
| $\text{colony}_{ni}$ | 0.507*** | 0.370*** | 0.364*** |
| | (6.02) | (4.85) | (4.76) |
| $\text{lang}_{ni}$ | 0.180 | 0.115 | 0.109 |
| | (1.84) | (1.26) | (1.20) |
| $\text{comrelig}_{ni}$ | 1.370*** | 1.217*** | 1.210*** |
| | (9.02) | (8.46) | (8.42) |
| $\text{contig}_{ni}$ | -0.150 | -0.0863 | -0.0952 |
| | (-1.43) | (-0.86) | (-0.94) |
| $\text{comleg}_{ni}$ | 0.142* | 0.0775 | 0.0724 |
| | (2.41) | (1.37) | (1.28) |
| $\text{FTA}_{ni,t}$ | 0.302*** | 0.265*** | 0.260*** |
| | (3.96) | (3.60) | (3.52) |
| $\text{BIT}_{ni,t}$ | -0.0289 | -0.0441 | -0.0443 |
| | (-0.45) | (-0.73) | (-0.74) |
| Imp×Year FE | √ | √ | √ |
| Exp×Year FE | √ | √ | √ |
| Obs | 87448 | 87448 | 87448 |
| % Zeros | 0.749 | 0.749 | 0.749 |
| $R^2$ | 0.9056 | 0.9216 | 0.9221 |
| Estimator | PPML | PPML | PPML |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z-statistics in parentheses. Standard errors are clustered by trading-pair. The dependent variable "Value" $C_{ni,t}$ is the value of the aggregated bilateral flow of greenfield investments from country $i$ to country $n$, including zero flows. The estimates are obtained with PPML using the *PPML panel sg* command written by Thomas Zylkin which simultaneously allows to absorb pair-wise as well as origin-by-time and destination-by-time FEs. The model includes origin×time and destination×time FEs. The sample size in this table is invariant to the number of covariates included and refers to the regression which features both imports and exports of cultural goods. The information which belong to groups with all zeros or missing values are automatically dropped by the estimator as FEs cannot be computed. fDIMarket database provides information on the value of each greenfield. When no official figures are provided by the parent company, the value is estimated by FDIIntelligence unit. Information about the estimation algorithm can be found on fDIMarket website.

the cultural content embodied in these types of products: hence, it is reasonable to expect the impact of CP as mostly driven by the trade (in either direction) of core cultural goods as they are likely to better capture proximity in cultural tastes.However, optional cultural goods represent the lion share of cultural trade from and between developing countries: failing to account for these flows would exclude many South countries from the analysis, limiting the impact of CP on specific FDI channels (especially North-North). As shown in Table 9, the pattern of results is stable across different measures of cultural trade, showing the capacity of both types of cultural goods to reflect the same underlying forces.

The exercise proposed in Table 9 serve the additional purpose of minimizing potential concerns regarding the measurement error introduced by the gross nature of cultural trade used in the

Table 8: Impact of Cultural Proximity on Intensive Margin of Investment

| Dep. Var. | Average Value $\bar{V}_{ni,t}$ | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| $\ln \text{CultIMP}_{ni,t}$ | 0.0705*** | | 0.0390* |
| | (3.96) | | (2.11) |
| $\ln \text{CultEXP}_{ni,t}$ | | 0.147*** | 0.137*** |
| | | (6.99) | (6.11) |
| $\ln \text{dist}_{ni}$ | -0.308*** | -0.194*** | -0.166** |
| | (-6.48) | (-3.72) | (-3.20) |
| $\text{colony}_{ni}$ | 0.155 | 0.0529 | 0.0290 |
| | (1.29) | (0.45) | (0.25) |
| $\text{lang}_{ni}$ | 0.0701 | 0.0399 | 0.0222 |
| | (0.74) | (0.43) | (0.24) |
| $\text{comrelig}_{ni}$ | 0.825*** | 0.773*** | 0.750*** |
| | (5.79) | (5.16) | (5.09) |
| $\text{contig}_{ni}$ | 0.0805 | 0.0998 | 0.0874 |
| | (0.60) | (0.74) | (0.66) |
| $\text{comleg}_{ni}$ | 0.0513 | 0.0325 | 0.0215 |
| | (0.76) | (0.49) | (0.32) |
| $\text{FTA}_{ni,t}$ | 0.161 | 0.127 | 0.120 |
| | (1.75) | (1.43) | (1.34) |
| $\text{BIT}_{ni,t}$ | 0.295*** | 0.292*** | 0.284*** |
| | (4.50) | (4.44) | (4.33) |
| Imp×Year FE | √ | √ | √ |
| Exp×Year FE | √ | √ | √ |
| Obs | 87448 | 87448 | 87448 |
| % Zeros | 0.749 | 0.749 | 0.749 |
| $R^2$ | 0.4555 | 0.5016 | 0.4961 |
| Estimator | PPML | PPML | PPML |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z-statistics in parentheses. Standard errors are clustered by trading-pair. The estimates are obtained with PPML using the *PPML panel sg* command written by Thomas Zylkin which simultaneously allows to absorb pair-wise as well as origin-by-time and destination-by-time FEs. This table replicates results as in Table 4, but using "Average Value" $\bar{V}_{ni,t}$, the average value of bilateral greenfield investments from country $i$ to country $n$, as dependent variable. fDIMarket database provides information on the value of each greenfield. When no official figures are provided by the parent company, the value is estimated by FDIIntelligence unit. Information about the estimation algorithm can be found on fDIMarket website.

analysis. Indeed, facing a world trading system where global supply chains are prevalent, one may argue that Chinese exports of fashion products or toys (included in the category of optional cultural goods) to an import country not only (and not necessarily) reflect Chinese cultural content, and therefore the cultural attractiveness of China for the importer, but also some third country's cultural content embedded in the fashion or pottery design performed in that country before actual manufacturing happening in China. This concern is legitimate as long as few countries in our sample have a comparative advantage in the manufacture of a number of cultural products, fostering a disproportionate concentration of production in (and export from) these countries of cultural goods embedding foreign cultural value added. This might actually the case for the several Asian countries and for some of the products included in the sub-category of optional cultural goods (see Table 1). It is well known that countries in the so called Factory Asia have an international specialisation in the manufacturing of low tech goods, including

Table 9: Different Measures of CP: Core VS Optional Cultural Trade

| Dep. Var. | Count $C_{ni,t}$ | | |
|---|---|---|---|
| | Total cultural trade | Core cultural trade | Optional cultural trade |
| | (1) | (2) | (3) |
| $\ln \text{CultIMP}_{ni,t}$ | 0.0690*** | 0.0925*** | 0.0525*** |
| | (5.90) | (8.22) | (4.34) |
| $\ln \text{CultEXP}_{ni,t}$ | 0.305*** | 0.285*** | 0.249*** |
| | (21.91) | (20.18) | (19.43) |
| $\text{FTA}_{ni,t}$ | 0.118* | 0.0990 | 0.110 |
| | (2.19) | (1.89) | (1.93) |
| $\text{BIT}_{ni,t}$ | 0.0115 | 0.0329 | -0.0174 |
| | (0.29) | (0.83) | (-0.41) |
| $\ln \text{dist}_{ni}$ | -0.179*** | -0.198*** | -0.225*** |
| | (-5.13) | (-5.75) | (-6.38) |
| $\text{colony}_{ni}$ | 0.366*** | 0.244*** | 0.488*** |
| | (6.85) | (4.49) | (8.65) |
| $\text{lang}_{ni}$ | 0.181** | 0.161** | 0.216*** |
| | (3.53) | (3.12) | (4.01) |
| $\text{comrelig}_{ni}$ | 0.883*** | 0.711*** | 1.012*** |
| | (9.21) | (7.64) | (9.92) |
| $\text{contig}_{ni}$ | -0.0977 | -0.139* | -0.111 |
| | (-1.61) | (-2.44) | (-1.74) |
| $\text{comleg}_{ni}$ | 0.153*** | 0.184*** | 0.153*** |
| | (4.06) | (4.97) | (3.88) |
| Imp×Year FE | √ | √ | √ |
| Exp×Year FE | √ | √ | √ |
| Obs | 87448 | 67192 | 76951 |
| % Zeros | 75% | 69% | 71% |
| $R^2$ | 0.91 | 0.92 | 0.91 |
| Estimator | PPML | PPML | PPML |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z-statistics in parentheses. Standard errors are clustered by trading-pair. The dependent variable "Count" $C_{ni,t}$ is the bilateral number of Greenfield FDI projects from country $i$ to country $n$. It includes the zero flows.
The estimates are obtained by PPML using the *PPML panel sg* command written by Thomas Zylkin which simultaneously allows to absorb pair-wise as well as origin-by-time and destination-by-time FEs. The model includes origin×time and destination×time FEs. The first column replicates column (3) of table 4. The second column refers to the effect on greenfield FDI of 'core' cultural trade, while the third refers to 'optional' cultural trade, as defined by UNCTAD (2010)
The sample is reduced due to the large number of null values that are dropped when taken in logarithmic form.

for instance toys (see Baldwin and Lopez-Gonzalez, 2015). The average revealed comparative advantage (RCA) across optional cultural goods for the period of our analysis is equal to 1.2 for China and above the threshold value of 1 also for India, Indonesia, Malaysia, Thailand and Vietnam.[20] The stability of our results across core and optional cultural goods suggests that this potential source of measurement error is not biasing our results as the average RCA of the listed Asian countries across core cultural goods is always well below one (for instance it is equal to 0.378 of China and 0.165 for Vietnam).[21]

[20] RCA is computed following the Balassa index.

[21] A better test of the implications of relying on gross cultural trade would require the use of value added trade data. Unfortunately available sources such as the OECD/WTO TiVA database fail to match the country coverage and product desegregation required by the research design of the present study.

# 5 Extensions

This section proposes the two extensions to the analysis conducted so far. First we propose two empirical tests of the 'destination-side' mechanisms as introduced in the conceptual framework laid out in Section 3.1. Then, we test whether the role of asymmetric and time-dependent component of CP changes at different levels of its symmetric and time-invariant component.

## 5.1 Destination-side mechanisms

The empirical analysis so far has established the relative importance of the two directions of asymmetric CP in explaining Greenfield FDI from an origin country $i$ to a destination $n$. In particular the attractiveness of the $i$'s culture for individuals in $n$ - $A_{in,t}$ proxied by $CultEXP_{ni,t}$ - seems to play a much stronger role than the attractiveness of the destination for the origin, $A_{ni,t}$ proxied by $CultIMP_{ni,t}$. This is somehow at odds with the standard theories of bilateral FDI which tend to focus on 'origin-side' mechanisms and calls for a more careful consideration of 'destination-side' mechanisms. In this section we propose an empirical test of the 'destination consumers demand' and the 'destination political economy' channels introduced in Section 3.1.

According to the 'destination consumers demand' channel, $A_{in,t}$ can be relevant to explain FDI from $i$ to $n$ because the preferences of consumers in $n$ for the the affiliate's production in their country would be a positive function of $i$'s cultural attractiveness for them. This leads us to expect $A_{in,t}$ to be more relevant with respect to $A_{ni,t}$ when the FDI projects are intended to target consumer demand in the destination country rather than to serve as an intermediary step in a global supply chain type of production. In the case of horizontal FDI the attractiveness of the origin's culture for consumers in the destination country could be a stronger driver of the investment decision as it might positively affect the expected revenues of the FDI project. This is confirmed empirically by the estimation results presented in Table 10.

Both columns replicate results as in column (3) of Table 4 on two different subsamples. Column (1) includes only FDI projects in those sectors that are more likely to target the consumers demand in the destination country, i.e. that include consumption (final) goods and services. Conversely, the estimation sample used to derive the results presented in column (2) is restricted to those sectors where the importance of local consumption is lower compared to the location advantages of different kind: such sectors include mainly intermediate goods.[22] Taking the ratio between the point estimates of the coefficients for $\ln CultEXP_{ni,t}$ and $\ln CultIMP_{ni,t}$ as a measure of the relative importance of $A_{in,t}$ with respect to $A_{ni,t}$ in explaining $C_{ni,t}$ we notice that this ratio is higher when the estimation sample is restricted to those sectors that are more likely to target the consumers demand in the destination country. We take this a suggestive evidence of

[22]The estimation sample in the first column includes only FDI projects classified in the following sectors: beverages, consumer electronics, consumer product, financial services, food and tobacco, leisure and entertainment, software and ICT devices, and transportation. T'he estimation sample in the second column instead includes only the following sectors: automotive components, biotech, building and construction material, ceramics, glasses, chemical, coal, oil gas, electronic component, engines and turbines, industrial machinery, metals, minerals, plastic, rubber, semiconductors.

Table 10: Destination Consumers Demand Channel

| Dep. Var. | Count $C_{ni,t}$ | |
|---|---|---|
| FDI targeting consumers in $n$ | More likely | Less likely |
| | (1) | (2) |
| $\ln \text{CultIMP}_{ni,t}$ | 0.0768*** | 0.0731*** |
| | (5.85) | (4.12) |
| $\ln \text{CultEXP}_{ni,t}$ | 0.317*** | 0.255*** |
| | (20.12) | (14.70) |
| $\ln \text{dist}_{ni}$ | -0.258*** | -0.0730 |
| | (-7.34) | (-1.42) |
| $\text{colony}_{ni}$ | 0.315*** | 0.369*** |
| | (4.48) | (5.50) |
| $\text{lang}_{ni}$ | 0.244*** | 0.0386 |
| | (3.97) | (0.46) |
| $\text{comrelig}_{ni}$ | 1.047*** | 0.872*** |
| | (9.60) | (6.50) |
| $\text{contig}_{ni}$ | -0.153* | -0.0963 |
| | (-2.21) | (-1.13) |
| $\text{comleg}_{ni}$ | 0.204*** | 0.0174 |
| | (4.64) | (0.31) |
| $\text{FTA}_{ni,t}$ | 0.0138 | 0.171* |
| | (0.24) | (2.15) |
| $\text{BIT}_{ni,t}$ | 0.0467 | -0.0522 |
| | (1.10) | (-0.83) |
| Imp×Year FE | √ | √ |
| Exp×Year FE | √ | √ |
| Obs | 78697 | 62989 |
| % Zeros | 0.82 | 0.83 |
| $R^2$ | 0.90 | 0.88 |
| Estimator | PPML | PPML |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z-statistics in parentheses. Standard errors are clustered by trading-pair. Both columns replicate results as in column (3) of Table 4. The estimation sample in the first column includes only FDI projects classified in the following sectors: beverages, consumer electronics, consumer product, financial services, food and tobacco, leisure and entertainment, software and ICT devices, and transportation. the estimation sample in the second column instead includes only the following sectors: automotive components, biotech, building and construction material, ceramics, glasses, chemical, coal, oil gas, electronic component, engines and turbines, industrial machinery, metals, minerals, plastic, rubber, semiconductors.

the existence of the hypothesized 'destination consumers demand' channel in determining the role of CP for FDI.

The 'destination political economy' channel, on the other hand, rationalizes the role of $A_{in,t}$ in determining greenfield FDI from $i$ to $n$, through the potential political and economic support granted by the government in $n$ to an FDI project coming from $i$. In a political economy model this would need to respond - at least to some extent - to the preferences of voters in $n$, affected by their appreciation of the culture in $i$. This mechanism implies a stronger relative importance of the origin's cultural attractiveness for the destination when politicians in the destination country are subject to a higher degree of accountability with respect to their citizens, i.e. when their allocation of support across projects coming from different sources is likely to more closely

reflect voters' preferences. The estimates reported in Table 11 represent an empirical test of this implication.

Table 11: Destination Political Economy Channel

| Dep. Var. | Count $C_{ni,t}$ | |
|---|---|---|
| Accountability in $n$ | Low | High |
| | (1) | (2) |
| $\ln \text{CultIMP}_{ni,t}$ | 0.107*** | 0.0526 |
| | (6.03) | (1.36) |
| $\ln \text{CultEXP}_{ni,t}$ | 0.294*** | 0.498*** |
| | (13.91) | (9.35) |
| $\ln \text{dist}_{ni}$ | -0.419*** | -0.425* |
| | (-5.72) | (-2.28) |
| $\text{colony}_{ni}$ | 0.528*** | 0.761** |
| | (3.44) | (3.09) |
| $\text{lang}_{ni}$ | 0.323*** | -0.135 |
| | (3.48) | (-0.51) |
| $\text{comrelig}_{ni}$ | 0.870*** | -0.234 |
| | (4.29) | (-0.38) |
| $\text{contig}_{ni}$ | -0.0701 | 0.112 |
| | (-0.52) | (0.35) |
| $\text{comleg}_{ni}$ | 0.145 | -0.120 |
| | (1.60) | (-0.62) |
| $\text{FTA}_{ni,t}$ | 0.0298 | 1.276*** |
| | (0.30) | (3.51) |
| $\text{BIT}_{ni,t}$ | 0.0822 | -0.0827 |
| | (0.93) | (-0.54) |
| Imp×Year FE | √ | √ |
| Exp×Year FE | √ | √ |
| Obs | 9817 | 2376 |
| % Zeros | 0.76 | 0.68 |
| $R^2$ | 0.85 | 0.99 |
| Estimator | PPML | PPML |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z-statistics in parentheses. Standard errors are clustered by trading-pair. Both columns replicate the specification as in column (3) of Table 4. The estimation sample used to derive the estimates reported in the first (second) column is restricted to destination countries with an accountability score below (above) the sample median. Accountability is measured with the accountability index, from the World Bank CPIA indicators on Corruption, Accountability and Transparency perception.

Both columns replicate the specification as in column (3) of Table 4. The estimation sample used to derive the estimates reported in the first (second) column is restricted to destination countries with an accountability score below (above) the sample median. Accountability is measured with the accountability index, from the World Bank CPIA indicators on Corruption, Accountability and Transparency perception. The ratio between the point estimates of the coefficients for $\ln \text{CultEXP}_{ni,t}$ and $\ln \text{CultIMP}_{ni,t}$ is higher for the subsample of high accountability destination countries, suggesting a relative higher importance of $A_{in,t}$ when politicians in the destination country are more accountable vis-à-vis their citizens and therefore providing empirical evidence for the existence of the hypothesized 'destination political economy' channel.

## 5.2 Heterogeneous impact of the asymmetric and time-dependent dimension of CP

This section tests how the asymmetric and time-dependent component of CP affects bilateral investment flows at different degree of the symmetric and time-invariant component of CP. In order to do so, we explore the effect of trade in cultural goods at different values (above and below the median value) of three symmetric, and time-invariant measures of cultural proximity previously used in the literature: religious proximity, the Melitz and Toubal (2014) "Common Spoken Language" (CSL) measure of linguistic proximity, and the composite index of linguistic proximity (AP Index) by Adsera and Pytlikova (2015).[23] Moreover, to identify the impact of time-contingent shocks in CP all regressions include a full set of fixed effects as in Table 5. The inclusion of dyadic fixed effects absorbs all the cross section variability in our sample, a necessary feature if we are interested in exploring the time-varying dimension of cultural trade. Results are reported in Table 12 below.

Table 12: Heterogeneous impact of the asymmetric and time-dependent dimension of CP

| Dep. Var. | Count $C_{ni,t}$ | | | | | |
|---|---|---|---|---|---|---|
| | Religion[1] | | CSL[2] | | AP index[3] | |
| | (1-50 pct) | (51-100 pct) | (1-50 pct) | (51-100 pct) | (1-50 pct) | (51-100 pct) |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\ln \text{CultIMP}_{ni,t}$ | 0.00639 | -0.000994 | 0.00920 | -0.0151 | -0.00908 | -0.0434 |
| | (0.53) | (-0.07) | (0.82) | (-1.03) | (-0.57) | (-0.92) |
| $\ln \text{CultEXP}_{ni,t}$ | 0.0554*** | 0.0122 | 0.0604*** | 0.00995 | 0.0713*** | -0-0779 |
| | (3.34) | (0.75) | (3.59) | (0.66) | (3.51) | (-1.26) |
| $\text{FTA}_{ni,t}$ | 0.136* | -0.0640 | 0.0315 | -0.0336 | 0.0130 | -0.0475 |
| | (2.06) | (-1.09) | (0.50) | (-0.66) | (0.14) | (-0.55) |
| $\text{BIT}_{ni,t}$ | 0.0273 | 0.0754 | 0.223* | 0.0187 | 0.0859 | 0.289 |
| | (0.27) | (0.65) | (2.32) | (0.19) | (0.64) | (0.77) |
| Imp×Year FE | √ | √ | √ | √ | √ | √ |
| Exp×year FE | √ | √ | √ | √ | √ | √ |
| Country Pair FE | √ | √ | √ | √ | √ | √ |
| Obs | 23209 | 23916 | 22657 | 23465 | 12487 | 23465 |
| % Zeros | 59.78% | 55.25% | 64.04% | 51.00% | 45.77% | 4.47% |
| $R^2$ | 0.9687 | 0.9770 | 0.9721 | 0.9791 | 0.9730 | 0.9895 |
| Estimator | PPML | PPML | PPML | PPML | PPML | PPML |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z-statistics in parentheses. Standard errors are clustered by trading-pair. The dependent variable "Count" $C_{ni,t}$ is the number of the aggregated bilateral flow of greenfield investments from country $i$ to country $n$, including null flows. The estimates are obtained with PPML using the *PPML panel sg* command written by Thomas Zylkin which simultaneously allows to absorb pair-wise as well as origin-by-time and destination-by-time FEs.
[1] Division along the median of the distribution of religious proximity between country $i$ and country $n$.
[2] Division along the median of the distribution of Common Spoken Language as in Melitz and Toubal (2014) between country $i$ and country $n$.
[3] Division along the Composite Index of Linguistic Proximity as in Adsera and Pytlikova (2015) between country $i$ and country $n$.

Consistently with the results presented in Table 5 the reported estimates suggest that only time

[23] The choice of these measures is constrained by our intention to split the estimation sample. The majority of the "traditional" measures used in existing literature have a binary structure and for this reason they are not suitable to split our sample in a simple and effective way.

contingent shocks in terms of cultural attractiveness of the origin country for the destination seem to trigger investments. However, it seems that those results are mainly driven by pairs characterized by low level of time-invariant and symmetric CP: time contingent shocks to the cultural attractiveness of the origin country for the destination only play a role when the level of pre-existing or historical cultural ties is relatively weak. This is consistent with a relationship of substitutability between time-contingent, asymmetric and time-invariant, symmetric dimensions of CP in triggering FDI, with the former operating as a bridgehead between otherwise culturally distant countries.

# 6 Conclusions

Many countries are pursuing policies to attract foreign direct investments because they reckon FDI will contribute to their economic growth by creating a more competitive business environment, triggering technology spillovers, increasing capital accumulation and generating more job opportunities. The growth-enhancing role of FDI is well documented in the literature and is particularly evident for developing countries. Over the last 15 years the share of FDI originating from developing countries over total flows has increased from 8% to 26% while recent research has showed that much of this investment takes place between developing economies (Gold et al., 2017)

The overall economic benefits of FDI have motivated a thorough investigation of its determinants and CP has been established as an important driver of the firm's decision to invest abroad. However, the definition of CP used assumed that it was symmetric and stable over time. The resulting standard measures of CP - including the composite indexes (as the one proposed by Kogut and Singh, 1988, based on Hofstede, 2003's cultural dimensions) - employed in the existing empirical studies are therefore inadequate to capture a broader and more refined notion of CP. In this paper we have assessed the effect of CP on greenfield FDI explicitly accounting for its asymmetric and time-dependent dimensions. In line with Disdier et al. (2010), we used bilateral trade in cultural goods as a proxy for asymmetric and time-dependent CP. The exercise contributes to the literature as the effects of asymmetric bilateral cultural measures remain largely understudied and the few papers that include FDI as outcome variable as well as an asymmetric measure of bilateral cultural relationship have been confined mainly to samples of OECD economies. The use of two comprehensive datasets on trade and greenfield FDI - namely BACI (CEPII) and Financial Times FDI Market dataset, respectively - allows the present study to feature a very extended country coverage which also includes South-South FDI, for which CP may be particularly relevant.

Relying on the PPML estimation technique with high-dimensional fixed effects our results have shown that asymmetry in cross-country cultural proximity matters for FDI flows: more precisely, investment projects from a source to a destination country tend to increase more with cultural exports from source to destination rather than with imports. In other words, the evidence points to a stronger role of the cultural attractiveness of the country where the investment is coming from for individuals in the destination economy. This result suggests that higher relevance in

explaining patterns of FDI should be attributed to the cultural preferences of the individuals in the destination country, both as consumers potentially buying the outcome produced by the subsidiary as well as voters, affecting the allocation of political pressures across competing investment projects.

Our analysis leaves at least two interesting questions open to future research. First, while the study of asymmetry in CP is limited in the context of the present paper to a descriptive assessment, it undoubtedly proves that such phenomenon exists in the data, namely that cultural relationships are indeed asymmetric. More can be done to identify a statistically robust and convincing measure of the degree of asymmetry in cultural relationships and to study its determinants and effects in the realm of economic phenomena. Second, our findings shed new light on the role played by individuals in the destination country to trigger inward FDI. While this paper focuses on the cultural dimension of these preferences, further theoretical investigation can be conducted to broadly assess their contribution within a fully micro-founded general equilibrium model of bilateral FDI.

## References


**Abu-Lughod, Lila**, "Writing against culture," in Richard G. Fox, ed., *Recapturing Anthropology: Working in the Present*, School of American Research Press., 1996, pp. 137–162.

**Adsera, Alicia and Mariola Pytlikova**, "The role of language in shaping international migration," *The Economic Journal*, 2015, *125* (586), F49–F81.

**Alesina, Alberto and Paola Giuliano**, "Culture and Institutions," *Journal of Economic Literature*, December 2015, *53* (4), 898–944.

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

**Appadurai, Arjun**, *Modernity at Large: Cultural Dimensions of Globalization*, Minneapolis-London: University of Minnesota Press, 1996.

**Artuç, Erhan, Frédéric Docquier, Çaglar Özden, and Christopher Parsons**, "A global assessment of human capital mobility: the role of non-OECD destinations," *World Development*, 2015, *65*, 6–26.

**Aubry, Amandine, Maurice Kugler, and Hillel Rapoport**, "Migration, FDI and the Margins of Trade," *Boston: Center for International Development, Harvard University*, 2014.

**Baldwin, Richard and Javier Lopez-Gonzalez**, "Supply-chain Trade: A Portrait of Global Patterns and Several Testable Hypotheses," *The World Economy*, 2015, *38* (11), 1682–1721.

**Beine, Michel, Simone Bertoli, and Jesús Fernández-Huertas Moraga**, "A practitioners' guide to gravity models of international migration," *The World Economy*, 2016, *39* (4), 496–512.

**Belot, Michèle and Sjef Ederveen**, "Cultural barriers in migration between OECD countries," *Journal of Population Economics*, 2012, *25* (3), 1077–1105.

**Bertoli, Simone and Jesús Fernández-Huertas Moraga**, "Multilateral resistance to migration," *Journal of Development Economics*, 2013, *102*, 79–100.

**Blonigen, Bruce A. and Jeremy Piger**, "Determinants of Foreign Direct Investment," *Canadian Journal of Economics*, 2014, *47* (3), 775–812.

**Briant, Anthony, Pierre-Philippe Combes, and Miren Lafourcade**, "Product complexity, quality of institutions and the protrade effect of immigrants," *The World Economy*, 2014, *37* (1), 63–85.

**Combes, Pierre-Philippe, Miren Lafourcade, and Thierry Mayer**, "The trade-creating effects of business and social networks: evidence from France," *Journal of international Economics*, 2005, *66* (1), 1–29.

**de Sousa, José and Julie Lochard**, "Does the Single Currency Affect Foreign Direct Investment?," *The Scandinavian Journal of Economics*, 9 2011, *113* (3), 553–578.

**Desbordes, Rodolphe and Shang-Jin Wei**, "The effects of financial development on foreign direct investment," *Journal of Development Economics*, 2017, *127*, 153–168.

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

**Du, Julan, Yi Lu, and Zhigang Tao**, "Institutions and FDI location choice: The role of cultural distances," *Journal of Asian Economics*, 2012, *23* (3), 210–223.

**Feenstra, Robert C**, *Advanced international trade: theory and evidence*, Princeton university press, 2015.

**Felbermayr, Gabriel J. and Farid Toubal**, "Cultural proximity and trade," *European Economic Review*, 2010, *54* (2), 279 – 293.

**Fernández, Raquel**, "Culture and Economics," in Steven N. Durlauf and Lawrence E. Blume, eds., *The New Palgrave Dictionary of Economics*, Basingstoke: Palgrave Macmillan, 2008.

**_** , "Does Culture Matters?," in Jess Benhabib, Matthew O. Jackson, and Alberto Bisin, eds., *Handbook of Social Economics*, Vol. 1A, Elsevier, Elsevier, North-Holland, 2011, pp. 481–510.

**Giovannetti, Giorgia and Mauro Lanati**, "Do High-Skill Immigrants trigger High-Quality Trade?," *The World Economy*, 2016.

**Giuliano, Paola, Antonio Spilimbergo, and Giovanni Tonon**, "Genetic distance, transportation costs, and trade," *Journal of Economic Geography*, 2014, *14* (1), 179–198.

**Gold, Robert, Holger Görg, Aoife Hanley, and Adnan Seric**, "South–South FDI: is it really different?," *Review of World Economics*, 2017, pp. 1–17.

**Gould, David M**, "Immigrant links to the home country: empirical implications for US bilateral trade flows," *The Review of Economics and Statistics*, 1994, pp. 302–316.

**Guiso, Luigi, Paola Sapienza, and Luigi Zingales**, "Does Culture Affect Economic Outcomes?," *Journal of Economic Perspectives*, June 2006, *20* (2), 23–48.

**_ , _ , and _** , "Cultural biases in economic exchange?," *The Quarterly Journal of Economics*, 2009, *124* (3), 1095–1131.

**Head, Keith and John Ries**, "FDI as an outcome of the market for corporate control: Theory and evidence," *Journal of International Economics*, 2008, *74* (1), 2 – 20.

**_ and Thierry Mayer**, "Gravity Equations: Workhorse, Toolkit, and Cookbook," in Gita Gopinath, Elhanan Helpman, and Kenneth S. Rogoff, eds., *Handbook of International Economics*, Vol. 4, Elsevier, 2014, chapter 3.

**Hofstede, Geert**, *Cultures and organizations. Intercultural cooperation and its importance for survival. Software of the mind*, Mc Graw-Hill, 1991.

**_** , *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*, Sage publications, 2003.

**House, Robert J, Paul J Hanges, Mansour Javidan, Peter W. Dorfman, and Vipin Gupta**, *Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies*, Sage publications, 2004.

**Javorcik, Beata S, Çağlar Özden, Mariana Spatareanu, and Cristina Neagu**, "Migrant networks and foreign direct investment," *Journal of Development Economics*, 2011, *94* (2), 231–241.

**Kogut, Bruce and Harbir Singh**, "The effect of national culture on the choice of entry mode," *Journal of international business studies*, 1988, pp. 411–432.

**Larch, Mario, Joschka Wanner, Yoto Yotov, and Thomas Zylkin**, "The Currency Union Effect: A PPML Re-assessment with High-Dimensional Fixed Effects," School of Economics Working Paper Series 2017-7, LeBow College of Business, Drexel University April 2017.

**Lee, Hyun-Hoon and John Ries**, "Aid for Trade and Greenfield Investment," *World Development*, 2016, *84*, 206–218.

**Leung, Kwok, Rabi S. Bhagat, Nancy R. Buchan, Miriam Erez, and Cristina B. Gibson**, "Culture and International Business: Recent Advances and Their Implications for Future Research," *Journal of International Business Studies*, 2005, *36* (4), 357–378.

**Li, Chengguang, Felix C. Brodbeck, Oded Shenkar, Leonard J. Ponzi, and Jan Hendrik Fisch**, "Embracing the foreign: Cultural attractiveness and international strategy," *Strategic Management Journal*, 2017, *38* (4), 950–971.

**Mayer, Thierry and Gianmarco IP Ottaviano**, "The happy few: The internationalisation of european firms," *Intereconomics*, 2008, *43* (3), 135–148.

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

**Santos Silva, João M C and Silvana Tenreyro**, "The log of gravity," *The Review of Economics and statistics*, 2006, *88* (4), 641–658.

**_ and _** , "Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator," *Economics Letters*, 2011, *112* (2), 220–222.

**Shenkar, Oded**, "Cultural distance revisited: Towards a more rigorous conceptualization and measurement of cultural differences," *Journal of international business studies*, 2001, *32* (3), 519–535.

**Tung, Rosalie L and Alain Verbeke**, "Beyond Hofstede and GLOBE: Improving the quality of cross-cultural research," *Journal of International Business Studies*, Oct 2010, *41* (8), 1259–1274.

**UNCTAD**, *Creative Economy Report. A Feasible Development Option*, UNDP, 2010.

_ , "51 World Investment Report," Technical Report, UNCTAD 2017.

**Yotov, Yoto V, Roberta Piermartini, José-Antonio Monteiro, and Mario Larch**, "An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model," *World Trade Organization, Geneva*, 2016.

# Appendices

## A Data: sources and general features

The data used throughout both the descriptive and the analytical parts of the paper come from a variety of sources. Table A-1 displays the major sources and related links where additional information on the different databases used to create our final dataset: most of the other data come from sources that are well known in empirical gravity literature.

The focus of the analysis is on testing the role and the extent of the non-reciprocal component of CP on international economic flows, with the specific focus on greenfield FDI. For this reason we aggregate the projects according to the country of origin, destination and year in which the investment has been made. Then, we label missing dyadic flows at this stage as null investment channels, to obtain a square bilateral FDI matrix accounting for 184×185 countries of origin and destination. Cultural Trade data are then merged accordingly. Given that some territorial units in fDIMarket are not matched in BACI, some countries are dropped throughout the empirical analysis (see Table F-1 in Appendix B with the complete list of unmatched and excluded countries). In this respect, our strategy is similar to the one adopted by Aubry et al. (2014), Desbordes and Wei (2017), and Lee and Ries (2016) among the others. As a consequence, our FDI data reveals a pattern that is consistent with the findings from the recent theoretical and empirical literature in international economics (see for instance Mayer and Ottaviano, 2008), i.e. that only few firms are able to undertake FDI as a form of internationalization. [24]

Table A-1: Main Sources of Data used in the Empirical Section

| Variables | Dataset / Source / Website / Reference and Accessibility |
|---|---|
| **FDI Variables** | FDIMarket / FDI Intelligence Unit, The Financial Times / `http://www.fdiintelligence.com/` / FDI Market License |
| **Trade Variables** | BACI / CEPII / `http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=1` / UN COMTRADE access required |
| **Gravity Variables** | Gravdata / CEPII / `http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=8` / Free |
| **Bilateral Distance** | Geodist / CEPII / `http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=6` / Free |
| **Migrant Stock** | WB Global Bilateral Migration Dataset / The World Bank / `http://data.worldbank.org/data-catalog/global-bilateral-migration-database` / Artuç et al. (2015) / Free |
| **Language I** | Lingweb / CEPII / `http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=19` / Melitz and Toubal (2014) / Free |
| **Language II** | Data S1 / The Economic Journal / `http://onlinelibrary.wiley.com/doi/10.1111/ecoj.12231/abstract` / Adsera and Pytlikova (2015) / Free |
| **Cultural Distance** | Hofstede Index / The Journal of Population Economics / `https://link.springer.com/article/10.1007/s00148-011-0356-x` / Belot and Ederveen (2012) / Free |
| **BITs** | UNCTAD Investment Policy Hub / `http://investmentpolicyhub.unctad.org/IIA` / Free |
| **CPIA** | Country Policy and Institutional Assessment / The World Bank / `https://data.worldbank.org/data-catalog/CPIA` / Free |

Notes: This table lists the main sources in the data used throughout the dataset. Additional information are available upon request to the corresponding author. Concerning the sources of the single variables referring to a particular dataset used in this paper, the authors encourage to search directly in the websites provided.

However, the databases related to our variables of interest, cultural trade and greenfield FDI respectively, present some peculiarities that demand for some crucial choices in terms of data aggregation and classification, in order to obtain the least distortionary measures possible. In

[24] As this is particularly true for greenfield FDI, the result is that null bilateral flows account for more than 94% of the possible bilateral channels in our dataset. See Table A-2 below for a detailed report concerning the incidence of null flows in our dataset.

the remaining of this section, we explore the main issues related to cultural trade (that constitute our main variable of interest) and greenfield FDI respectively.

**Data on trade in cultural goods** Trade data come from the BACI dataset by CEPII[25], a proper workhorse in empirical gravity analysis in international trade. It is not the purpose of this appendix to describe the features of the BACI dataset as it is, for which we suggest the interested reader to check directly on the web link provided in Table A-1 above. Much more interesting for the purpose of this paper is to define what can be labelled as *Cultural Good* and what classification scheme is batter able to fit to the purpose of this paper that is, to investigate the role of imperfect reciprocity in cultural proximity in international economic flows.

Many countries and international organizations developed their own classification scheme, based on precise principles and content of the single class of product: for this reason, identifying the most suitable scheme for the sample considered in this paper is not an easy task. Yet, the choice of the classification is particularly sensible. Given the world coverage of our analysis, we restricted our search to two alternative classifications for cultural goods promoted by United Nations agencies, the UNESCO and the UNCTAD, [26] each of them based on different criteria and different categories of goods to be included in the count. Disdier et al. (2010) classified cultural goods using the definition proposed by UNESCO. Despite we build upon their seminal work, we depart from that approach and adopt the scheme proposed by UNCTAD (2010). There exist two main reasons for this choice: (i) a technicality related to the *time coverage* of the data, and (ii) a more substantial issue concerning the *sample selection*.

As for time coverage, the decision to prefer the UNCTAD classification leans on the different coding system adopted by the two different classifications. With respect to this point, UNESCO adopts the 2007 Harmonized Commodity Description and Coding System (HS 2007), that would call for the adoption of a conversion table to arrange the data along our time period. Conversely, UNCTAD (2010) adopts the HS 2002 coding system, that is more suitable for the time period at stake, as it allows not to convert the trade flows prior of 2007.[27] The conversion may distort the data, since the way they are collected is not always consistent across different coding systems: for this reason, the adoption of the UNCTAD (2010) classification could turn out to be not only less burdensome from a computational point of view, but also less prone to distortions.

Much more relevant for the choice of the classification scheme is the the sample coverage issue. The dataset used throughout this paper has global coverage,[28] with a large number of developing and transition economies in addition to developed ones. Conversely, Disdier et al. (2010) confine their analysis to a much more homogeneous group of OECD countries. This could not seem a major concern, but it is important to acknowledge that cultural goods are neither homogeneous nor equally produced worldwide. Both UNESCO and UNCTAD classifications uphold this fact by splitting cultural goods into **"core"** and **"optional"** cultural goods, with the former generally dominated by developed economies. By construction, in both classifications "optional" cultural goods encompass a wide range of products that are more likely to be produced in, and traded by developing countries too.[29]

[25] `http://www.cepii.fr/cepii/en/bdd_modele/presentation.asp?id=1`

[26] Other criteria can be found in the classification schemes developed by national and smaller international institutions (see UNCTAD, 2010, for a review).

[27] Nonetheless, as we adopted lag values of cultural trade as instruments in our IV analysis, we could not eventually avoid the burden of converting trade data prior to 2002. See Section 4.2.1.

[28] See Appendix F for the list of excluded countries.

[29] The definition of "core" goods made in UN agencies' and sovra-national organizations' classifications in general derive from this consideration, since most of the minor classification tend to include those "high cultural content" goods in their schemes. Conversely, "optional" goods refers to those goods that are included by certain countries or agencies' classification, but not by others (the inclusion of a class of goods depends on the productive system of the country that develop the classification). However, since all those schemes refer to developed countries,

A potential drawback of the wider conception of what can be considered as cultural good is that the UNCTAD classification has a much more diluted cultural content when compared to the UNESCO's. In fact, despite the latter encompasses a narrower set of traded goods, they are the ones with the larger cultural content. Nonetheless, given the world coverage of our sample, developed countries account for less than 30% of the whole set of countries included. For this reason, in order to balance the cultural composition of trade flows, and to construct a comparable measure of cultural trade across different development stages, the classification that is able to guarantee a relatively higher weight to those goods more evenly distributed across developed, developing and least developed economies should be preferred. This problem was not relevant in Disdier et al. (2010) because of the relative homogeneity of the sampled countries. Comparing the two classifications suggests that "core" goods account for 60% of total cultural goods in UNESCO's classification; barely 20% in UNCTAD's. For this reason, "[...]the UNESCO classification is better at capturing the experience of countries in the global North, while UNCTAD's better reflects opportunities for countries in the South.[...]" (UNCTAD, 2010, p. 111). This issue is more explicitly tackled in Section 4.2.3, where separate regressions on "core" and "optional" trade are run separately and compared to the results of our benchmark specification, where cultural goods encompasses both groups of goods.

**Greenfield FDI data** Data on FDI (that constitute the dependent variable in our empirical analysis) come from the *fDIMarket* database, that includes a detailed collection of all (and only) greenfield investments occurred worldwide in the period 2003-2014 (the first available year for greenfield FDI and the last year available for cultural trade data - our variable of interest - respectively). In figures, *fDIMarket* contains more than $169,000$ investment projects, carried on by roughly 67700 different companies worldwide in the period considered. The dataset include a large amount of information related to each recorded investment, included the the declared capital expenditure and the estimated number of jobs created at the moment the investment is carried out. Beyond the "quantitative" information, the dataset includes several additional investment-level entries such as location (up to NUTS 3 level of disaggregation), economic activity of the parent company as well as the (broad) sector in which such activity can be associated to in the host country. The high level of detail would ideally allow a much finer aggregation than the broad national-sectoral unit most of the more common datasets allow, but this type of analysis goes beyond the scope of the current research.

However, despite the exceptionally wide coverage of the dataset and its reliability in terms of missing records,[30] *fDIMarket* data present some important issues that worth to be introduced.

The first issue relates to the cross sectional dimension: Table A-2 shows the incidence of null flows over the full set of potential country pairs in the dataset, at a yearly break down. The estimation via OLS is therefore excluded by the zero-inflated structure of the full dataset, that would distort the estimates downward (see for instance Head and Mayer, 2014, for a thorough discussion on the choice of the correct estimator for gravity analysis in the context of zero-inflation). To the best of our knowledge, the incidence of null flows in the full dataset is larger than any other previous study: nonetheless, in the empirical section the sample is reduced by the estimation routines to those observations for which the FDI flow is non-zero in at least one year out of 12. This refinement substantially reduces the amount of zeroes to slightly less than 70%, allowing us to obtain consistent estimates via PPML (See Santos Silva and Tenreyro, 2011, for a comprehensive proof of the consistency of PPML estimator in presence of both over-dispersion of the data and over-inflation of null values in the dependent variable.).

they tend to mirror goods prevalently traded by advanced economies, leaving apart those goods that may have a diluted cultural content.

[30] UNCTAD itself bases part of its investments' reports on *fDIMarket*'s figures. Not only, the database constitutes one of the sources of the UNCTADSTAT's FDI dedicated section.

Table A-2: Percentage of "Zeroes" by Year

| Year | Null | Non-Null | Total | Incidence |
|---|---|---|---|---|
| 2003 | 32,453 | 1,587 | 34,040 | 95.34% |
| 2004 | 32,442 | 1,598 | 34,040 | 95.31% |
| 2005 | 32,405 | 1,635 | 34,040 | 95.20% |
| 2006 | 32,289 | 1,751 | 34,040 | 94.86% |
| 2007 | 32,151 | 1,889 | 34,040 | 94.45% |
| 2008 | 31,751 | 2,289 | 34,040 | 93.28% |
| 2009 | 31,960 | 2,080 | 34,040 | 93.89% |
| 2010 | 31,931 | 2,109 | 34,040 | 93.80% |
| 2011 | 31,833 | 2,207 | 34,040 | 93.52% |
| 2012 | 31,916 | 2,124 | 34,040 | 93.76% |
| 2013 | 31,756 | 2,284 | 34,040 | 93.29% |
| 2014 | 31,901 | 2,139 | 34,040 | 93.72% |
| Total | 384,788 | 23,692 | 408,480 | 94.20% |

Notes: This table breaks down the incidence of null flows by year. It becomes apparent that the issue of null flows is pervasive in the FDIMarket dataset as we constructed it. The high incidence of zeroes and the data over-dispersion in the sample prevent us from using OLS. We resort to use a PPML estimation technique as suggested by citetSilvaTenreyro, and raised to workhorse strategy by authors (see for instance Head and Mayer (2014), Yotov et al. (2016) among the others.

The second issue concerns the reliability of the "quantitative" information available, namely the *Capital Expenditure* (CAPEX) and *Job Creation* entries. Section 3.1 provides a theoretical justification for the use of count instead of the value of FDI flows as dependent variable: nonetheless, being able to test the theoretical prediction about the role of asymmetry in CP would call for a comparison across different measures of bilateral FDI. *fDIMarket* database is one of the few existing datasets that could potentially allow for this issue. Nonetheless, such an exercise calls for additional prudence: as stressed by both Desbordes and Wei (2017) and Lee and Ries (2016), *fDIMarket* collects information on all existing greenfield FDI *projects* as they are officially released by the respective investing companies. Unfortunately, in most of the cases no communication is made about the true CAPEX value. In all those cases, CAPEX is imputed according to an algorithm summarily described on the *DIMarket*'s website. Such imputation is likely to introduce non-trivial distortions in the data, the more relevant (a) the higher is the percentage of estimated projects over the total number of projects in a given bilateral corridor; (b) the lower the number of projects from the country of origin. Table A-3 below provides the tabulation of the projects for which only the *imputed* CAPEX was available, broken down by year. Given the incidence of estimated observations, we suggest a particular care when handling estimates obtained using value related dimensions as dependent variables, though the picture they provide may be particularly interesting. In Section 4.2.2, those results are presented and commented in light of our measure of CP.

Table A-3: Percentage of Imputed Values by Year

| Year | Imputed | Real Value | Observations | Incidence |
|---|---|---|---|---|
| 2003 | 6,325 | 3,182 | 9,507 | 67% |
| 2004 | 7,270 | 3,143 | 10,413 | 70% |
| 2005 | 7,849 | 2,883 | 10,732 | 73% |
| 2006 | 9,534 | 3,301 | 12,835 | 74% |
| 2007 | 8,968 | 4,006 | 12,974 | 69% |
| 2008 | 13,416 | 3,794 | 17,210 | 78% |
| 2009 | 12,063 | 2,723 | 14,786 | 82% |
| 2010 | 12,843 | 2,629 | 15,472 | 83% |
| 2011 | 14,101 | 2,757 | 16,858 | 84% |
| 2012 | 13,088 | 2,181 | 15,269 | 86% |
| 2013 | 14,319 | 2,399 | 16,718 | 86% |
| 2014 | 13,044 | 2,344 | 15,388 | 85% |
| Total | 132,820 | 35,342 | 168,162 | 79% |

Notes: The table report the percentage of estimated capital investment. The number of observations refers to the number of single projects collected by FDIMarket for the period 2003-2014. The large incidence of estimated values makes the estimates obtained using values as dependent variables not fully reliable: as a matter of facts, in addition to the lack of clarity in the imputation technique, imputation brings in a component of uncertainty per se.

# B Cultural trade as a proxy of the symmetric component of CP

Building upon Disdier et al. (2010), we identified the exchange of cultural goods as classified by UNCTAD (2010) as a good proxy of CP. In this Appendix we show how trade in cultural goods strongly relates to the symmetric component of CP as defined in Section 2. In other words, we provide a rough indication of the dependency of cultural attractiveness on cultural similarities. To that end we regress cultural trade on various conventional symmetric (and time invariant) proxies for cultural distance such as a dummy for *common border* $\text{contig}_{ni}$, the *log of weighted distance* $\ln \text{dist}_{ni}$, a measure of *religious proximity* $\text{relig}_{ni}$, a dummy $\text{rta}_{ni}$, which takes the value of 1 if both countries belong to a *regional trade agreement*, 0 otherwise, a binary variables for *common legal origin* $\text{comleg}_{ni}$, and finally a binary variable for *past colonial relationship* $\text{colony}_{ni}$ which takes the value of 1 if the two countries have ever been in a colonial relationship, 0 otherwise. All these variables are sourced from CEPII databases. Among the covariates the regression also includes a time varying component ($\ln\text{mig}_{ni,t}$), namely the *stock of bilateral immigrants* resident in the exporting country (Source: World Bank). Because data are available every 10 years (with the notable exception of the year 2013), our empirical exercise is a Pooled regression for the years 2010 and 2013 only, which nonetheless guarantees a still reasonably high number of observations. Furthermore, as in Felbermayr and Toubal (2010) we enrich the number of proxies by adding more refined measures of linguistic proximity obtained from Melitz and Toubal (2014): along with the standard dummy that equals 1 if a two countries share the same official language and 0 otherwise ($\text{COL}_{ni}$ *"common official language"*), we include $\text{CSL}_{ni}$ *"common spoken language"* as the probability that a pair of people at random from two countries understands one another in some language and $\text{CNL}_{ni}$ *"common native language"* as the probability that a random pair from two countries speak the same native language. Lastly, we employ a comprehensive measure of cultural distance widely used in the literature, namely the *Hofstede Index* $\text{Hofstede}_{ni}$ (Hofstede, 1991). This composite Index has been one of the main workhorses for the empirical of test the impact of cultural proximity on economic exchanges such as trade and FDI (see for instance Du et al. (2012)), but other than being at the same time pre-determined and symmetric, has the drawback of covering a fairly limited sample (see for a discussion Shenkar (2001)). The data are from Belot and Ederveen (2012).[31] Results for this exercise are reported in Table B-1.

The estimates in Table B-1 indicate that trade of cultural goods relate to almost all the proxies of CP we included, whose impacts have the expected sign. The first column reports the OLS results with log of imports of cultural goods as dependent variable. The coefficients are all statistically significant with the exception of $\text{CNL}_{ni}$: this is likely to be imputed to an high degree of collinearity between linguistic distance measures. The loss of information of zero bilateral trade due to the logarithmic specification could be a serious concern in our case, as the zeros in trade of cultural goods stand for a large share of the total available information. The main issue with the elimination of the zeros is a possible selection bias. Indeed, it might be that proxies for cultural proximity are associated with the intensity of trade in cultural goods only in the instances of positive trade and have no role in explaining the cases of the zeros. To address this issue we report the PPML results in Column 2. Despite the change in the sample size, almost all the effects retain the expected sign. The only exceptions are the measures of language proximity and the RTA dummy that, in any case, maintain a pairwise correlation coefficient that is positive and statistically different from zero (see Table B-2 below). Lastly, the inclusion of the Hofstede Index in the third column causes a considerable loss of information as the sample reduces to 19 OECD countries. The Index seems to be capturing most of the effect of religious and linguistic proximity and - most importantly for our purposes - is negatively related to the imports of cultural goods.

[31] See Belot and Ederveen (2012) for the details related to the construction of the Hofstede Index. See Section A for a more thorough description of the data and the complete list of sources and data accessibility.

Table B-1: Testing the Validity of Cultural Trade as a Proxy of CP

| Dep. Var. | $\ln \text{CultIMP}_{ni,t}$ | $\ln \text{CultIMP}_{ni,t}$ | $\ln \text{CultIMP}_{ni,t}$ |
|---|---|---|---|
| | (1) | (2) | (3) |
| $\ln \text{mig}_{ni,t}$ | 0.115*** | 0.0761*** | 0.0880** |
| | (20.83) | (4.30) | (2.89) |
| $\ln \text{dist}_{ni}$ | -1.225*** | -0.695*** | -0.921*** |
| | (-49.15) | (-10.61) | (-6.77) |
| $\text{contig}_{ni}$ | 0.317*** | 0.260** | 0.440* |
| | (3.74) | (2.86) | (2.34) |
| $\text{FTA}_{ni,t}$ | 0.266*** | 0.0807 | 0.683** |
| | (6.24) | (0.77) | (2.96) |
| $\text{comrelig}_{ni}$ | 0.236*** | 0.440* | 0.235 |
| | (3.55) | (2.28) | (1.26) |
| $\text{comleg}_{ni}$ | 0.281*** | 0.303*** | 0.411** |
| | (8.66) | (4.43) | (2.68) |
| $\text{colony}_{ni}$ | 0.500*** | 0.383*** | 0.763*** |
| | (5.67) | (3.65) | (3.45) |
| $\text{COL}_{ni}$ | 0.374*** | 0.0786 | -0.0000199 |
| | (6.13) | (0.55) | (-0.00) |
| $\text{CSL}_{ni}$ | 0.683*** | -0.350 | -0.394 |
| | (6.52) | (-1.45) | (-0.74) |
| $\text{CNL}_{ni}$ | 0.0691 | 0.209 | -0.402 |
| | (0.48) | (0.71) | (-0.92) |
| $\text{Hofstede}_{ni}$ | | | -1.034*** |
| | | | (-4.01) |
| Imp×Year FE | √ | √ | √ |
| Exp×year FE | √ | √ | √ |
| Sample | Full | Full | Reduced |
| Obs | 24620 | 54525 | 684 |
| % Zeros | - | 0.5485 | - |
| $R^2$ | 0.7476 | 0.8993 | 0.9118 |
| Estimator | OLS | PPML | OLS |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. t (z) -statistics in parentheses. Standard errors are clustered by trading-pair. The model includes importer×time and exporter×time FEs. The first and third columns' estimates are estimated with OLS. The sample size in this table reflect the way the different estimators deal with null flows as well as the sample size. The information which belong to groups with all zeros or missing values are automatically dropped by the estimator as FEs cannot be computed. The sample in the third column is reduced due to those countries for which the Hofstede Index of Cultural Proximity is available (see Belot and Ederveen, 2012).

Table B-2: Testing Validity of Cultural Trade as a Proxy of CP - Correlations

| Correlation with: | cult.trade $T_{ni,t}$ | |
|---|---|---|
| | Baseline Covariates Set | Linguistic and CP proxies |
| | (1) | (2) |
| $\ln \text{mig}_{ni}$ | 0.0955* | 0.0955* |
| | (0.0000) | (0.0000) |
| $\ln \text{dist}_{ni}$ | -0.0218* | -0.0218* |
| | (0.0000) | (0.0000) |
| $\text{contig}_{ni}$ | 0.0771* | 0.0771* |
| | (0.0000) | (0.0000) |
| $\text{FTA}_{ni,t}$ | 0.0363* | 0.0363* |
| | (0.0000) | (0.0000) |
| $\text{comrelig}_{ni}$ | -0.0049 | -0.0049 |
| | (0.2433) | (0.2433) |
| $\text{comleg}_{ni}$ | -0.0037 | -0.0037 |
| | (0.3691) | (0.3691) |
| $\text{colony}_{ni}$ | 0.0265* | 0.0265* |
| | (0.0000) | (0.0000) |
| $\text{lang}_{ni}$ | 0.0130* | |
| | (0.0018) | |
| $\text{COL}_{ni}$ | | 0.0101* |
| | | (0.015) |
| $\text{CSL}_{ni}$ | | 0.0359* |
| | | (0.0000) |
| $\text{CNL}_{ni}$ | | 0.0275* |
| | | (0.0000) |
| $\text{Hofstede}_{ni}$ | | -0.2507* |
| | | (0.0000) |
| Obs | 57672 | 703 |

Notes: * $p < 0.01$. SE in parentheses are clustered by trading-pair. The table show pairwise correlation coefficients between trade in cultural goods and all standard coefficients of proximity. Coefficients in the first column refers to the whole sample for which all variables are available. This means that it is limited to just year 2010 and year 2013 because of bilateral stock of migrants availability. Coefficients in the second column refers instead to the reduced sample for which the Hofstede index is available.

# C Extensions to the detour on asymmetry

**Asymmetry in CP and export capacity** This Appendix investigates the correlation between the degree of asymmetry in CP and the relative cultural export capacity between trading partners. This is done by dividing the set of countries which appear in at least one pair for which a value of asymmetry is available into four classes, depending on the value of their exports of cultural good with respect to the 3 quartiles of the distribution of cultural exports. The first class consists of countries below the first quartile of cultural exports, the second class of those between the first and the second quartile, the third class of those between the second and third quartile, and finally the fourth class of those countries above the third quartile of the distribution. The set of country pairs are then partitioned according to all possible combinations of two elements with repetitions from the four classes defined above. One pair could be classified either as containing two first class countries (both at the bottom of the cultural export distribution), one first and one fourth class country (the former at the bottom and the latter at the top of the cultural trade distribution) and so on and so forth for all 10 possible combinations. Finally, the value of asymmetry is regressed on the ten dummies identifying the elements of this partition (First-First, Second-Second, ..., First-Second, ...), taking those pairs with two bottom cultural exporters (First-First) as the base group. Results are reported in Table C-1.

Looking at the first column of Table C-1, we notice that on average across all pairs including two bottom cultural exporters the value of asymmetry is equal to 2.078, below both the mean and median values of asymmetry, equal to 2.932 and 2.614 respectively. Less asymmetry appears to be present in the CP between countries with a similar but higher value of cultural exports, and also between a country in the fourth class (top cultural exporter) and one in the third (quasi-top cultural exporter). Higher levels of asymmetry in CP instead are expected among countries which are relatively more heterogeneous in terms on cultural export capacity. Higher asymmetry in bilateral CP is associated with wider heterogeneity in export capacity and, to a lesser extent, with average export capacity within the pair. These patters are generally confirmed when restricting the analysis to bilateral cultural relationships characterized by attractiveness premia with the same sign (both positive and negative) as well as with different sign. These results are presented in the second, third and fourth columns of Table C-1.

**Asymmetry across different samples** The motivation of this extension is to show how the width and degree of homogeneity within the sample of countries may be crucial when the impact of CP on the economy is investigated. We argue that the empirical assessment of the role of asymmetric CP for economic transactions needs to be conducted with the widest possible country coverage. A empirical analysis conducted on a narrow and homogeneous set of countries could potentially overestimate the degree of asymmetry embedded in cultural relationship and therefore undermine the assessment of the role of such asymmetric component in determining economic outcomes. In order to show this we replicate the construction of our empirical measure of asymmetry in CP starting from a sample with the same country coverage of the database used by Felbermayr and Toubal (2010) to construct their asymmetric measure of CP based on Eurovision Song Contest scores.[32] This subsample includes only European countries, that can be considered as a relatively homogeneous group under many respects, and especially when compared with the rest of the World. We denote by $|\hat{\gamma}_{ni}^{full} - \hat{\gamma}_{in}|^{FT}$ the resulting measure of asymmetry. $|\hat{\gamma}_{ni} - \hat{\gamma}_{in}|^{full}$ indicates instead the asymmetry whose components have been estimated on the whole sample. Table C-2 reports both measures of asymmetry and their difference for a number of selected country pairs. The + and − signs below the first two columns reflect the sign of the attractiveness premium exerted by country $i$ and country $n$ on each other. Take for instance the UK and France. The asymmetry computed from the whole sample is very low

[32] The country coverage is identical with the exception of Yugoslavia due to availability of cultural trade data.

Table C-1: Asymmetry Across Different Types of Cultural Traders

| Dep. Var. | Asymmetry $\lvert\hat{\gamma}_{ni} - \hat{\gamma}_{in}\rvert$ | | | |
|---|---|---|---|---|
| Attractiveness premia | All types | Both positive | Both negative | Opposite sign |
| | (1) | (2) | (3) | (4) |
| Second-Second | -0.400** | -0.279 | -0.0767 | -0.561** |
| | (-3.13) | (-1.35) | (-0.54) | (-2.75) |
| Third-Third | -0.610*** | -0.143 | -0.946*** | -0.399 |
| | (-4.90) | (-0.74) | (-5.45) | (-1.60) |
| Fourth-Fourth | -0.828*** | -0.172 | - | - |
| | (-5.59) | (-0.82) | - | - |
| First-Second | 1.048*** | 1.104*** | 0.299* | 1.532*** |
| | (6.77) | (3.96) | (2.10) | (7.12) |
| Second-Third | 0.188 | 0.00573 | 0.110 | -0.00420 |
| | (1.56) | (0.03) | (0.79) | (-0.02) |
| Third-Fourth | -0.586*** | 0.0328 | - | 0.973 |
| | (-4.75) | (0.17) | - | (1.18) |
| First-Third | 1.682*** | 1.380*** | 0.889*** | 1.721*** |
| | (12.21) | (4.79) | (6.50) | (9.06) |
| Second-Fourth | 0.779*** | 0.607** | 1.093 | 0.889*** |
| | (5.97) | (3.07) | (1.12) | (4.61) |
| First-Fourth | 2.690*** | 1.270*** | 1.651*** | 2.043*** |
| | (21.84) | (5.07) | (10.23) | (11.96) |
| Constant (First-First) | 2.078*** | 1.423*** | 1.392*** | 3.194*** |
| | (19.70) | (7.76) | (12.86) | (20.20) |
| Obs | 4137 | 1486 | 793 | 1858 |
| $R^2$ | 0.3424 | 0.1274 | 0.2285 | 0.2421 |
| Estimator | OLS | OLS | OLS | OLS |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z-statistics in parentheses. Standard errors are clustered by trading-pair. In this table the proxy for asymmetry ($|\hat{\gamma}_{ni} - \hat{\gamma}_{in}|$) is regressed on a constant and 9 dummies. As an illustration, the dummy "Fourth-Fourth" takes value one for those country pairs where both countries have a value of cultural exports above the third quartile of the distribution of cultural exports. As a further illustration the dummy "First-Fourth" takes value one for those country pairs where one country is a bottom exporter of cultural goods (below the first quartile of the cultural exports distribution) and the other is a top cultural exporter (above the third quartile). When point estimates and t statistics are not reported it is because the respective dummy coefficient has no variability (always equal to 0) in the corresponding estimation sample. The case in which both countries in the pair are bottom exporters (below the first quartile of the cultural exports distribution) is set as base level and the related dummy variable is omitted from the regression.

and equal to 0.17. The first + sign below the asymmetry score indicates that the attractiveness premium that France exerts on the UK with respect to the average country is positive. The same is true the other way round, as indicated by the second + sign. When computed on a smaller sample featuring only European countries, the value of asymmetry increases by more than 180% and becomes equal to 0.48 (still relatively small compared to the average asymmetry over the whole sample).

The last column of the table shows the extent of the bias induced by considering only a subsample of (relatively) homogenous countries: a negative sign in the difference between $|\hat{\gamma}_{ni} - \hat{\gamma}_{in}|^{full}$ and $|\hat{\gamma}_{ni} - \hat{\gamma}_{in}|^{FT}$ means that the degree of asymmetry in the country pair under consideration decreases when other, more heterogeneous countries are considered. Failing to consider the role of the rest of the world within the system of cultural affinity could result in a sever bias in cultural relationship between countries.

Beyond the few examples reported in Table C-2, Figure  gives a sense of the sign of the bias on

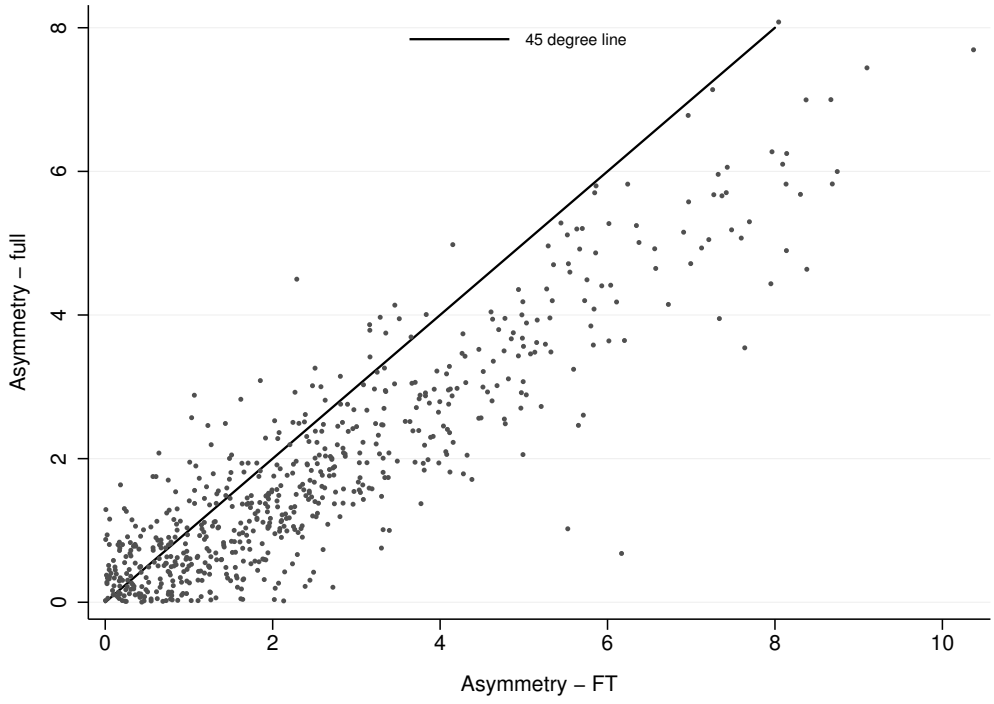
Table C-2: Asymmetry Across Different Samples

| Country $n$ | Country $i$ | Asymmetry - full $\lvert\hat{\gamma}_{ni} - \hat{\gamma}_{in}\rvert^{\text{full}}$ | Asymmetry - FT $\lvert\hat{\gamma}_{ni} - \hat{\gamma}_{in}\rvert^{\text{FT}}$ | Differential $\lvert\hat{\gamma}_{ni} - \hat{\gamma}_{in}\rvert^{\text{full}} - \lvert\hat{\gamma}_{ni} - \hat{\gamma}_{in}\rvert^{\text{FT}}$ |
|---|---|---|---|---|
| Finland | Italy | 1.16 | 2.35 | -1.19 |
| | | + + | + + | |
| United Kingdom | France | 0.17 | 0.48 | -0.31 |
| | | + + | + + | |
| Russian | United Kingdom | 0.95 | 1.60 | -0.65 |
| | | + + | + + | |
| Germany | Turkey | 0.33 | 1.46 | -1.13 |
| | | + + | + + | |
| Spain | Russian | 2.19 | 2.20 | -0.01 |
| | | + + | - + | |
| Norway | Sweden | 1.49 | 1.95 | -0.46 |
| | | + + | + + | |
| Croatia | Sweden | 0.31 | 1.89 | -1.58 |
| | | + + | + - | |
| Belgium | Malta | 2.88 | 5.02 | -2.14 |
| | | + + | + - | |
| Ireland | United Kingdom | 2.70 | 3.32 | -0.62 |
| | | + + | + + | |
| Ukraine | Ireland | 3.04 | 3.45 | -0.41 |
| | | + - | - - | |

Notes: The table lists a selection of country pairs and shows the extent of the bias in the empirical assessment of asymmetry due to adopting a sample of relatively homogeneous countries. A positive (negative) value of the differential across the full sample and the restricted one implies that the restriction is actually over-(under-) estimating the true extent of CP. The sample of countries used in Felbermayr and Toubal (2010), which only includes European countries is taken as the restricted set of relatively homogeneous countries. The + and − signs below the two columns of symmetry report the sign of the attractiveness premium exerted by country $i$ and country $n$ on each other.

all the country pairs generated from the restricted sample for which both measures of asymmetry are estimated. This is done by plotting, for each pair the value of asymmetry coming from the full sample (on the vertical axis) against the value of the asymmetry generated from the restricted sample (on the horizontal axis). With the bulk of the observations below the 45 degree line, especially moving away from the origin, we conclude that the overestimation of asymmetry in CP implied by an empirical framework with limited country coverage can be highly widespread.

Figure C-1: Asymmetry Full Sample VS Asymmetry Felbermayr and Toubal (2010) Sample



# D Further addressing the omitted variable bias

An important econometric issue in our regressions is the potential endogeneity of our proxy for CP which mainly arises because of the potential omission of unobserved factors that might be correlated both with the error term (and thus FDI) and with trade in cultural goods. Both the proposed IV analysis and the inclusion of dyadic FEs in Section add robustness to our estimates and confirm our main conclusions. Here we further test the consistency of our benchmark results by augmenting the specification with the inclusion of observable variables of dimension *nit* that might capture (part) of these unobserved time-varying dyadic factors.

A variable which potentially shapes both cultural trade as well as FDI is represented by the migrants' networks. Migrants are able to form important linkages between the country of origin and the one of destination. To this regard, the literature identified a positive impact of migrants' networks on both FDI and international trade (see for instance Javorcik et al., 2011; Gould, 1994; Giovannetti and Lanati, 2016), which is predominantly imputed to the "insider knowledge" that migrants provide to reduce information costs in international transactions. The time varying impact of migrants' networks on FDI cannot be entirely absorbed through our comprehensive set of fixed effects and its exclusion from the list of regressors may introduce an omitted variable bias.[33] The results are reported in Table D-1 below, that replicates the specifications of Table 4, but comprises the bilateral stocks of immigrants from both $n$ to $i$ and $i$ to $n$ as additional regressors.

Including the stocks of immigrants does not alter our overall conclusions. In particular, the positive impact of exports in cultural goods which proxy for the destination side mechanisms driving the firm's decision to invest is always statistically significant and does not vary in magnitude as we control for the network effect (column 1-3). In a nutshell, the destination side mechanisms driving FDI seem to be independent from the network channel. This points to the goodness of our proxy in capturing the role of cultural proximity and score in favor of its robustness to the inclusion of alternative measures of time varying CP.

Finally, in Table D-2 we augment our baseline specification with the total volume of bilateral non cultural trade. In particular, ln bil_trade_NC captures the effect of the sum of bilateral non cultural imports and exports between origin and destination at time $t$ on FDI. The evidence from Disdier et al. (2010) shows that bilateral flows of cultural products can be highly related to the overall flows of bilateral trade, while at the same time bilateral economic exchanges such as aggregate trade are likely to be positively associated with FDI. The statistics indicate that the volume of bilateral non-cultural trade does not impact FDI and its inclusion substantially leaves our results unaffected, which we find as reassuring.

[33] Their inclusion, however, reduces the explanatory power of our econometric exercise, as data on bilateral migrants' stocks with a global country coverage are generally only available with a 10 year interval between observation (Source: The World Bank). Therefore, we only include the migrants' stock as a robustness check.

Table D-1: Addressing Omitted Variable Bias: Including Migration

| Dep. Var. | Count $C_{ni,t}$ | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| $\ln$ migstock$_{ni,t}$ | 0.0810*** | | 0.0579** |
| | (5.13) | | (2.63) |
| $\ln$ migstock$_{in,t}$ | | 0.0788*** | 0.0293 |
| | | (4.29) | (1.33) |
| $\ln$ CultIMP$_{ni,t}$ | 0.0507** | 0.0368 | 0.0204 |
| | (3.27) | (1.90) | (0.93) |
| $\ln$ CultEXP$_{ni,t}$ | 0.290*** | 0.296*** | 0.290*** |
| | (15.12) | (12.94) | (11.37) |
| $\ln$ dist$_{ni}$ | -0.0566 | -0.0693 | -0.0574 |
| | (-1.25) | (-1.46) | (-1.13) |
| colony$_{ni}$ | 0.283*** | 0.308*** | 0.292*** |
| | (4.26) | (4.41) | (3.87) |
| lang$_{ni}$ | 0.117* | 0.0704 | 0.0725 |
| | (2.01) | (1.11) | (1.08) |
| comrelig$_{ni}$ | 0.930*** | 0.910*** | 0.960*** |
| | (7.48) | (7.04) | (6.82) |
| contig$_{ni}$ | -0.0391 | -0.0447 | -0.0140 |
| | (-0.55) | (-0.60) | (-0-18) |
| comleg$_{ni}$ | 0.156*** | 0.189*** | 0.187*** |
| | (3.45) | (3.84) | (3.61) |
| FTA$_{ni,t}$ | 0.129 | 0.144* | 0.138 |
| | (1.94) | (2.10) | (1.84) |
| BIT$_{ni,t}$ | 0.0277 | -0.0154 | -0.0315 |
| | (0.51) | (-0.26) | (-0.93) |
| Imp×Year FE | √ | √ | √ |
| Exp×year FE | √ | √ | √ |
| Obs | 9619 | 8756 | 5853 |
| % Zeros | 67% | 67% | 60% |
| $R^2$ | 0.91 | 0.92 | 0.92 |
| Estimator | PPML | PPML | PPML |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z-statistics in parentheses. Standard errors are clustered by trading-pair. The dependent variable "Count" $C_{ni,t}$ is the bilateral number of Greenfield FDI projects from country $i$ to country $n$. It includes the zero flows. This table replicates the baseline specification adding the bilateral stock of migrants from $n$ to $i$ as additional regressors. The reduced number of observations is due to the availability of the migration data, that allow to use only two points in time (2010 and 2013) for the period covered in the analysis (Source: The World Bank).

Table D-2: Addressing Omitted Variable Bias: Including Non-Cultural Trade

| Dep. Var. | Count $C_{ni,t}$ | |
|---|---|---|
| | (1) | (2) |
| $\ln \text{CultIMP}_{ni,t}$ | 0.0690*** | 0.0838*** |
| | (5.90) | (6.01) |
| $\ln \text{CultEXP}_{ni,t}$ | 0.305*** | 0.324*** |
| | (21.91) | (14.64) |
| $\ln \text{bil\_trade\_NC}_{ni,t}$ | | -0.0352 |
| | | (-1.24) |
| $\ln \text{dist}_{ni}$ | -0.179*** | -0.176*** |
| | (-5.13) | (-5.08) |
| $\text{colony}_{ni}$ | 0.366*** | 0.367*** |
| | (6.85) | (6.90) |
| $\text{lang}_{ni}$ | 0.181*** | 0.176*** |
| | (3.53) | (3.50) |
| $\text{comrelig}_{ni}$ | 0.883*** | 0.876*** |
| | (9.21) | (9.21) |
| $\text{contig}_{ni}$ | -0.0977 | -0.0947 |
| | (-1.61) | (-1.56) |
| $\text{comleg}_{ni}$ | 0.153*** | 0.154*** |
| | (4.06) | (4.08) |
| $\text{FTA}_{ni,t}$ | 0.118* | 0.117* |
| | (2.19) | (2.17) |
| $\text{BIT}_{ni,t}$ | 0.0115 | 0.00749 |
| | (0.29) | (0.19) |
| Imp×Year FE | √ | √ |
| Exp×year FE | √ | √ |
| Obs | 87448 | 87448 |
| % Zeros | 0.749 | 0.749 |
| $R^2$ | 0.9221 | 0.9221 |
| Estimator | PPML | PPML |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z-statistics in parentheses. Standard errors are clustered by trading-pair. The dependent variable "Count" $C_{ni,t}$ is the bilateral number of Greenfield FDI projects from country $i$ to country $n$. It includes the zero flows. The estimates in column (1) replicates column (3) in our baseline results in Table 4; column (2) provides the result of the same equation, augmented to include total bilateral trade of non-cultural goods.

# E Relevance of the instruments

Table E-1 below mimics a first stage regression for the IV analysis, by showing the relevance of the instruments in explaining the endogenous variables to our analysis. Since the *IVPPML* command does not compute first stage regression, we regressed the endogenous variables on all the instruments as well as on the covariates of the second stage.

Table E-1: Relevance of the Instrument: First Stage Endogenous Variables on Instruments

| Dep. Var. | $\text{Cult.Import}_{ni,t}$ | $\text{Cult.Export}_{ni,t}$ |
|---|---|---|
| | (1) | (2) |
| $\ln \text{CultIMP}_{ni,t-8}$ | 0.560*** | |
| | (14.73) | |
| $\ln \text{CultEXP}_{ni,8}$ | | 0.560*** |
| | | (14.74) |
| $\ln \text{dist}_{ni}$ | -0.664*** | -0.663*** |
| | (-9.15) | (-9.14) |
| $\text{colony}_{ni}$ | -0.116 | -0.116 |
| | (-1.37) | (-1.37) |
| $\text{lang}_{ni}$ | 0.123 | 0.124 |
| | (0.90) | (0.91) |
| $\text{comrelig}_{ni}$ | 0.0534 | 0.0539 |
| | (0.44) | (0.44) |
| $\text{contig}_{ni}$ | 0.0773 | 0.0776 |
| | (1.13) | (1.14) |
| $\text{comleg}_{ni}$ | 0.0481 | 0.0479 |
| | (0.78) | (0.78) |
| $\text{FTA}_{ni,t}$ | 0.324** | 0.325** |
| | (2.94) | (2.95) |
| $\text{BIT}_{ni,t}$ | 0.0485 | 0.0484 |
| | (0.59) | (0.58) |
| Imp×Year FE | √ | √ |
| Exp×Year FE | √ | √ |
| Obs | 11117 | 11117 |
| % Zeros | 12.2% | 12.2% |
| $R^2$ | 0.9502 | 0.9502 |
| Estimator | PPML | PPML |

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. z-statistics in parentheses. Standard errors are clustered by trading-pair. This table shows the relevance of the selected instruments on the endogenous variables. The decision to adopt lagged values of the endogenous variables builds on Card (2001).
The estimates are obtained with PPML using the *PPML* command by Santos Silva and Tenreyro (2006) and Santos Silva and Tenreyro (2011) which perfectly deals with the reduced set of FE we are going to include in the instrumental analysis. Column (1) shows the correlation of the lagged value of import in cultural goods on current imports. Column (2) performs the same exercise on export. The sample is reduced in this specification, because of data availability for the lagged instruments. Time coverage: 2007-2014

# F Country excluded from the dataset

Table F-1: List of Countries Excluded from the Analysis

| In both direction: no flows of greenfield FDI (in or out) over the entire period |
|---|
| Anguilla, Netherland Antilles, Cocos and Keeling Islands, Cook Islands, Christmas Islands, Western Sahara, Falkland Islands, Faeroe Islands, Gibraltar, French Guiana, Kiribati, Marshall Islands, Northern Mariana Islands, Montserrat, Norfolk Islands, Niue, Nauru, Pitcairn, Palau, Saint Helena and Tristan da Cunha, San Marino, Saint Pierre et Miquelon, Tokelau, Tonga, Tuvalu, British Virgin Islands, Vanuatu, Wallis and Futuna |
| No outward flows over the whole period (excluded as source countries) |
| Aruba, Benin, Bhutan, Cape Verde, Central African Republic, Chad, Comoros, Republic of the Congo, Dominica, Eritrea, Grenada, Guinea, Guinea-Bissau, PRD Korea, Liberia, Maldives, Mauritania, New Caledonia, Niger, Paraguay, Sao Tome and Principe, Seychelles, Sierra Leone, Somalia, Saint Kitts and Nevis, Sain Lucia, Saint Vincent and the Grenadines, Suriname, Timor Leste, Turkmenistan |
| Countries excluded or aggregated for inconsistencies between CEPII and fDIMarket |
| Serbia and Montenegro (both excluded)<br>Belgium and Luxembourg (both excluded)<br>Sudan and Sud Sudan (South Sudan is Excluded)<br>Switzerland and Liechtenstein (Aggregated)<br>France and Monaco (Aggregated) |

Notes: The result of the exclusion of these countries is a rectangular dataset of $n \times m$ countries. In addition to these countries - excluded for data inconsistencies - other dyadic flows are excluded when no investment occurs between two countries during the period analyzed. This explains the discrepancy between the size of the dataset and the number of observations used in the estimation