XXXVII CONFERENZA ITALIANA DI SCIENZE REGIONALI

MIGRATIONS, DIASPORAS AND THEIR IMPACT ON KNOWLEDGE FLOWS IN THE EU REGIONS

Anna D'Ambrosio[1], Sandro Montresor[2], Mario Davide Parrilli[3], Francesco Quatraro[4]

Draft 15.07.2016

SUMMARY

We study the flows of knowledge between Spanish provinces and foreign countries by applying a gravity model to the analysis of patent coinventorships and patent citations (both inward and outward) using REGPAT data on the period 1998-2011. We show that both immigration and emigration stocks positively affect knowledge flows. However, the effect is positive and significant for immigrants in the before 2006 period and diminishes in the more recent years when the inflows of immigrants has been decreasing. The effect is instead positive and significant for emigrants in the after-2006 period, when the Spanish diaspora has boomed. For both immigrants and emigrants, the effect is stronger on more tacit kinds of knowledge flows, i.e. coinventorships, with respect to more codified knowledge flows as measured by citations. We also show that both human capital and social capital activate the information effect of migration stocks; the effect is strongest for non-Spanish speaking countries.

[1]University of Turin, Department of Economics and Statistics "Cognetti de Martiis", Lungo Dora Siena 100A, 10153 Turin and BRICK, Collegio Carlo Alberto, email: anna.dambrosio@unito.it (corresponding author)
[2]Kore University of Enna, Faculty of Economics and Law. Via Salvatore Mazza, 12, 94100 Enna. Email: sandro.montresor@unikore.it.
[3]University of Bournemouth, Fern Barrow, Poole BH12 5BB, Regno Unito, email: dparrilli@bournemouth.ac.uk
[4]University of Turin, Department of Economics and Statistics "Cognetti de Martiis", Lungo Dora Siena 100A, 10153 Turin and BRICK, Collegio Carlo Alberto, email: francesco.quatraro@unito.it.

1. Introduction

Diasporas and immigration flows are changing the workforce and the industrial base of regions. The rich literature on the pro-trade effect of migration has corroborated the hypothesis that mobility of people is associated with information and knowledge transfer (e.g. Gould, 1994; Rauch and Trinidade, 2002; Wagner et al., 2002; Peri and Requena-Silvente, 2010; Felbermayr et al, 2014). While this literature is strongly supporting the argument, the proof of this has always been indirect. In this paper, we seek to provide a direct measurement of the impact that diasporas (and migrations more in general) have on knowledge flows between the host regions and the origin countries. Using European Patent Office data and proxying knowledge flows with patent co-inventorships and citations between Spanish NUTS3 Regions and EU countries, we study whether immigrant and emigrant communities promote knowledge flows within each pair.

A recent branch of the literature has analyzed the patterns of knowledge flows and diffusion with the help of gravity models applied to the analysis of patent citations, coinventorship and applicant-inventor ties, highlighting the role of geographic distance, cultural and language barriers in affecting the intensity of knowledge flows (see for instance Maurseth and Verspagen, 2002, Pacci and Usai, 2009; Montobbio and Sterzi, 2013). The main argument for applying a gravity model to the analysis of transnational knowledge flows is that a component of the knowledge needed for innovation is tacit and localized, which hampers the diffusion of knowledge across regions; hence, we can expect geographical distance – as well as any other kind of distance – to hinder the transfer of such tacit knowledge (Montobbio and Sterzi, 2013). The distinction between co-inventorships and citations allows us testing this hypothesis: in analogy with previous literature (e.g. Quatraro and Usai, 2016; Montobbio and Sterzi 2013) we assume that patent co-inventorship proxies a more tacit kind of knowledge flow than patent citations: as argumented by Montobbio and Sterzi, 2013, face-to-face interactions can be considered to be as a "superior vehicle of knowledge communication because it is possible to have instantaneous feedback and direct correction of wrong interpretation (...)" (p.7). Obviously, this kind of face-to-face interaction is more likely to occur by co-inventorship than by citation. Hence, the effect of distance should be greater for co-inventorships than for patent citations. More generally, however, because information is not perfectly spread, there may well be barriers to the knowledge that inventors have of the existing inventions and studies of relevance to their work, even in the case that this information is entirely codified. Hence, we expect distance – proxying for information costs – to play a role also in the case of codified knowledge flows.

On the other hand, we can expect that there are "mass" factors accelerating the flows of knowledge between regions: greater absorptive capacity and a larger pool of qualified workers can be expected to act as "attractors" of knowledge flows (Montobbio and Sterzi, 2013: p.14).

Besides knowledge mass and distance, in analogy with the literature on gravity models in international trade (chiefly Anderson and Van Wincoop, 2003), we can think that bilateral knowledge flows are also affected by a "multilateral resistance term", i.e. the average level of barriers to knowledge flows that a province faces with all other partners. The basic implication would be that changes in the bilateral "costs" of exchanging knowledge would affect the observed flows of knowledge only to the extent that the bilateral barriers are affected more strongly than the average barriers towards all other countries. Just to make an example, consider the case of a province that is already well connected with a variety of inventors worldwide. Suppose an exogenous change in communication costs occurs. The flows of knowledge between the province and a given country will only be affected to the extent that the bilateral change in communication costs is greater than the average.

In this paper, we take a regional focus and propose that communities on the move, who carry with them both codified and tacit knowledge, may facilitate the transfer of both types of knowledge.

If distance hinders the exchange of information, immigrants and emigrants, by bridging such distance, may overcome part of these barriers to knowledge flows and effectively increase the knowledge flows incoming or outgoing the region. The argument resonates with the classical argument about the effects of immigration on the barriers to trade (cfr. Gould, 1994; Rauch and Trinidade,2002; Felbermayr et al, 2014).

The gravity model provides us with a framework to study the opportunities for knowledge flows to arise: greater human capital and absorptive capacity in the origin and destination areas are likely to increase the probability that there is an opportunity to exchange knowledge. Moreover, in analogy with Wagner et al. (2002) random encounter model, we can think of factors that facilitate the ability of immigrants to effectively promote the knowledge exchange. First and foremost, this is likely to increase in the human capital of the communities on the move and in the similarity of language between the origin and the destination country (e.g. Wagner et al. 2002). Another aspect that we expect to contribute to the ability of the diasporas to promote knowledge exchanges is the social capital of the recipient community. If it is easier for foreign workers to integrate in the host economic environment, we can expect that it will also be easier to *ceteris paribus* act as "boundary brokers" (Parrilli, 2012; Williams, 2007) and facilitate knowledge flows with an economic relevance. We expect that the more dense, cohesive and open the host community, the greater the effect of the immigrant community on the facilitation of knowledge flows. A similar argument applies to foreign workers: ceteris paribus, a more open and integrated foreign community is expected to more effectively promote knowledge flows. We are able to include proxies for the social capital of the host provinces thanks to the availability of yearly indices of social capital at the NUTS3 regions level computed by the Ivie (see below). Instead, we lack a systematic province-level measure of social capital in the migrating communities, hence we rely on their size as a proxy for their density, similarly to what has been argued in previous works on the migration-trade link (e.g Rauch and Trinidade, 2002: the larger the immigrant community, the more likely we expect it to be that it is densely interconnected both internally and with the host society).

2. Data

Our database originates from the intersection of different data sources. The data on the stocks of immigrant residents originate from the residents' register, the Padron Municipal, and are made publicly available by the Spanish Statistical Institute, the Instituto Nacional de Estadistica. The records include data on residents with foreign nationality. Hence, our immigration measure excludes all foreign residents who have acquired the Spanish nationality as well as, obviously, any irregularly residing foreigners. The data on emigrants are drawn from the Censo Electoral de Residentes Ausentes (CERA), i.e. the electoral registers of Spanish nationals who reside abroad but maintain their voting rights in Spain. The structure of these data forces the unit of analysis of the knowledge flows to be the province-country pair over the time period 1998-2011 (as regards immigrants) and 2006-2011 (as regards emigrants). Over the 1998-2011 period, immigration stocks in Spain have been booming, with an average growth rate of 17,9%, while emigration stocks have been growing an average rate of 4.17% (Eurostat). However, the 2008-2009 crisis period marked a drop in the growth rates of immigrant stocks which have been stagnating afterwards; on the contrary, emigrant stocks have been growing faster since the crisis years on. The extremely high levels of unemployment associated with the crisis in Spain are probably responsible for these changes.

The data on citation and co-inventorships originate from the OECD REGPAT and the OECD Citations database, which cover patent applications filed to the European Patent Office (EPO) and patent patent applications filed under the Patent Cooperation Treaty (PCT). The OECD Citations database provides information on patent citations found in patent applications filed directly to the EPO or via the PCT.

In order to construct our co-inventorship measure, we selected all patent applications where at least one inventor resides in Spain and one abroad. Based on the inventors' addresses, we

constructed a dyadic co-inventorship measure which counts the number of coinventorship ties between each province (NUTS3 region) and each country in each year between 1998 and 2011. In case of multiple inventors, a proportional share is assigned to each region and, as a result, cells are not going to be made of integers. This leads to a total of 3,650 dyads: 50 Spanish provinces (excluding Ceuta and Melilla) with 73 countries. This implies that when we are able to exploit the full length of our dataset (13 years), the number of observations amounts to 47,450. We consider the priority year of the patent application as the year of the coinventorship, meant to approximate the date to the invention.

As regards the citations, we distinguish between the *inward* citations, i.e. the cases where a Spanish patent is cited by a foreign patent, from the *outward* citations, i.e. the cases where a Spanish patent application cites a foreign patent. In each case, we aggregate citations by Spanish NUTS3 region and partner country for every year, yielding to a similar dyadic variable as the co-inventorship one. Patent citation data are limited to 26 EU countries, which restricts the numerosity of our dataset to a maximum of 16,900 observations (50 provinces * 26 countries over 13 years). Table 1 reports the countries included in the analysis.

<<<insert Table 1 here>>>

To test our arguments about the role of immigration and emigration in relation with absorptive capacity, we include province-level measures of per-capita GDP (INE data), stocks of inventors and of patent applications (REGPAT data, derived by aggregating the number of inventors and of patent applications by NUTS 3 region).

To provide a first test about our arguments on human capital of immigrants, we selected the data relevant for Spain from the OECD-DIOC database, which provides aggregate (and cross-sectional, referring exclusively to 2007) data on the immigrants' qualification by country of destination. While the measure is obviously imperfect, it allows us constructing two broad categories of immigrants: the highly qualified and the low-qualified. We define as “highly qualified” those immigrants coming from countries for which the ratio between highly qualified workers and low qualified workers (ISCED codes 3 to 6 over ISCED codes 0-2) exceeds the median. Notice that the average for Spain is well below the median. This binary variable can be interacted with the immigration stocks to allow for heterogeneity in the information effects by qualification[5]. Unfortunately, the corresponding measure of the qualifications of emigrants is too incomplete to be useful.

To test our arguments about social capital at the province level, we cannot resort to the relatively standard “generally speaking questions” on trust nor on the measures of associativity used in the European Value Survey nor in the World Value Survey (Cortinovis et al., 2016; Granato et al., 1996a; Knack and Keefer, 1997; Zak and Knack, 2001), which are not available at such a level of disaggregation nor on a yearly basis. Hence, we draw our province level social capital measures from the Ivie study[6] (Fernandez et al 2015; Pérez et al., 2005; Tortosa-Ausina and Peiro, 2012). In this study, social capital is modelled as an asset in which people invest if it is profitable, and as a function of a set of variables, including the effort of cooperation in terms of working hours, the duration of social capital stocks, the time horizon of the investment in social capital, the discount rate, the degree of connectivity of social networks, the inequality in society and the variation in the investment in social capital of the other individuals given a change in the social capital investment in one of them (see Fernandez et al 2015 for details). The resulting estimate of social capital volume provides a measure of the within-province variation of social capital over time, which suits the structure of our data.

Table 2 reports the summary statistics of our variables, and table 3 the correlation matrix.

[5] The information about the average immigrants' qualification is not available for Cyprus, Estonia, Kazakstan, Latvia, Madagascar, Malta, Malaysia, Netherlands, Serbia, Uzbekistan in the version at our current disposal, which reduces the observations to a maximum of 42,700 when the high-qualification variable is included.

[6] Instituto Valenciano de Investigaciones Económicas (www.ivie.es), in collaboration with the Fundacion Banco Bilbao Vizcaya (FBBVA, www.fbbva.es)

<<<insert Table 2 here>>>

<<<insert Table 3 here>>>

3. Model and Variables

Our empirical model is a gravity model of knowledge flows (Maurseth and Verspagen, 2002; Paci and Usai, 2009; Picci, 2010; Maggioni et al, 2011; Montobbio and Sterzi, 2013; Cappelli and Montobbio, 2013) which we use to predict each of the two proxies of knowledge flows, i.e. patent citations and co-inventorships. As we mentioned, the unit of analysis is the dyad is constituted by each Spanish province *i* (excluding Ceuta and Melilla) and each partner country *j* in each year *t*. In the case of co-inventorships, the set of partner countries corresponds to the one reported in Table 1; in the case of patent citations, the set of partner countries is limited by data availability to EU countries. In analogy with the literature on the pro-trade effect of immigrants, we augment the gravity model with the stocks of immigrants and/or emigrants meant to decrease the bilateral costs of transferring knowledge. More specifically, we apply a gravity model similar to the one used in Bratti et al. (2014) to analyze the trade of Italian provinces with the rest of the world. With respect to their model, however, we include a set of variables that refer more strictly to knowledge creation and exchange:

$$ln(1+Flow_{jit}) = \varphi_{rt}+\psi_{jt}+\chi_{rj}+\beta_1 * ln(1+Migra_{jit-1})+\beta_2 * ln(GDP_{it-1})+\beta_3 * ln(1+PatentStock_{it-1})+\beta_4 * lnDist_{ij} + \varepsilon_{jit} \quad [1]$$

Our dependent variable is the knowledge flow of the province-country pair for each year of the 1998-2011 period. $lnFlow_{ijt}$ corresponds to each of our measures of knowledge flows: the log of patent co-inventorships $lncoinv_{ijt}$; the log of citations from country j to province *i* in year *t,* $lncited_{ijt}$; and the log of citations from province *i* to region *j* in year *t,* $lnciting_{ijt}$.

The term $ln(1+Migra_{ijt-1})$ is included to account for our variables for interest, i.e. the log stocks of immigrants from country *j* to province *i* in year *t*, $ln(1+Immi_{ijt-1})$ and the log stocks of emigrants from province *i* to country *j*, $ln(1+Emi_{ijt-1})$. We run different specifications of our model. In a first set of specifications, we focus exclusively on immigrants in order to fully exploit the relatively long panel which we have at our disposal (13 years, 1998-2011). In a second set of specifications, we only look at emigrants, which shrinks the length of the panel to 5 years, 2006-2011. In a third set of specifications, we include immigrants and emigrants stocks jointly. This allows comparing the effects of the two migration types over 2006-2011.

As it is standard in gravity models with panel data (e.g. Head and Mayer, 2014; Debenedictis and Taglioni, 2011; Baldwin and Taglioni, 2007), we use time-varying country effects to proxy for the knowledge "mass" of the partner country. To proxy for the mass of knowledge available in each considered province, we include time-varying effects at the NUTS2 region *r* level (because of model saturation, and similarly to Bratti et al. 2014), along with the log of province GDP $lnGDP_{it-1}$and the log stock of patent applications in the province $ln(1+PatentStock_{it-1})$ (which we substitute in some specifications with the log stock of inventors $ln(1+InventorStock_{it-1})$. These are meant to proxy for the human capital and local absorptive capacity of the provinces. To proxy for time-invariant bilateral costs of sharing knowledge, we add a bilateral distance measure (calculated as great-circle distance between each province and the capital city of each partner country) and a bilateral fixed effect at the country *j*-NUTS2 region *r* level.

In order to test our hypotheses about the role of human capital, we interact each of our migration variables with an indicator variable *Spa* equal to 1 if the partner country is Spanish-speaking (either as an official language or because Spanish is currently spoken) and 0 otherwise. Spain being the only Spanish-speaking EU country, we cannot run this analysis for inward and outward citations.

Then, we test our hypotheses about the role of qualifications by interacting our immigration variable with the indicator variable *High_qual* introduced in the previous paragraph, which takes a

value 1 if the immigrant population is comparatively highly qualified (the ratio between ISCED 3-6 codes and ISCED 0-2 is greater than the median) and zero otherwise.

In order to test our hypotheses about the role of social capital, our specification is augmented to include $ln(SocCap_{it-1})$, which corresponds to the log volume of social capital in province *i* as calculated by the Ivie (see discussion in the previous paragraph), as well as its interaction with the immigration and emigration variables.

Our model faces a common issue of gravity modeling with panel data: on the one hand, our dependent variable is a count variable which would have to be treated through Maximum Likelihood Estimators; on the other hand, our specification requires the inclusion of a large set of fixed effects meant to absorb the wide heterogeneity across units. This makes it practically impossible to apply a Poisson Pseudo-Maximum Likelihood (Poisson PML) estimator as many fixed effects hamper model convergence. This also rules out the possibility to employ a Tobit model on grounds of the incidental parameters problem (see Greene, 2004). Hence, we apply an OLS regression with fixed effects dealing with the zeros by taking the natural log of the dependent variable to which we add one unit. This yields a log-log model where the estimated coefficients can be interpreted as elasticities.

4. Results

We begin our analysis from the immigration stocks, for which we have a longer time period at our availability. Table 4 reports the results of the basic gravity specification with and without the log of immigration stocks for the three types of knowledge flows that we consider. All specifications in this paper include the three sets of fixed effects mentioned above, i.e. time-varying country effects, time-varying NUTS2 region effects, and dyadic time-invariant effects. Columns 1-4 display the results of the gravity specifications without migration variables. Column 1 reports the results of the model with coinventorships as a dependent variable using the full set of countries and years at our availability; column 2 reports the results of the same model run on the subsample of EU countries, to enable comparisons with the specifications in columns 3 and 4, where the dependent variables inward and outward citations are only available for the EU. In all specifications, the province-level "attractor" variables (province GDP and the province stock of patents) have the expected positive and significant effects on knowledge flows. The magnitude of the coefficients suggests a very important role of the number of patents compared with a more generic indicator of absorptive capacity such as province GDP[7]. The coefficient of the distance variable results negative and significant, also consistent with the predictions of the gravity model.

<<<insert Table 4 here>>>

In columns 5-8 we report the specifications including immigration stocks. Adding $ln(1+Immi_{ijt-1})$, in all cases, increases the fit of the model, as measured by both the R-squared and by the AIC statistics. The coefficient is positive and significant, supporting our hypotheses that immigrants play a role as information brokers across country boundaries. Comparing the results for different kinds of knowledge flows, the table shows that the elasticity of coinventorship ties with respect to immigration is much greater than the elasticity of citation flows (either inwards or outwards): according to our results, increasing the immigrant population by 10% would increase the count of coinventorship ties by approximately 0.3% - with no major difference between the model on the full sample of countries and the one on EU countries - , while the number of citations would increase by slightly less than 0.1%. These results support our hypothesis that immigration affects tacit knowledge flows more strongly than codified knowledge flows. It is worth noticing that the coefficient of immigration stocks on inward and outward citations is very similar, although it is only significant at a 10% level for outward flows .

[7] In a set of unreported specifications, we also used the log of the inventor stocks as an alternative measure of absorptive capacity of the province, with similar results. However, the information statistics supported the use of the log of the number of patents filed by the province instread of the number of inventors.

Table 5 reports a similar set of estimates limited to the 2006-2011 period, to which we add the log of emigrant stocks. Similarly to Table 4, columns 1-4 show that the basic gravity specification applies to this subsample as well, and that the inclusion of migration variables (immigration or emigration) generally increases the fit of the model.

<<<insert Table 5 here>>>

Comparing the results in columns 5-8 of Table 5 with the results in columns 5-8 of Table 4, however, we see that, during this period marked by economic and financial recession and by stagnation in immigration flows, the effect of immigration on knowledge flows has downsized in magnitude (column 5) and has lost significance compared with the previous estimates (columns 6-8). During the same period, however, emigration flows from Spain have steeply increased in magnitude; the estimates reported in columns 9-12 suggest that this quantitative increase also had an effect on the knowledge flows occurred between Spanish provinces and the countries of emigration. Emigration is found to promote coinventorship both with worldwide inventors and with EU inventors as well as inward and outward citation flows between the origin province and the destination country. The elasticity of coinventorship with respect to emigration to EU countries is close to 0.09, implying that a 10% increase in the emigrant population would increase the count of EU coinventorship ties and outward citations by approximately 0.9%. This effect is almost twice the one found for emigration on inward and outward citations, both close to 0.05. The effect of coinventorship with respect to emigration flows worldwide is smaller in magnitude, though still sizeable, with a point estimate of 0.049. Overall, these results are in line with our expectations, confirming on the one hand that emigration promotes knowledge flows and, on the other hand, that its effect on tacit knowledge flows is larger than on codified knowledge. In columns 13-16, we include both measures of migration stocks jointly. The results confirm the previous findings, with a significant role of emigration stocks in all cases and no significant role for immigrants except for the specification on worldwide coinventorship.

<<<insert Table 6 here>>>

In tables 6, we test our hypotheses about the role of human capital. Our results clearly show that it is qualified immigrants who promote all types of knowledge flows. This result is not surprising, considering that patent coinventorships and citations obviously require high qualifications to occur. The results are very similar when comparing the whole sample with the EU subsample, and when comparing inward and outward citations. While the low-qualified workers do not significantly affect coinventorships, they are found to decrease patent citations, possibly because the regions employing larger amounts of non-qualified workers are more likely to specialize in sectors that do not lead to patents.

<<<insert Table 7 here>>>

In table 7, we test our hypotheses about the role of language commonality. The results contradict our expectations but highlight a result that is commonly observed in the literature on the pro-trade effects of immigration (Girma and Yu, 2002; Dunlevy, 2006): the positive effect of migration stocks is mainly observed for immigrants from countries speaking languages other than Spanish. The interpretation that is usually given, which could well apply to our case, language commonality *per se* reduces barriers to information flows, independently of the immigrants' contribution. The contribution of immigrants is, instead, positive when it comes to facilitating knowledge flows with non-Spanish speaking countries, as it is actually bridging barriers to exchanging information.

In tables 8 and 9, to test our hypotheses about the role of the host province social capital, we report the analysis including province-level measures of social capital, both in isolation and in interaction with the migration stocks. Table 8 reports the results for the full sample, while Table 9 refers to the 2006-2011 period as it includes emigration variables.

<<<insert Table 8 here>>>

Looking first at the full sample in Table 8, columns 1-4 show that social capital *per se* does not result to affect knowledge flows at standard significance levels. However, columns 5-8 show that social capital has an indirect effect on knowledge flows which operates through the interaction with the migration variable. Indeed, in these specifications, the main effect of the relevant migration variables vanishes or becomes negative, while the effect of the interaction between social capital and immigration results positive and significant. Interestingly, the main effect of social capital in these specifications results to be negative and significant.

<<<insert Table 9 here>>>

Turning to the estimates for the 2006-2011 period reported in Table 9, the results are qualitatively the same as those for the full sample, and they apply to emigration as well as to immigration stocks. Notice that the interaction term between immigration stocks and social capital results positive and significant also in the specifications where the main effect of immigration stocks in isolation was not significant – even if its effect is greatly reduced in the specifications including both immigration and emigration stocks. Overall, it could be argued that the positive effect of immigration and emigration on knowledge flows is activated by higher levels of social capital[8]. Notice, however, that the interpretation of the interaction term for immigration is different from the interpretation of the interaction term for emigration. In the first case, it is a feature of the host economy, which provides a favourable environment for receiving the knowledge carried by immigrants and for applying it to economic purposed. In the latter case, it is to be interpreted as a feature of the migrating community, the Spanish diaspora, which can be expected to improve their ability to integrate in the destination economies and to promote knowledge flows.

Our results also imply that high social capital with no openness of the local economy through human mobility (either as immigration or emigration) has a negative effect on the knowledge flows of the local system. Unfortunately, in our dataset we cannot distinguish whether this operates as a reduction of knowledge flows in absolute terms or as a greater concentration on domestic citations and coinventorships, which one could interpret as an indication of the prevalence of "bonding" over "bridging" types of social capital. This issue should be addressed in future extensions of the present study. More generally, however, the result that the effect of social capital is mainly mediated by the migration stocks suggests that social capital operates more at the level of the bilateral barriers to knowledge flows than at the level of increasing or decreasing the absorptive capacity of the local system.

Summarizing, the model specifications reported in Tables 4-6 and 8-9 support our hypotheses by highlighting a positive effect of migration stocks on knowledge flows, our hypotheses about the positive effect of human capital, as well as our hypotheses about the activating role of social capital; we showed that the effect of social capital per se, instead, does not significantly affect knowledge flows. The results in Table 7 contradicted our hypothesis about the positive role of language commonality; as we saw, this result can be easily interpreted in light of the gravity literature of international trade which highlights a facilitating role for immigrants when their contribution is not redundant with other trade-facilitating factors.

When comparing the results in Tables 4 and 5, the effect of immigration on knowledge flows seems to change significantly before and after 2006. To get clearer insights on this issue, in Table 10, we split our sample into two parts and study the effects of immigration on coinventorship before and after 2006. Indeed, the before-2006 effect is quite large (close to 0.06) and significant, while the post-2006 effect is smaller and, in the cases of EU countries, insignificant. The estimates in columns (3-4) and (9-10) confirm that social capital does not affect knowledge flows as an attractor

[8] The case where the log of social capital equals zero and the effect of immigration or emigration is strictly negative is actually theoretical, considering that, by the way the social capital variable is constructed, this would imply that the level of social capital in that year is exactly equal to its level in 1983, a case that is not realized in our data. The cases where the volume of social capital is less than in 1983 amount to a 1.15% of the total. They refer to the first two years of our panel, 1998 and 1999 in five provinces: Ourense, Pontevedra, Zamora, Albacete, Caceres and Cordoba.

variable; the estimates reported in columns (5-6) and (11-12) include the interaction term and highlight an inversion in the main effect of the immigration variables: before 2006, immigration stocks were affecting knowledge flows both directly and in interaction with social capital. After 2006, the main effect is negative and it requires social capital to be converted into a positive effect. The sign of the main effect for social capital remains negative across all specifications.

Finally, in Table 11 we report a similar set of estimates with the log of inward and outward citations as dependent variables. The results are qualitatively similar, while the effects are generally smaller, as we saw in the previous paragraphs, as these estimates refer to a more codified kind of knowledge flows. Again, the effect of immigration on both types of citations is positive and significant before 2006, and much smaller and statistically insignificant after 2006. The main effect of immigration is null in the pre-2006 sample, while it is negative in the post-2006 subsample. As before, social capital results to activate the positive effect of immigration even in the post-2006 subsamples.

<<<insert Table 10 here>>>

<<<insert Table 11 here>>>

5. Conclusions

This paper has adopted a gravity model to study whether migrants promote knowledge flows between Spanish provinces and foreign countries, as measured by patent coinventorship and citations. Our results show that both immigrants and emigrants can play an important role in opening up the local system to new knowledge, as we found evidence of a positive effect of both variables on all types of knowledge flows. In most caese, the quantitative dimension of the effect is non-negligible and is fully coherent with the findings of a branch of the international trade literature that assumes that such an information effect exists and promotes trade and FDI. However, the global macroeconomic conditions affect not only the magnitude, but the composition of the immigrant population, leading to a substantial reduction in the magnitude of the effect from the pre- to the post-2006 period. The available data on emigration, which only cover the post-2006 period during which emigration has grown steadily, highlight a strongly positive effect, suggesting that the brain drain caused by the recession has a "bright side", opening up the local system to new knowledge.

Our results also highlight a number of moderating effects that affect the ability of immigrants to promote trade. First and most obviously, their skills: our data clearly showed that the knowledge-promoting role is mainly to be attributed to highly qualified immigrants. Secondly, we highlighted that the information-facilitating role is mainly to be attributed to non-Spanish speaking immigrants, who apparently bear non-redundant knowledge that can otherwise not be accessed by the local system. Not only is this result coherent with the findings of the studies on the pro-trade effect of immigrants; at a closer look, it reveals that, even when focusing on high tech sectors such as those which lead to filing patents, where the role of veicular languages such as English can be expected to be widespread, barriers that can be attributed to language differences resist and are effectively bridged by bilingual migrants. These could for instance relate to difficulties in the translation of specific terms of the technical language. Third, we showed that social capital, which in isolation is not found to significantly affect knowledge flows, has a positive moderating effect on both immigration and emigration stocks, suggesting that more inclusive local systems are more able to benefit from the information effect promoted by migrants. To the best of our knowledge, this is a novel result, which should be further explored. In particular, data availability did not allow us distinguishing between whether social capital is of a "bridging" or "bonding" type, and what the implications of different types of social capital are on knowledge flows. Extensions of this work should seek to address this point.

Bibliography

Anderson, J. and E. van Wincoop (2003). Gravity with gravitas : a solution to the border puzzle. *American Economic Review* 93, 170-92.

Baldwin, R. and D. Taglioni (2007). Gravity for dummies and dummies for gravity equations. NBER working paper series, Working Paper 12516, http://www.nber.org/papers/w12516.

Cortinovis, N., Xiao, J., Boschma, R. and van Oort, F. (2016). Quality of government and social capital as drivers of regional diversification in Europe. *Papers in Evolutionary Economic Geography* # 16.10. http://econ.geog.uu.nl/peeg/peeg.html

Bratti, M., De Benedictis, L. and Santoni, G. (2014). "On the pro-trade effects of immigrants," *Review of World Economics (Weltwirtschaftliches Archiv),* Springer; Institut für Weltwirtschaft (Kiel Institute for the World Economy), vol. 150(3), pages 557-594, August.

Cappelli, R. and F. Montobbio (2013). European Integration and Knowledge Flows across European Regions. Technical report, Department Cognetti de Martiis Working Papers 201322.

De Benedictis, L. and Taglioni, D. (2011) The Gravity Model in International Trade, in De Benedictis, L. and Salvatici, L. (eds.) "*The trade impact of European Union preferential policies: An Analysis through gravity models*", Springer, pp.55-89, 2011.

Dunlevy, J. (2006). The impact of corruption and language on the pro-trade effect of immigrants: Evidence from the American States. *Review of Economics and Statistics* 88 (1), 182-186.

Felbermayr, G., V. Grossmann, and W. Kohler (2012, October). Migration, International Trade and Capital Formation: Cause or Effect? IZA Discussion Papers 6975, Institute for the Study of Labor (IZA).

Fernandez, J., Pérez, F., and Serrano, L. (2015) Crisis Económica, Confianza y capital social. Bilbao, Fundación BBVA. http://www.fbbva.es/TLFU/dat/DE_2015_IVIE_crisis_economica.pdf

Girma, S. and Z. Yu (2002). The link between immigration and trade: Evidence from the United Kingdom. *Weltwirtschaftliches Archiv* 138, 115–30.

Gould, D. M. (1994). Immigrant links to the home country: Empirical implications for U.S. bilateral trade flows. *The Review of Economics and Statistics* 76 (2), 302-316.

Granato, J., Inglehart, R., and Leblang, D. (1996b). The effect of cultural values on economic development: theory, hypotheses, and some empirical tests. *American Journal of Political Science*, 40(3):607–631.

Greene, William, (2004), Fixed Effects and Bias Due to the Incidental Parameters Problem in the Tobit Model, *Econometric Reviews*, 23, issue 2, p. 125-147

Head, K. and T. Mayer (2014). Gravity equations: Workhorse, toolkit, and cookbook. In G. Gopinath, E. Helpman, and K. Rogo (Eds.), *Handbook of International Economics.* Elsevier.

Knack and Keefer (1997). Does social capial an economic payoff?A cross country investigation. The *Quarterly Journal of Economics*, 112:1251–1288.

Maggioni, M., T. Uberti, and S. Usai (2011). Treating patents as relational data: knowledge transfers and spillovers across Italian provinces. *Industry & Innovation* 18, 39–67.

Maurseth, P. and B. Verspagen (2002). Knowledge Spillovers in Europe: A Patent Citations Anal- ysis. *Scandinavian Journal of Economics* 104(4), 531–45.

Montobbio, Fabio & Sterzi, Valerio, 2013. "The Globalization of Technology in Emerging Markets: A Gravity Model on the Determinants of International Patent Collaborations," *World Development*, Elsevier, vol. 44(C), pages 281-299.

Paci, R. and S. Usai (2009). Knowledge flows across the European regions. *Annals of Regional Science* 43, 669–690.

Parrilli MD (2012), Heterogeneous social capital: a window of opportunity for local economies, in Cooke P., Parrilli MD and Curbelo JL, *Innovation, global change and territorial resilience*, Edward Elgar.

Peri, G. and F. Requena-Silvente (2010). The trade creation effect of immigrants: Evidence from the remarkable case of Spain. *Canadian Journal of Economics* 43 (4), 1433-1459.

Pérez, F., Montesinos, V., Serrano, L., and Fernández de Guevara, J. (2005). La Medición del Capital Social: Una Aproximación Económica. Bilbao, Fundación BBVA.

Picci, L. (2010). The internationalization of inventive activity: A gravity model using patent data. *Research Policy* 39, 1070–1081.

Quatraro, F. and S. Usai (2016). Are knowledge flows all alike? Evidence from European regions. Forthcoming in *Regional Studies*.

Rauch, J. and V. Trinidade (2002). Ethnic Chinese networks in international trade. *Review of Economics and Statistics* 84 (1), 116-30.

Tortosa-Ausina, E. and Peiró, P. (2012). "Social Capital, Investment and Economic Growth: Evidence for Spanish Provinces," Working Papers 2012122, Fundacion BBVA / BBVA Foundation.

Wagner, D., K. Head, and J. Ries (2002). Immigration and the trade of provinces. *Scottish Journal of Political Economy* 49 (5), 507-525.

Williams, A. (2007). Listen to me, learn with me: International migration and knowledge transfer. *British Journal of Industrial Relations* 45, 361–382.

Zak, P. and Knack, S. (2001). Trust and growth. *The Economic Journal*, 111(470):295–321.

Table 1. List of countries

European Union and Western Europe	Eastern Europe	Africa	Middle East and Central Asia	South and Eastern Asia	Latin America and the Caribbean	North America	Ocerania
Andorra (EU)	Croatia	Algeria	Armenia	China	Argentina	Canada	Australia
Austria (EU)	Cyprus (EU)	Egypt	United Arab Emirates	India	Bolivia	United States	New Zealand
Belgium (EU)	Bulgaria (EU)	Morocco	Israel	Japan	Brazil		
Denmark (EU)	Czech Republic (EU)	Madagascar	Jordan	Korea, Republic of	Chile		
Finland (EU)	Estonia (EU)	South Africa	Kazakhstan	Malaysia	Colombia		
France (EU)	Hungary (EU)		Uzbekistan	Pakistan	Costa Rica		
Germany (EU)	Lithuania (EU)			Saudi Arabia	Cuba		
Greece (EU)	Latvia (EU)			Singapore	Ecuador		
Ireland (EU)	Poland (EU)			Thailand	Guatemala		
Iceland	Romania (EU)				Honduras		
Italy (EU)	Russia				Mexico		
Luxembourg (EU)	Serbia, Republic of				Peru		
Malta (EU)	Slovakia				Uruguay		
Netherlands (EU)	Slovenia (EU)				Venezuela		
Norway	Turkey						
Portugal (EU)	Ukraine						
Switzerland							
Sweden (EU)							
United Kingdom (EU)							

Table 2. Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
lncoinv	51100	.0758488	.3501802	0	5.141664
lncited	18200	.0729906	.3363888	0	4.70048
lnciting	18200	.091417	.3912281	0	4.779123
lnimmi1	51100	2.745386	2.876719	0	12.29727
lnemi1	21900	2.741219	2.364601	0	10.80286
lnYi	47450	16.12232	.9198684	14.07695	19.07961
inter_i	51100	25.50159	39.06854	0	636.0176
inter_e	21900	28.23069	33.33248	0	428.8287
lnKsoc	51100	1.711651	.7743499	-.3319823	4.246955
lnpat	51100	1.995907	1.373428	0	6.144035
lninv	51100	3.53662	1.903611	0	8.488999
lnimmiSpa	51100	.705774	1.992892	0	12.06447
lnimmiNSpa	51100	2.039612	2.680104	0	12.29727
lnemiSpa	21900	.8279354	1.961881	0	10.80286
lnemiNSpa	21900	1.913284	2.216	0	9.777244
lnimmihq_e~r	42700	1.718907	2.463298	0	11.7834
lnimmiNhq~ur	42700	1.398797	2.689631	0	12.29727

Table 3 – Correlation matrix

	lncoinv	lncited	lnciting	lnimmi1	lnemi1	lnYi	lnpat	lninv	lndist	lnKsoc
lncoinv	1.0000									
lncited	0.6046*	1.0000								
lnciting	0.6256*	0.6403*	1.0000							
lnimmi1	0.2610*	0.2789*	0.2906*	1.0000						
lnemi1	0.3654*	0.3884*	0.4459*	0.5304*	1.0000					
lnYi	0.2699*	0.2785*	0.2629*	0.3503*	0.2936*	1.0000				
lnpat	0.2808*	0.3021*	0.2955*	0.2822*	0.2504*	0.8374*	1.0000			
lninv	0.2554*	0.2672*	0.2582*	0.2518*	0.2289*	0.8033*	0.9331*	1.0000		
lndist	-0.1432*	-0.1674*	-0.1936*	-0.1749*	-0.0455*	0.0029	-0.0289*	-0.0259*	1.0000	
lnKsoc	0.0795*	0.0794*	0.0777*	0.3500*	0.0068	0.3875*	0.3528*	0.3321*	0.0032	1.0000

* p-value<0.5

Table 4. Estimation results. All years (1998-2011)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var	lncoinv	lncoinv	lncited	lnciting	lncoinv	lncoinv	lncited	lnciting
L.lnYi	0.0666***	0.1226***	0.0846***	0.0701***	0.0512***	0.1021***	0.0789***	0.0643***
	(0.0097)	(0.0226)	(0.0156)	(0.0168)	(0.0092)	(0.0230)	(0.0152)	(0.0164)
L.lnpat	0.0332***	0.0597***	0.0230***	0.0417***	0.0299***	0.0556***	0.0219***	0.0405***
	(0.0041)	(0.0097)	(0.0056)	(0.0064)	(0.0038)	(0.0092)	(0.0055)	(0.0063)
lndist	-0.1547***	-0.2520***	-0.1718***	-0.1595***	-0.1220***	-0.2191***	-0.1627***	-0.1500***
	(0.0253)	(0.0486)	(0.0333)	(0.0359)	(0.0244)	(0.0488)	(0.0326)	(0.0350)
L.lnimmi1					0.0295***	0.0318***	0.0088**	0.0091*
					(0.0052)	(0.0097)	(0.0041)	(0.0049)
Region-time effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-time effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-region effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	47450	16900	16900	16900	47450	16900	16900	16900
R^2	0.585	0.607	0.527	0.551	0.588	0.610	0.528	0.552
AIC	994.7367	10794.5535	1322.2654	5564.3406	620.9364	10699.0842	1310.9126	5555.2710

BIC	16802.4162	16719.6162	7247.3281	11489.4034	16428.6160	16639.6171	7243.7104	11488.0689

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The estimates for the dependent variable lncoinv in columns (1) and (5) refer to all years and all countries; the estimates for the dependent variables lncited and lnciting, and for the dependent variable lncoinv in columns (2) and (6) refer to all years but only to EU countries.

Table 5. Estimation results (2006-2011)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	lncoinv	lncoinv	lncited	lnciting	lncoinv	lncoinv	lncited	lnciting	lncoinv	lncoinv	lncited	lnciting	lncoinv	lncoinv	lncited	lnciting
L.lnYi	0.0750***	0.1340***	0.1042***	0.1028***	0.0582***	0.1227***	0.0961***	0.1027***	0.0340***	0.0613**	0.0609***	0.0632***	0.0256**	0.0597**	0.0588***	0.0673***
	(0.0117)	(0.0273)	(0.0194)	(0.0220)	(0.0122)	(0.0293)	(0.0202)	(0.0230)	(0.0105)	(0.0254)	(0.0175)	(0.0196)	(0.0115)	(0.0279)	(0.0188)	(0.0213)
L.lnpat	0.0413***	0.0734***	0.0299***	0.0419***	0.0403***	0.0732***	0.0298***	0.0419***	0.0404***	0.0705***	0.0282***	0.0403***	0.0398***	0.0704***	0.0282***	0.0404***
	(0.0061)	(0.0144)	(0.0092)	(0.0097)	(0.0061)	(0.0143)	(0.0092)	(0.0097)	(0.0061)	(0.0143)	(0.0090)	(0.0096)	(0.0061)	(0.0143)	(0.0090)	(0.0096)
lndist	-0.1883***	-0.2633***	-0.2110***	-0.2370***	-0.1535***	-0.2436***	-0.1968***	-0.2369***	-0.1339***	-0.1401**	-0.1376***	-0.1698***	-0.1149***	-0.1373**	-0.1339***	-0.1771***
	(0.0303)	(0.0594)	(0.0421)	(0.0474)	(0.0311)	(0.0621)	(0.0432)	(0.0488)	(0.0278)	(0.0548)	(0.0379)	(0.0422)	(0.0296)	(0.0585)	(0.0399)	(0.0447)
L.lnimmi1					0.0184***	0.0117	0.0084	0.0000					0.0118**	0.0021	0.0028	-0.0053
					(0.0060)	(0.0124)	(0.0060)	(0.0076)					(0.0058)	(0.0122)	(0.0059)	(0.0075)
L.lnemi1									0.0485***	0.0898***	0.0535***	0.0489***	0.0456***	0.0893***	0.0528***	0.0502***
									(0.0067)	(0.0176)	(0.0111)	(0.0136)	(0.0064)	(0.0174)	(0.0110)	(0.0135)
N	18250	6500	6500	6500	18250	6500	6500	6500	18250	6500	6500	6500	18250	6500	6500	6500
R^2	0.637	0.659	0.576	0.604	0.638	0.660	0.576	0.604	0.641	0.665	0.580	0.606	0.641	0.665	0.580	0.606
AIC	3263.1050	5181.3709	2231.8398	4390.4059	3219.9550	5179.4901	2230.6769	4392.4058	3078.6445	5081.3857	2177.1240	4358.3769	3062.5455	5083.2596	2178.7915	4359.4828
BIC	11762.4744	8096.5806	5147.0495	7305.6156	11727.1362	8101.4794	5152.6662	7314.3951	11585.8258	8003.3750	5099.1132	7280.3662	11577.5387	8012.0284	5107.5603	7288.2516

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All specifications include country-time, region-time and region-country effects. All specifications in the table refer to the period 2006-2011 for which data about the stock of emigrants are available. The estimates for the dependent variable lncoinv in columns (1), (5), (9) and (13) refer to all countries; the estimates for the dependent variables lncited and lnciting, and for the dependent variable lncoinv in columns (2), (6), (10) and (14) refer only to EU countries.

Table 6. Estimation results. Specifications including interactions with *High_qual* dummy

	(1)	(2)	(3)	(4)
	lncoinv	lncoinv	lncited	lnciting
L.lnYi	0.0547***	0.1109***	0.0886***	0.0723***
	(0.0105)	(0.0268)	(0.0178)	(0.0191)
L.lnpat	0.0327***	0.0609***	0.0241***	0.0448***
	(0.0044)	(0.0108)	(0.0063)	(0.0073)
lndist	-0.1315***	-0.2438***	-0.1848***	-0.1710***
	(0.0272)	(0.0569)	(0.0380)	(0.0408)
L.lnimmihq_edur	0.0523***	0.0561***	0.0221***	0.0220**
	(0.0081)	(0.0131)	(0.0073)	(0.0086)
L.lnimmiNhq_edur	-0.0047	-0.0155	-0.0241***	-0.0216***
	(0.0072)	(0.0206)	(0.0060)	(0.0069)
N	39650	13650	13650	13650
R^2	0.600	0.624	0.549	0.565
AIC	4130.3519	10007.9552	2434.2585	6313.9698
BIC	17347.0472	14957.0988	7390.9235	11263.1134

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All specifications include country-time, region-time and region-country effects. The specification in col. 1 refers to the full sample of countries; the specifications in columns 2-4 refer to EU countries only.

Table 7. Estimation results. Specifications including interactions with *Spanish-speaking* dummy

	(1)	(2)	(3)
	lncoinv	lncoinv	lncoinv
L.lnYi	0.0533***	0.0300***	0.0240**
	(0.0093)	(0.0103)	(0.0114)
L.lnpat	0.0306***	0.0400***	0.0402***
	(0.0039)	(0.0061)	(0.0061)
lndist	-0.1275***	-0.1477***	-0.1313***
	(0.0245)	(0.0285)	(0.0302)
L.lnimmiSpa	-0.0267***		-0.0232***
	(0.0051)		(0.0053)
L.lnimmiNSpa	0.0391***		0.0171***
	(0.0060)		(0.0065)
L.lnemiSpa		-0.0238***	-0.0073*
		(0.0056)	(0.0043)
L.lnemiNSpa		0.0742***	0.0660***
		(0.0091)	(0.0086)
N	47450	18250	18250
R^2	0.591	0.646	0.647
AIC	259.5331	2825.2012	2778.4015
BIC	16075.9801	11340.1943	11309.0185

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All specifications include country-time, region-time and region-country effects. The specification in col. 1 refers to the full time period at our disposal (1998-2011); the specifications in columns 2-3 refer to the period for which emigration data are available, i.e. 2006-2011.

Table 8. Estimation results (1998-2011). Specifications including social capital and its interactions with the migration variables.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	lncoinv	lncoinv	lncited	lnciting	lncoinv	lncoinv	lncited	lnciting
L.lnYi	0.0513***	0.1025***	0.0791***	0.0645***	0.0525***	0.1070***	0.0814***	0.0675***
	(0.0093)	(0.0231)	(0.0153)	(0.0165)	(0.0093)	(0.0235)	(0.0155)	(0.0167)
L.lnpat	0.0303***	0.0564***	0.0224***	0.0410***	0.0305***	0.0559***	0.0221***	0.0406***
	(0.0039)	(0.0093)	(0.0055)	(0.0062)	(0.0039)	(0.0093)	(0.0054)	(0.0061)
lndist	-0.1220***	-0.2193***	-0.1628***	-0.1502***	-0.1243***	-0.2138***	-0.1600***	-0.1465***
	(0.0244)	(0.0488)	(0.0326)	(0.0351)	(0.0246)	(0.0484)	(0.0322)	(0.0347)
L.lnimmi1	0.0297***	0.0324***	0.0092**	0.0094*	0.0024	-0.0253	-0.0201**	-0.0288***
	(0.0053)	(0.0098)	(0.0041)	(0.0049)	(0.0078)	(0.0163)	(0.0080)	(0.0091)
L.lnKsoc	-0.0051	-0.0123	-0.0071	-0.0066	-0.0355***	-0.0929***	-0.0480***	-0.0601***
	(0.0060)	(0.0141)	(0.0084)	(0.0104)	(0.0067)	(0.0198)	(0.0126)	(0.0142)
L.lninter_i					0.0118***	0.0265***	0.0135***	0.0176***
					(0.0024)	(0.0060)	(0.0034)	(0.0041)
N	47450	16900	16900	16900	47450	16900	16900	16900
R^2	0.588	0.610	0.528	0.552	0.589	0.613	0.529	0.554
AIC	618.7795	10698.6248	1315.4836	5556.2982	496.7267	10577.1114	1259.1034	5486.1188
BIC	16426.4590	16646.8928	7271.4866	11496.8311	16321.9411	16525.3794	7207.3714	11434.3867

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All specifications include country-time, region-time and region-country effects. The estimates for the dependent variable lncoinv in columns (1) and (5) refer to all years and all countries; the estimates for the dependent variables lncited and lnciting, and for the dependent variable lncoinv in columns (2) and (6) refer to all years but only to EU countries.

Table 9. Estimation results (2006-2011). Specifications including social capital and its interactions with the migration variables.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
	lncoinv	lncoinv	lncited	lnciting	lncoinv	lncoinv	lncited	lnciting	lncoinv	lncoinv	lncited	lnciting	lncoinv	lncoinv	lncited	lnciting	lncoinv	lncoinv	lncited	lnciting
L.lnYi	0.0590***	0.1242***	0.0985***	0.1035***	0.0644***	0.1374***	0.1042***	0.1144***	0.0326***	0.0574**	0.0622***	0.0616***	0.0304***	0.0514**	0.0595***	0.0577***	0.0272**	0.0613**	0.0621***	0.0715***
	(0.0124)	(0.0298)	(0.0207)	(0.0236)	(0.0126)	(0.0307)	(0.0215)	(0.0243)	(0.0109)	(0.0265)	(0.0184)	(0.0207)	(0.0107)	(0.0257)	(0.0178)	(0.0205)	(0.0118)	(0.0284)	(0.0193)	(0.0221)
L.lnpat	0.0404***	0.0733***	0.0299***	0.0420***	0.0388***	0.0691***	0.0281***	0.0385***	0.0403***	0.0703***	0.0282***	0.0403***	0.0379***	0.0660***	0.0264***	0.0376***	0.0373***	0.0650***	0.0259***	0.0362***
	(0.0061)	(0.0144)	(0.0092)	(0.0097)	(0.0060)	(0.0142)	(0.0090)	(0.0094)	(0.0061)	(0.0142)	(0.0090)	(0.0096)	(0.0059)	(0.0135)	(0.0087)	(0.0093)	(0.0059)	(0.0135)	(0.0087)	(0.0093)
lndist	-0.1541***	-0.2448***	-0.1986***	-0.2375***	-0.1507***	-0.2385***	-0.1958***	-0.2323***	-0.1319***	-0.1343**	-0.1393***	-0.1678***	-0.0945***	-0.0799	-0.1158***	-0.1335***	-0.0853***	-0.0910	-0.1173***	-0.1491***
	(0.0312)	(0.0626)	(0.0436)	(0.0492)	(0.0309)	(0.0622)	(0.0431)	(0.0486)	(0.0284)	(0.0562)	(0.0390)	(0.0435)	(0.0277)	(0.0544)	(0.0370)	(0.0416)	(0.0296)	(0.0581)	(0.0388)	(0.0439)
L.lnimmi1	0.0189***	0.0126	0.0098	0.0005	-0.0237***	-0.0709***	-0.0266*	-0.0686***									-0.0038	-0.0356*	-0.0137	-0.0487***
	(0.0060)	(0.0125)	(0.0061)	(0.0077)	(0.0091)	(0.0228)	(0.0137)	(0.0169)									(0.0083)	(0.0197)	(0.0117)	(0.0140)
L.lnemi1									0.0489***	0.0914***	0.0529***	0.0496***	-0.0250**	-0.0493*	-0.0068	-0.0375	-0.0201	-0.0371	-0.0012	-0.0208
									(0.0066)	(0.0176)	(0.0109)	(0.0134)	(0.0124)	(0.0291)	(0.0179)	(0.0236)	(0.0124)	(0.0287)	(0.0178)	(0.0236)
L.lnKsoc	-0.0067	-0.0120	-0.0189*	-0.0060	-0.0616***	-0.1271***	-0.0690***	-0.1013***	0.0050	0.0111	-0.0042	0.0052	-0.0620***	-0.1196***	-0.0601***	-0.0763***	-0.0727***	-0.1461***	-0.0732***	-0.1120***
	(0.0071)	(0.0171)	(0.0111)	(0.0135)	(0.0107)	(0.0298)	(0.0193)	(0.0231)	(0.0063)	(0.0152)	(0.0104)	(0.0129)	(0.0094)	(0.0235)	(0.0150)	(0.0182)	(0.0118)	(0.0324)	(0.0205)	(0.0250)
L.lninter_i					0.0168***	0.0325***	0.0142***	0.0269***									0.0046*	0.0109*	0.0052	0.0148***
					(0.0032)	(0.0082)	(0.0050)	(0.0062)									(0.0025)	(0.0065)	(0.0039)	(0.0048)
L.lninter_e													0.0322***	0.0626***	0.0266***	0.0387***	0.0291***	0.0581***	0.0241***	0.0327***
													(0.0053)	(0.0123)	(0.0076)	(0.0096)	(0.0051)	(0.0117)	(0.0073)	(0.0093)
N	18250	6500	6500	6500	18250	6500	6500	6500	18250	6500	6500	6500	18250	6500	6500	6500	18250	6500	6500	6500
R^2	0.638	0.660	0.576	0.604	0.640	0.663	0.578	0.607	0.641	0.665	0.580	0.606	0.645	0.674	0.583	0.610	0.646	0.674	0.583	0.611

AIC	3220.3222	5180.4661	2228.6956	4394.1134	3136.6238	5123.2949	2213.0671	4350.4005	3079.5162	5080.3070	2178.9241	4360.1613	2858.8766	4908.4686	2131.8669	4287.4665	2848.2372	4904.8697	2133.8746	4275.7301
BIC	11735.3154	8109.2350	5157.4644	7322.8823	11659.4290	8058.8433	5148.6155	7285.9489	11594.5094	8002.2962	5107.6929	7288.9301	11389.4936	7844.0170	5067.4153	7223.0148	11386.6662	7853.9772	5082.9821	7224.8376

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All specifications include country-time, region-time and region-country effects. All specifications in the table refer to the period 2006-2011 for which data about the stock of emigrants are available. The estimates for the dependent variable lncoinv in columns (1), (5), (9) and (13) refer to all countries; the estimates for the dependent variables lncited and lnciting, and for the dependent variable lncoinv in columns (2), (6), (10) and (14) refer only to EU countries.

Table 10. Estimation results. Dependent variable: *lncoinv.* pre and post-2006 subsamples.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	lncoinv	lncoinv	lncoinv	lncoinv	lncoinv	lncoinv	lncoinv	lncoinv	lncoinv	lncoinv	lncoinv	lncoinv
L.lnYi	0.0437***	0.0582***	0.0433***	0.0590***	0.0448***	0.0644***	0.0865***	0.1227***	0.0854***	0.1242***	0.0916***	0.1374***
	(0.0091)	(0.0122)	(0.0090)	(0.0124)	(0.0091)	(0.0126)	(0.0227)	(0.0293)	(0.0225)	(0.0298)	(0.0230)	(0.0307)
L.lnpat	0.0212***	0.0403***	0.0220***	0.0404***	0.0221***	0.0388***	0.0387***	0.0732***	0.0408***	0.0733***	0.0402***	0.0691***
	(0.0035)	(0.0061)	(0.0036)	(0.0061)	(0.0036)	(0.0060)	(0.0085)	(0.0143)	(0.0085)	(0.0144)	(0.0085)	(0.0142)
lndist	-0.1024***	-0.1535***	-0.1011***	-0.1541***	-0.1046***	-0.1507***	-0.1956***	-0.2436***	-0.1924***	-0.2448***	-0.1921***	-0.2385***
	(0.0228)	(0.0311)	(0.0225)	(0.0312)	(0.0227)	(0.0309)	(0.0477)	(0.0621)	(0.0471)	(0.0626)	(0.0470)	(0.0622)
L.lnimmi1	0.0573***	0.0184***	0.0575***	0.0189***	0.0360***	-0.0237***	0.0603***	0.0117	0.0609***	0.0126	0.0090	-0.0709***
	(0.0107)	(0.0060)	(0.0108)	(0.0060)	(0.0122)	(0.0091)	(0.0152)	(0.0124)	(0.0153)	(0.0125)	(0.0189)	(0.0228)
L.lnKsoc			-0.0070	-0.0067	-0.0235***	-0.0616***			-0.0204	-0.0120	-0.0819***	-0.1271***
			(0.0067)	(0.0071)	(0.0058)	(0.0107)			(0.0155)	(0.0171)	(0.0177)	(0.0298)
L.lninter_i					0.0109***	0.0168***					0.0281***	0.0325***
					(0.0033)	(0.0032)					(0.0077)	(0.0082)
N	25550	18250	25550	18250	25550	18250	9100	6500	9100	6500	9100	6500
R^2	0.595	0.638	0.595	0.638	0.595	0.640	0.605	0.660	0.606	0.660	0.608	0.663
AIC	-5978.0403	3219.9550	-5980.2437	3220.3222	-6022.7969	3136.6238	3659.6127	5179.4901	3658.0268	5180.4661	3601.7009	5123.2949
BIC	4362.2699	11727.1362	4360.0665	11735.3154	4325.6617	11659.4290	7324.3680	8101.4794	7329.8981	8109.2350	7280.6882	8058.8433

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All specifications include country-time, region-time and region-country effects. All specifications in the table include *lncoinv* as a dependent variable. The subsamples used for the estimates in columns (1)- (6) include all countries; the subsamples used for the estimates in columns (7)- (12)

include EU countries only. The estimates in columns (1), (3), (5) and in columns (7), (9), (11) are based on the pre-2006 sample; the estimates in columns (2), (4), (6) and (8), (10), (12) are based on the post-2006 sample.

Table 11. Estimation results. Dependent variables: *lncited* and *lnciting.* Pre and post-2006 subsamples.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	lncited	lncited	lncited	lncited	lncited	lncited	lnciting	lnciting	lnciting	lnciting	lnciting	lnciting
L.lnYi	0.0693***	0.0961***	0.0689***	0.0985***	0.0710***	0.1042***	0.0389***	0.1027***	0.0382***	0.1035***	0.0407***	0.1144***
	(0.0148)	(0.0202)	(0.0146)	(0.0207)	(0.0149)	(0.0215)	(0.0145)	(0.0230)	(0.0143)	(0.0236)	(0.0145)	(0.0243)
L.lnpat	0.0091*	0.0298***	0.0099*	0.0299***	0.0097*	0.0281***	0.0346***	0.0419***	0.0359***	0.0420***	0.0357***	0.0385***
	(0.0054)	(0.0092)	(0.0054)	(0.0092)	(0.0053)	(0.0090)	(0.0057)	(0.0097)	(0.0057)	(0.0097)	(0.0056)	(0.0094)
lndist	-0.1463***	-0.1968***	-0.1451***	-0.1986***	-0.1450***	-0.1958***	-0.0882***	-0.2369***	-0.0861***	-0.2375***	-0.0860***	-0.2323***
	(0.0315)	(0.0432)	(0.0310)	(0.0436)	(0.0309)	(0.0431)	(0.0312)	(0.0488)	(0.0305)	(0.0492)	(0.0304)	(0.0486)
L.lnimmi1	0.0136***	0.0084	0.0138***	0.0098	-0.0036	-0.0266*	0.0223***	0.0000	0.0227***	0.0005	0.0015	-0.0686***
	(0.0050)	(0.0060)	(0.0051)	(0.0061)	(0.0095)	(0.0137)	(0.0058)	(0.0076)	(0.0058)	(0.0077)	(0.0100)	(0.0169)
L.lnKsoc			-0.0077	-0.0189*	-0.0284***	-0.0690***			-0.0132	-0.0060	-0.0383***	-0.1013***
			(0.0092)	(0.0111)	(0.0106)	(0.0193)			(0.0101)	(0.0135)	(0.0112)	(0.0231)
L.lninter_i					0.0095**	0.0142***					0.0115**	0.0269***
					(0.0046)	(0.0050)					(0.0050)	(0.0062)
N	9100	6500	9100	6500	9100	6500	9100	6500	9100	6500	9100	6500
R^2	0.517	0.576	0.517	0.576	0.518	0.578	0.535	0.604	0.535	0.604	0.535	0.607
AIC	-1987.2774	2230.6769	-1984.2216	2228.6956	-1996.4676	2213.0671	-765.7101	4392.4058	-766.1255	4394.1134	-779.9100	4350.4005
BIC	1677.4779	5152.6662	1694.7658	5157.4644	1682.5198	5148.6155	2899.0451	7314.3951	2905.7458	7322.8823	2899.0773	7285.9489

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All specifications include country-time, region-time and region-country effects. All specifications in the table refer to EU countries. The dependent variable of the models in columns (1)- (6) is inward citations *lncited*; the dependent variable of the models in columns (7)- (12) is outward citations

lnciting. The estimates in columns (1), (3), (5) and in columns (7), (9), (11) are based on the pre-2006 sample; the estimates in columns (2), (4), (6) and (8), (10), (12) are based on the post-2006 sample.