Invest in Friends or Foreigners? Social Connectedness and Foreign Direct Investment*

Felix Dornseifer†

Oliver Rehbein‡

December 16, 2022

Abstract

Using Facebook friendship links as a proxy for social ties for a large set of country pairs across the world, this paper shows that cross-country social connectedness is an important determinant of foreign direct investment (FDI). We find that social connections can account for a significant share of more traditional cross-country determinants such as distance. Importantly, we show that social connections can help to overcome three different types of investment hurdles: information asymmetries, bad institutions, and global (macroeconomic) uncertainty. To strengthen identification, we exploit migration and genetic distance as instruments.

JEL classification: F21, F23, G41

Keywords: FDI, investment, social connectedness, information asymmetries, institutions, uncertainty

* We thank Andrea Ariu, Jonathan Comouth, Christiane Hellmanzik, Vanessa Hellwig, Maren Kaliske, Christian Laux, Simon Rother and seminar participants at TU Dortmund, the University of Bonn and the Vienna University of Economics and Business (WU) for valuable comments and suggestions. We also thank participants of the RGS Doctoral Conference 2022, the 21st Doctoral Meeting in International Trade and International Finance in Paris and the 26th Spring Meeting of Young Economists in Orléans.

†TU Dortmund, Department of Economics, Vogelpothsweg 87, 44227 Dortmund. Phone: +49 (0) 231 755-6822. Email: felix.dornseifer@tu-dortmund.de

‡Vienna University of Economics and Business, Institute for Finance, Banking and Insurance, Welthandelsplatz 1, 1020 Vienna, Austria. Email: oliver.rehbein@wu.ac.at

1 Introduction

Foreign Direct Investment (FDI) peaked with a total value of more than 2 trillion USD worldwide in 2015. It has been the largest source of external financing for developing countries over the past decade (UNCTAD, 2019). Such long-term strategic investments require an exceptionally high degree of information, especially when crossing geographical, linguistic, or cultural borders. In line with an emergent literature demonstrating that social connections play an important role in transmitting information in financial markets (Kuchler and Stroebel, 2021), we show an economically significant relationship between social connectedness and FDI: a 1% increase in social connectedness across borders is associated with a 0.16% increase in FDI, controlling for a large set of covariates. We extensively document that social connections are especially important in overcoming three different investment frictions: information asymmetries about the investment, bad institutions in the destination country, and (global) economic and political uncertainty.

To measure social connectedness, we use a new data set on Facebook friendship links between countries (Bailey et al., 2018). The Social Connectedness Index (SCI) provides a snapshot of the probability that a person in country A is friends with a person in country B. In 2020, the number of monthly active Facebook users exceeded one third of the world's population for the first time (Facebook, 2020). As a result, friendship links on Facebook provide a picture of social links on an unprecedented scale. We hypothesize that these links are - directly or indirectly - leveraged in investment decisions, especially in the presence of major economics frictions. We combine this data with bilateral FDI data from the World Investment Report and add a large set of gravity and cultural variables from the Centre d' Études Prospectives et d'Informations Internationales's (CEPII), the World Bank, the International Monetary Fund (IMF), the Varieties of Democracy Institute (V-Dem), the World Governance Indicators Project (WGI) and the World Uncertainty Index (WUI).

Using a gravity model, we show that social connectedness is associated with a significant in-

crease in FDI, even when controlling for a battery of factors that are known to be the main determinants of FDI, such as distance and a common language. The bilateral panel structure of the data also allows us to control for an extensive set of fixed effects that can account for most relevant determinants and time trends of FDI in our large set of countries. Our results show that without accounting for social connectedness, the effects of classical gravity covariates, such as distance and language are significantly overestimated. We address remaining endogeneity concerns by using historic bilateral migration and genetic distance as instrumental variables (IV). The results bolster our findings and show that social connections can be traced back to the movement and interaction of people long before Facebook started to gather data on how individuals are linked across the globe.

We first present evidence that information asymmetries matter. The effect of social connections on FDI is much stronger for greenfield investments, which require more information about the local environment in the destination country. Social connectedness is also significantly more important when trade between countries is low and where no regional trade agreements exist. It is also more important in countries that do not share a common language, or legal history and are politically dissimilar. In all, these sets of results suggest that social connectedness is much more important when the demand for information is high and when information flows between countries are hampered by language or history.

We next demonstrate that social connections are very important in overcoming bad institutional quality. We first demonstrate that social connectedness is significantly more important when procedures for starting businesses are longer, more difficult, and more costly. In addition, financial markets are very important: Well-developed financial markets significantly reduce the importance of social connectedness for FDI. Better institutions, such as corruption and political stability in the destination country also diminish the social connectedness effect. Interestingly, social connectedness is also more important when the exclusion of certain socio-economic groups is high. These results suggest that social connections are helpful to overcome bad institutions, which provides important insight for public policy as well as for corporate investment strategies.

Further, our results suggest that institutions enable investment without necessitating (informal) social networks, which are likely difficult to form and maintain over large distances.

Last, we show that social connections can mitigate uncertainty, which is associated with declines in output in general (Baker et al., 2016) and foreign direct investment in particular (Choi et al., 2021). We use the World Uncertainty Index (Ahir et al., 2022) to show that social connections mitigate uncertainty in the destination country, especially when greenfield investment is prominent. However, our results suggest that social connectedness is most important in overcoming *global* uncertainty. Since uncertainty is on the rise across the globe, social networks are likely to play an increasingly important role in global investment and output, if current trends with regard to global uncertainty continue.

This paper relates to three separate strands of literature: the determinants of FDI, the role of social networks and the importance of investment frictions. Empirical research on FDI determinants has been mainly focused on either the destination country or the originating country (Biswas, 2002; Blonigen and Piger, 2014; Paul and Feliciano-Cestero, 2021). On the country-pair level, the most important determinants are trade (Büthe and Milner, 2008; Berger et al., 2013; Kox and Rojas-Romagosa, 2020) and distance (Egger and Pfaffermayr, 2004). A large literature shows a negative impact of distance on different forms of investment, for example in international investment portfolios (Coval and Moskowitz, 1999; Daude and Fratzscher, 2008) and equity holdings (Ahearne et al., 2004; Portes and Rey, 2005). We expand this literature by highlighting the importance of social networks as a key determinant for FDI and demonstrate that social connectedness can fully account for the relationship between distance and FDI in most of our specifications.

Social networks play an important role in many economic areas, especially when the exchange of information is important (Kuchler and Stroebel, 2021). One strand of evidence points to the importance of networks for household finance decisions, for example for stock market participation, choice of particular stocks, and household debt (Brown et al., 2008; Georgarakos et al., 2014; Ouimet and Tate, 2020; Balakina et al., 2022). The literature also documents the

measurable effects of networks on professional investors (Hong et al., 2005; Pool et al., 2015) and trade (Greif, 1989; Rauch, 2001; Combes et al., 2005). While many of these papers rely on survey data, or a relatively specific set of networks, newly available data from social networks, especially Facebook, allows measurement of social networks on a much broader scale. This type of broad social connectedness is strongly correlated with (housing market) investment (Bailey et al., 2019; Kuchler et al., 2022a), bank lending (Rehbein and Rother, 2020) and trade (Bailey et al., 2021).[1] Using this social connectedness data from Facebook (Bailey et al., 2018), we add to this emerging literature in two key ways: First, we demonstrate the importance of social connections in foreign direct investment. Second, we demonstrate additional channels through which social connectedness might function. While Bailey et al. (2021) test whether social connectedness can overcome bad institutions in trade, they do not find significant effects. This appears to be different for investment: we show that social connectedness can help to overcome various institutional frictions such as contract enforcement and even the negative effects stemming from discrimination of minorities. Importantly, we also demonstrate that social connectedness can help to overcome investment frictions stemming from uncertainty, which to the best of our knowledge has not been shown before in any context.

A very large literature is concerned with the importance of institutions in general (Rodrik, 2000; Acemoglu et al., 2001; Rodrik et al., 2004) and for trade (Anderson and Marcouiller, 2002; Berkowitz et al., 2006; Ranjan and Lee, 2007) and foreign direct investment in particular (Daude and Stein, 2007; Nunn, 2007). There is broad agreement that bad institutions pose a significant hurdle to investment, and as a result to economic growth. Yet, there is little to no evidence on how this friction can be overcome, other than by improving the institutions themselves. We provide evidence that in the presence of bad institutions, social connections can be leveraged to partially overcome some of these frictions. This is an important insight because

[1]The literature on the economic effects of social networks is large and goes beyond Facebook's SCI data. Field experiments (Allcott et al., 2020), the distribution of hyperlinks (Hellmanzik and Schmitz, 2017) or patent citations (Agrawal et al., 2008) have been used to study the effect on different outcome variables. Paniagua et al. (2017) describe the effect of the *use* of social networks on FDI. Our approach does not rely on direct effects of online social networks, but rather uses them as a reflection of underlying real-world relationships, through which information flows. All of these methods have the disadvantage of covering only a small fraction of global social connectedness.

social connections can arguably be more easily influenced by both private and public policy than institutions, especially from the originating country. For example, it may be significantly cheaper for a firm to send employees abroad to build a social network, than to actively attempt to improve contract enforcement standards in the destination country.

Increasingly, the (macroeconomic) literature has focused on uncertainty as an important determinant of investment flows. Especially recent advances in using text as data have led to an increasing variety of uncertainty measures (Jurado et al., 2015; Segal et al., 2015; Baker et al., 2016; Bekaert et al., 2022). Not surprisingly, uncertainty has also been shown to be an important determinant of foreign direct investment (Julio and Yook, 2016; Hsieh et al., 2019; Avom et al., 2020; Choi et al., 2021; Nguyen and Lee, 2021; Jardet et al., 2022).[2] And while there exists evidence that uncertainty might also spread via social networks (Ma et al., 2022), we show that the existence of social networks can help to *overcome* investment frictions resulting from high uncertainty, especially global uncertainty. In our view, this is an important extension of the uncertainty literature, not only because it suggests that social connections can help to bridge ever-increasing uncertainty, but also because a weakening social network might amplify already existing effects of uncertainty.

2 Data and Strategy

The baseline approach explains FDI flows from an origin country i to a destination country j by the social connectedness of these two countries. The underlying gravity model controls not only for standard gravity covariates but also for additional economic and cultural covariates. For this baseline result we use a panel data set covering FDI flows from 2010 to 2019 on a country-pair level.

[2]Cascaldi-Garcia et al. (2020) provide a summary of all measures and their numerous effects on various economic outcomes.

2.1 Data

Inward foreign direct investment: Every year, trends in FDI are examined and published in the World Investment Report by the UNCTAD. This is one of the most comprehensive databases on FDI and contains information on inflows as well as outflows. Our study concentrates on FDI inflows from 2010 to 2019. FDI is defined as an investment that gives its investor, resident in country i, control or a significant degree of influence over the management of a firm in country j (UNCTAD, 2020). Since FDI flows per investment are presented on a net basis, reverse investment, negative retained earnings or disinvestment can lead to negative FDI. To be able to estimate a gravity model, which depends on working with count data, we replace these negative values of FDI by zero.

Social connectedness: To proxy for social connections between countries we use Facebook's Social Connectedness Index, which is described in details for the US county level in Bailey et al. (2018). The methodology carries over to the country level and defines the connectedness between country i and country j as follows:

$$\text{SCI}_{i,j} = \frac{\text{Facebook Connections}_{i,j}}{\text{Facebook Users}_i \cdot \text{Facebook Users}_j} \tag{1}$$

The SCI can be interpreted as the relative probability of a Facebook user in country i being friends with a Facebook user in country j. To preserve confidentiality, Facebook multiplies the result of Equation 1 by a random factor. Accordingly, no conclusions can be drawn on the actual user data, and the SCI takes values between one and 1 billion. Nevertheless, the measure is well suited to compare the degree of social connections between different country pairs. The SCI enters the model as a time-invariant variable with data from August 2020, open to the public at the Humanitarian Data Exchange.[3] To account for outliers in the distribution of the

[3]The latest available data of the SCI is for October 2021 (see Humanitarian Data Exchange). However we use data from August 2020, because it is closer in time to our FDI data. After an inquiry for historical data, Facebook clarified that for large aggregations such as the country-to-country level there is little change of the SCI, such that it makes little difference which vintages are used in practice.

SCI, the data are winsorized at the 99$^{\text{th}}$ percentile. 2.74 billion monthly active users worldwide in the third quarter of 2020 (Facebook, 2020) suggests that the usage of the social network is internationally pervasive and thus provides an unprecedented picture of social networks across the globe. Moreover, several studies show that Facebook friendships most likely map real-world connections between relatives, friends and colleagues (Bailey et al., 2018; Kuchler et al., 2022b).

Gravity covariates: The classic gravity covariates are obtained from the CEPII Gravity database. These include the population weighted distance in km between countries i and j, which is the distance between the centroid of country i's population and the centroid of country j's population. They also include indicator variables on the country-pair level for a common border, a colonial relationship after 1945, a common colonizer and the same official language.

Economic covariates: As economic covariates we include total bilateral trade, taken from the Direction of Trade Statistics of the IMF, absolute GDP difference and the absolute difference in GDP growth rate in percentage points, both taken from World Bank Open Data. The latter controls for the economic cycle and for differences in the economic development of the origin and destination country. The more aligned the economic cycle of two countries are, the better is the investor's ability to evaluate the investment in the destination country. Total trade volume between the origin and destination country is constructed by the sum of country i's imports from country j and country j's imports from country i. Total trade is intended to control for the effect of existing business relationships. Because engaging in FDI and establishing lasting interest in a foreign country requires more commitment than simply trading goods or services, investors might favor destinations which they already know well because of an existing trade record. Since investment decisions require a lead time, all these economic covariates enter as averaged values over the three years prior to the investment. We also control for the existence of a regional trade agreement using data from the CEPII Gravity database (also lagged by three years) and a dummy variable indicating whether a country pair shares the same currency to control for exchange rate uncertainty that may deter cross-border investments (Schiavo, 2007).

Culture covariates: We include a dummy for sharing a common ethnological language, spoken by at least 9% of the population and an indicator for a common legal origin from the CEPII Gravity database to control for cultural differences between countries. Religion is an important factor that influences behavior of people and governments (La Porta et al., 1999). We use data on religious diversity from the Pew Research Center to construct a shared religion index. We also control for proximity of the countries' political systems using data from the Varieties of Democracy Institute.

Instrumental vairables: To bolster identification we implement an instrumental variable approach using migration data from the Global Bilateral Migration Database of the World Bank as well as data on genetic distance from Spolaore and Wacziarg (2018). We combine the migration data over five consecutive census rounds from 1960 to 2000 to get a bilateral measure for migratory movements between destination and origin countries of FDI. The genetic distance builds on the time elapsed between a destination and origin country of investment since having the last common ancestors.

Institutional covariates: We include variables from the Ease of Doing Business Index from the World Bank which is included in the CEPII Gravity database to capture the days, procedures and costs that occur when starting a business in the destination and origin country of the investment. Next to that we add data on the financial development in the origin and destination country, taken from the IMF Financial Development Index and covariates from the Worldwide Governance Indicators (WGI). Lastly we use a measure for the exclusion of socio-economic groups taken from the Varieties of Democracy Institute.

Uncertainty measures: To proxy for country-specific and worldwide uncertainty, the World Uncertainty Index (WUI) is included. We address stock market uncertainty by refering to the VIX Index that expresses the stock market's expectation of volatility of the U.S. equity index S&P 500. This index is published by the Chicago Board Options Exchange (CBOE).

2.2 Descriptive Statistics

The final data set for the baseline regression contains cross-sectional data on 151 origin countries and 109 destination countries. Overall, we cover a total of 5,904 country pairs with 40,343 observations over 10 years from 2010 to 2019.[4] For almost 40% of the country pairs (2,305) the data spans all 10 years of our panel. However, there are also 661 country pairs for which only one year of data is available. Table A1 in the appendix presents all data sources and provides a definition of each variable. Summary statistics are shown in Table 1.

– Table 1 around here –

The median FDI is 0.78 million USD and the distribution is highly right-skewed. Trade volume and differences in GDP and its growth rate show a large dispersion, which reflects a high variety in the economic performance of the countries across the world. Interpreting these numbers on their own is not very insightful because they are constructed as bilateral proxies for differences in economic development. Sizable differences reflect the fact that we have a large and diverse set of countries in the dataset. We check the most important correlations for SCI in table A2 in the appendix: It is most strongly correlated with distance (-0.29), the dummy for a common official language (0.29) and the dummy for a common border (0.26). The further apart two countries are, the weaker are the social ties. Language appears to be a similarity that promotes and simplifies social contacts. This is also supported by a correlation of 0.25 between the SCI and a dummy for a common ethnological language which is spoken by at leat 9% of the population. The indicator variable for a common colonizer as well as the one for regional trade agreements are comparable in size. Next to that average migration between 1960

[4]Due to governmental censorship of Facebook, the SCI data is not available for China, Iran, North Korea, Tajikistan and Turkmenistan. Additionally, some small or off-shore countries are excluded because they are known as tax-havens and often do not represent the real origin or destination of investments. Furthermore some countries are not included due to a lack of data on covariates. This applies to the following countries: Andorra, Curacao, Cyprus, Dominica, French Guiana, Guadeloupe, Guam, Isle of Man, Jersey, Kiribati, Luxembourg, Macao, Martinique, Mauritius, Mayotte, Montenegro, Panama, Puerto Rico, Reunion, Rwanda, San Marino, St. Kitts and Nevis, Tuvalu and Western Sahara.

and 2000 as well as between 1990 and 2010 are strongly correlated with the SCI (0.24). All other coefficients range between 0.18 and -0.19.

2.3 Empirical Specification

Gravity models are often used in economics to explain various directional flows, such the flow of goods, money, labor and people between countries. Gravity models have also been frequently used to describe FDI flows (see for example Portes and Rey, 2005; Bergstrand and Egger, 2007; Head and Ries, 2008; Okawa and Van Wincoop, 2012). In its most basic form, a gravity model explains FDI flows by the GDP of the origin and the destination country and the distance between them. Thereby the GDP serves as a proxy for the supply and demand for cross-border investments while distance represents the frictions of such investments (i.e. transaction costs). To study the relationship between social connectedness and FDI, we follow this literature and estimate the following gravity equation:

$$\mathrm{FDI}_{i,j,t} = \exp[\beta_1 \cdot \log(\mathrm{SCI}_{i,j}) + \beta_2 \cdot \mathrm{G}_{i,j,t} + \beta_3 \cdot \mathrm{A}_{i,j,t} + \delta_{i,t} + \delta_{j,t}] \cdot \epsilon_{i,j,t} \tag{2}$$

The left-hand side depicts FDI as the dependent variable between country i and j at time t. On the right-hand side, we include log(SCI) as the main explanatory variable, the vector of gravity ($\mathrm{G}_{i,j,t}$) and the vectors of economic and cultural covariates ($\mathrm{A}_{i,j,t}$), both described in Section 2.1. $\delta_{i,t}$ and $\delta_{j,t}$ capture *origin-time* and *destination-time* fixed effects. These fixed effects subsume all country-time varying variables which are usually deemed important for FDI inflows, such as population, economic vitality and laws restricting inward and outward investments. Country-level differences in the use of Facebook that otherwise would have an impact on the SCI are one of the most important factors which are absorbed by these fixed effects. Differences between countries in the tendency to engage in FDI are also absorbed by these fixed effects.

To account for zero FDI between several country pairs, we use a Pseudo Maximum Likelihood (PPML) estimator to evaluate Equation 2. Silva and Tenreyro (2006) show that in comparison

to linear estimation methods (e.g. Ordinary Least Squares) the use of a PPML estimator reduces the bias due to heteroskedasticity. Despite the fact that the underlying data does not follow a poisson distribution, it yields robust results as long as the specification of the conditional mean is correct. Next to that the PPML estimation is able to deal with count data that exhibits a large proportion of zeros.[5] All continuous variables that are not scaled between zero and one enter the model logarithmized. We choose to do so because FDI and SCI follow an approximately log-linear relationship, as plotted in Figure 1.

3 Main Results

3.1 Baseline Estimation

We estimate Equation 2 and show the results in Table 2, which is our baseline estimation. Column (1) only includes origin-country-year and destination-country-year fixed effects. It shows that 72.3% of the variation in Inward FDI is explained by fixed effects, which is in line with other literature (see e.g. Bailey et al., 2021). Column (2) introduces the SCI as a control for social connections between the origin and destination country. It shows an elasticity of FDI with respect to the SCI of 0.54, which is statistically significant on the 1% level. It implies that a 1% increase in social connections leads to a rise in FDI of 0.54%. The SCI also has a noticeable explanatory power: The within R^2 suggests that 13.4% of the variation in FDI that is not already explained by fixed effects can be attributed to social ties.

In the next step, we include a full set of control covariates and display the results in Column (3). Including control variables reduces the size of the social connectedness effect, but it remains statistically significant and economically important. An increase in the SCI by 1% is still associated with an increase in FDI inflows by 0.16%. Interestingly, distance, one of the most important control variables in predicting FDI is not statistically significant when including social connectedness as a control.[6] While it is very time- and cost-intensive to ship goods or to

[5]The estimation of Equation 2 uses the algorithm of Correia et al. (2020) implemented in Stata by the command `ppmlhdfe`.

[6]In line with existing literature, it is highly significant without controlling for social connectedness (see

provide services over large distances, this does not necessarily apply to FDI. Intuitively, FDI is less reliant on transportation routes, but perhaps more reliant on information. Alfaro and Chen (2018) suggest that the elasticity of FDI with respect to distance corresponds to half of the elasticity of trade with respect to distance. Thus, controlling for the flow of information through social networks (and business networks via trade) eliminates the effect of distance on FDI entirely.

Consistent with the literature, most other covariates enter with the expected sign. By far the most important economic covariate is the logarithm of trade volume showing an elasticity of 0.5. Differences in GDP (-0.10) and GDP growth (-0.06) between the origin and destination countries are also statistically significant. The shared religion index (0.85), the political distance (-0.53), and a common legal origin (0.21) also suggest that cultural and political proximity ease FDI flows. As an exception, a common border does not seem to be a promoting factor for investment in our setting. While the trade literature suggests a positive relationship of being neighbors, the results may be different for FDI, especially when controlling for social connectedness and trade.

– Table 2 around here –

To further highlight the point that social connectedness is one of the most important factors in explaining FDI, Table A3 inserts each control variable one after another into the regression. We can then get an impression of the importance of each factor by looking at the R^2, i.e. the explanatory power that the variable adds to the model. SCI enters with the second largest explanatory power of all variables (added R^2 of 0.037), right behind trade volume (0.044), but ahead of distance (0.033).

Overall, the baseline results clearly show that social connectedness is significantly positively correlated with the FDI flows between countries. Despite controlling for a large set of fixed effects and a battery of gravity, economic and cultural covariates, the elasticity of FDI with

Table A3).

respect to the SCI is 0.16. That means a doubling of social ties between country i and j leads to an increase in FDI by 16%. In light of the fact that social connections between countries have not been investigated as one determinant of cross-border investments so far, this effect is large and speaks to the importance of social connectedness in gathering information about potential foreign direct investment.

3.2 Instrumental Variables

While we include a battery of known determinants of FDI to minimize omitted variable bias, there is also a potential reverse causality concern. The SCI data measures the probability of cross-border Facebook friendships in August 2020 and is used to explain FDI flows from 2010 to 2019. It is possible that cross-border friendships were established because an investment had been made. Hence, FDI would affect the SCI and not vice versa which would cause an endogeneity problem. Bailey et al. (2021) already provide a qualitative discussion of these concerns regarding the relationship of social ties and trade. They argue that by size alone reverse causality can be ruled out. Connections following a trade contract would need to account for an implausible large increase in connections on Facebook as the number of people involved in trade is much smaller than the number of Facebook users. This same argument applies similarly to our analysis regarding FDI.

Nevertheless, we complement this qualitative discussion with an instrumental variable approach, using a control function. See section A1 for details on the implementation. Finding an instrument for social connections is difficult because there is a broad range of reasons for cross-border friendships on Facebook. For example, many of these friendships can be traced back to single historical events or factors that are unique for single countries or country pairs.[7] The challenge is to find an instrument that influences all countries to a sufficient degree and is therefore valid for the whole data set. We suggest two two possible instruments in the following subsections before we present the results.

[7]For example, Bailey et al. (2021) report a high social connectedness between Australia and Denmark, stemming from the marriage of the Danish crown prince to his Australian-born wife.

First, we use historic migration data from the Global Bilateral Migration Database of the World Bank as an instrument. The data is available for consecutive five census rounds, spanning from 1960 to 2000 (Özden et al., 2011). To capture both immigration and emigration, we sum up the number of people that have entered country i from country j and vice versa to get a bilateral migration matrix, representing the total flow of migration. As the data represents the stock of immigrated people in the respective year in which the census took place, we compute the average over all five available census rounds. We argue that migratory movements have an influence on today's probability of cross-border friendships since in many cases migrants maintain close and persistent connections to their home countries (Waldinger, 2015; Bailey et al., 2021).[8] To be a valid instrument for the SCI, migration needs not to have an influence on FDI decisions, except through their influence on the SCI. Our main argument relies on the time dimension of the census data. Since we use average data between 1960 and 2000, migration and FDI are 30 years apart on average. A significant relationship between people changing their usual country of residence and the FDI decision is therefore very unlikely. We provide a more detailed discussion of the potential direct effects of migration on FDI in section A2 of the appendix.

As a second instrument we include genetic distance from Spolaore and Wacziarg (2018). Genetic distance is a measure that builds on the time elapsed between a destination and origin country of investment having last common ancestors. Sharing common ancestors from long-past time periods is unlikely to directly influence investment relationships today via channels other than social contacts. More importantly, FDI is very unlikely to alter genetic patterns across the globe in the distant past. As a result, genetic ancestry is likely to be a good instrument for the effect of social connections on FDI.

Column (4) of Table 2 reports the results of estimating Equation 5 using historic migration as an instrument. The results of the first stage regression can be found in Table A5. The instrument exhibits a strong first stage (F-value > 800), which confirms the assumption that the relevance condition holds. Using migration as an instrument reveals an elasticity of 0.25, which is slightly

[8]Evidence for that is also given by the increasing number of remittances over the past years that are sent from migrants to home countries (Ratha et al., 2019).

larger compared to 0.16 in the baseline result, but very similar when subtracting the residual component -0.12.[9] Column (5) displays the results using genetic distance as the instrument. Again, we observe a highly significant first stage (F-value > 600). We observe a coefficient of 0.52 and a coefficient for the residuals of -0.38. In total that amounts to 0.14, which is close to the baseline coefficient of 0.16.

All specifications, using migration and genetic distance as instruments, are robust to an OLS approach which is presented in Table A6 in the appendix. In addition, Tables A7 and A8 provide evidence for the fact that the results are not exceptional to our approach but can be replicated with different other sets of migration census data. The use of a control function approach with different IVs allows some conclusions while calling for caution in interpretation in some points. All of them indicate that using the SCI of 2020 there is a causal effect of social connectedness on FDI, given that the exclusion restriction is plausible.[10] However, the design of the control function approach allows to disentangle the influence of social ties on FDI into a part that is explained by the instrument and a residual. The latter can be interpreted as the bias or endogeneity carried by the SCI. Since both residuals are not statistically different from zero, this could be taken as an indication that the instruments perform well in explaining the influence of social connectedness on FDI.

3.3 Robustness

A growing source of uncertainty in FDI data is the increasing use of tax routing and special-purpose-enterprises (letter-box companies) causing "phantom FDI". In those cases the origin or destination of the FDI is not the ultimate designation which disturbs our analysis. Damgaard et al. (2019) identify three such countries which are part of our sample: The Netherlands, Hong Kong and Singapore. We remove these countries and present the results of the regression on the adjusted sample in Table A9. The results are very similar to the original estimates, the elasticity of the SCI regarding FDI yields 0.13 which is only slightly lower compared to the

[9] See section A1 for a detailed discussion on the role of the residuals in the implementation.
[10] See section A2 for a detailed discussion.

baseline regression.

To show that the baseline result is robust to other alterations, Table A10 presents an approach using OLS. Since Figure 1 suggests a logarithmic relationship of the SCI and FDI, the dependent variable enters the regression in logarithms. That leads to a loss of all country pairs showing zero FDI, i.e. it is only an investigation of the intensive margin. The results show an even larger effect of social connectedness on FDI. To avoid the drop of all observations exhibiting zero FDI, we include a specification where we add one to every observation before taking the log. The results are presented in Columns (4)-(6). The elasticity of FDI with respect to the SCI yields 0.11 and is statistically significant on a 1%-level. This is close to our baseline estimate of 0.16.

Moreover, we use additional data sets to contextualize and compare the impact of the SCI on investment decisions. At first, we use the Coordinated Direct Investment Survey of the IMF (CDIS) to estimate the influence of the SCI on FDI stock data. Table A11 shows a coefficient of 0.29, significant on a 1% level, which is larger than in the baseline regression. It is the largest significant variable in this approach. In a second step we use data of the Coordinated Portfolio Investment Survey of the IMF (CPIS) for 2019. Similar to above the CPIS represents stock data and will most likely date back several years. Next to that, portfolio investment is very different from FDI. It requires much less due diligence and commitment. Not only can it be smaller in size but it also does not necessarily fulfill the criteria of establishing lasting interest and managerial control. While for FDI personal visits are often unavoidable, portfolio investment is often made without the investor being on site. We thus expect the coefficient to be smaller and less significant. The results in Columns (4)-(6) confirm this expectation. The elasticity of the investment regarding the SCI yields 0.17 and is significant on the 10%-level only.

Based on these estimates, the baseline results seem to be very robust. Social connectedness appears to be a significant predictor of FDI, independent of the type of estimation and the type of data set used to arrive at this finding. The results thus reinforce the importance of social

connectedness for investment decisions.

4 How Social Connectedness Affects FDI: The Channels

4.1 Social Connectedness and Information Asymmetries

Information asymmetries and frictions are known as a major inefficiency in markets that influence all kinds of business decisions. For many fields in business and economics, authors find that social connections are able to reduce information asymmetries and to improve decision-making: Bailey et al. (2021) find that social ties have a larger effect on trade for goods with larger information frictions, Rehbein and Rother (2020) show that social ties are more important when banks require more information, Hellmanzik and Schmitz (2017) show that controlling for social proximity reduces the influence of factors that control for information asymmetries in portfolio investments. Regarding institutional investments in private equity funds, Freiburg and Grichnik (2012) present evidence for the fact that social ties reduce information asymmetries.

Greenfield investment vs. M&A: To test how sensitive the influence of the SCI is to informational frictions, we exploit variation in the fraction of different kinds of investments in the sample. FDI can be split up into greenfield investments and investments that consist of mergers and acquisitions (M&A). Both types require different kinds of information. Greenfield FDI is an investment that builds an enterprise, subsidiary or branch in a foreign country from scratch. To do so it is necessary to gather information on how to set up a business and on all steps that are required. An M&A project is the purchase or takeover of an already existing business infrastructure. Information on the performance, balance sheet, accounting conditions, future revenues, and earnings have to be evaluated. Compared to that, the information needed for greenfield investments is typically less standardized and differs massively between countries. While M&A deals are relatively standardized and often accompanied by law firms, investment banks, or consulting agencies, greenfield investments typically follow a less uniform process. It is therefore subject to much higher information frictions. If there is a channel through which

the SCI promotes FDI by limiting these frictions, the SCI should show a greater effect for the greenfield rather than M&A FDI.

From 2010 to 2019 the ratio of these two kinds of investments differed significantly. Figure 2 plots the value of greenfield and M&A projects. Both types of investments show an upward trend over the whole period. The value of greenfield projects has fluctuated between 630 and 982 billion USD with an average of almost 800 billion USD. Compared to M&A, the development is relatively smooth. The value of M&A projects fluctuates significantly more and we identify the period from 2015 to 2018 as periods of high values for M&A deals compared to greenfield investments. In these years the share of M&A projects accounts for at least 45% of all FDI. The share for all remaining years vary between 25% and 39%. Therefore, we construct a dummy indicating for a low M&A share from 2010 to 2014 and for 2019 and a dummy indicating for high M&A share from 2015 to 2018.

Table 3 presents the results for the low-M&A period in Column (1) and the results for the high-M&A period in Column (2). For comparison, Column (3) replicates the baseline result. As expected, the SCI plays a more important role for FDI in the low-M&A period. The coefficient of 0.24 is significant at the 1% level. This reflects a significant increase compared to the baseline result. In contrast to that, the elasticity of FDI regarding the SCI is only 0.07 and statistically insignificant for the high-M&A period. At the same time, FDI in the latter period appears to be more responsive to hard information associated with the economic conditions of the two countries involved, such as GDP and GDP growth.

– Table 3 around here –

In line with the theory of limiting information frictions, greenfield FDI is more dependent on social connections than M&A deals. It underlines the assumption that social ties are able to better support the non-standardized processes related to greenfield investments than to help in negotiating M&A-deals. This result is also in line with papers arguing that M&A-deals

sometimes reflect more opportunistic investment, which may not require as much information about the destination country (Baker et al., 2009; Davies et al., 2018).

Economic and Cultural Proximity: The need for information might also be higher if fewer business relationships and cultural similarities exists between countries. To investigate the effect of economic and cultural proximity, we proceed by interacting social connectedness with the relevant variables in the following manner:

$$\begin{aligned} \mathrm{FDI}_{i,j,t} = \exp[&\beta_1 \cdot \log(\mathrm{SCI}_{i,j}) + \beta_2 \cdot \log(\mathrm{SCI}_{i,j}) \cdot \mathrm{Z}_{i,t} + \beta_3 \cdot \log(\mathrm{SCI}_{i,j}) \cdot \mathrm{Z}_{j,t} \\ &+ \beta_4 \cdot \mathrm{G}_{i,j,t} + \beta_5 \cdot \mathrm{A}_{i,j,t} + \delta_{i,t} + \delta_{j,t}] \cdot \epsilon_{i,j,t} \end{aligned} \tag{3}$$

This equation adds the interaction terms between the SCI and the control variables $Z_{i,t}$ and $Z_{j,t}$ to the baseline regression 2. As in the baseline, $G_{i,j,t}$ and $A_{i,j,t}$ represent vectors including all gravity and additional covariates. In the following sections we present different sets of measures for $Z_{i,t}$ and $Z_{j,t}$. For completeness, we include the interaction term for both the origin and destination country of FDI, i and j.

We start by investigating the importance of social connectedness if fewer business relationships exist between countries. We thus interact trade quantile dummies with the SCI. Trade intensity roughly reflects existing business contacts across borders and therefore transmitting information that is also important for FDI decisions. The results presented in Table 4, Column (1) show that the SCI reveals a coefficient of 0.16, similar to the baseline result. However, this elasticity almost doubles for observations in the lowest quantile of trade relationships, as the elasticity rises by 0.14. We complement the analysis by interacting the SCI with the dummy for a regional trade agreement. We again find significant effects. Elasticity yields 0.27 when there is no agreement in place compared to 0.08 when such an agreement exists. Both specifications show that especially in the absence of existing business contracts, social connectedness is able

to provide important information for FDI decisions.

– Table 4 around here –

We next extend the analysis by interacting the SCI with three cultural covariates: common official language, common legal origin, and political distance. Columns (3) to (5) present the results. Especially a common official language provides information flows and decreases the importance of social connections for FDI. Less intuitive but also strong is the role of a common legal origin. It shapes a country's way to organize lawmaking. If a country pair can look back on a similar legal origin, information flows easier and the SCI has a smaller effect on FDI. The same holds true for political distance. The farther apart two countries are in terms of their political system, the more important social connections become. Overall the results clearly point to the relevance of social connections in overcoming information frictions in foreign direct investment. By exploiting differences between the two main types of FDI, we show that social connections are able to overcome information frictions that hamper cross-border investments.

4.2 Overcoming Bad Institutions with Social Connections

The idea that institutions are a relevant factor in influencing FDI was formulated theoretically by Dunning in 1988 (Dunning, 1988). Since then, numerous articles have presented evidence for the supporting role of institutional quality in FDI (Ali et al., 2010; Buchanan et al., 2012; Daude and Stein, 2007). Besides that, institutions are known to facilitate growth.[11] One of the mechanisms by which this is achieved is the creation of an institutional framework that promotes investment decisions, for example through the enforcement of contracts. However, when institutional measures of contract enforcement are absent, contracts may be alternatively enforced through social mechanisms (Chandrasekhar et al., 2018). To test whether social connections can substitute for missing or bad institutions and promote contract enforcement, we interact unilateral controls for the quality of institutions with the SCI. We gather a number of

[11] see Guiso et al. (2008); Economides and Egger (2009) for an overview of the relevant literature.

unilateral control variables to find out to what extent social connections in combination with institutional quality exert their influence on FDI. Definitions and sources of all variables used throughout this section can be found in Table A1, and descriptive statistics are presented in Table A12.

Ease of Doing Business: The CEPII Gravity Data include information on how many procedures and days are required in every country to start a business and how costly this process is. We propose that in countries where institutions and bureaucracies are well developed, starting a business is faster and less costly. Faster and cheaper procedures may be a sign of good contract and property rights enforcement and may thus require fewer social connections.[12] We thus interact the social connectedness index with the ease of doing business variable:

Table 5 presents the results of interacting the ease of doing business variables with the SCI. Column (1) includes interactions with the time needed to start a business, Column (2) with the procedures of starting a business, and Column (3) with the associated costs of doing business. Since all countries require time and procedures greater than zero and almost all of them exhibit costs, the coefficients of the SCI itself are not very meaningful and we, therefore, refrain from interpretation. The important finding of this approach is that all three interaction terms of the SCI with the variables referring to the destination countries are large in size and significant on a 1% level. The opposite applies to the estimates for the origin countries. All of them are not statistically significant and small. Since time and procedures are included in logarithms, their coefficients can again be interpreted as elasticities. FDI shows elasticities of 0.13 with respect to the days and 0.21 regarding the procedures needed to start a business. A one standard deviation change in the costs is associated with an increase of 0.11[13] of the SCI's effect on FDI. This approach shows that the effect of social connectedness is highly sensitive to measures of the resources that are needed to establish a new business. The results suggest that social relationships are able to reduce barriers by facilitating the process of investing abroad,

[12] It may also require less information to be collected before investing, which would make it more in line with an information-asymmetry interpretation from section 4.1.

[13] $0.673 \cdot 0.17 = 0.114$

especially when the hurdles to do so are large in terms of time, procedures, and costs.

– Table 5 around here –

Financial Market Development: We next interact FDI with a variable for financial development in the origin and destination country, taken from the IMF Financial Development Index. This index evaluates the financial market development by measuring its depth, access, and efficiency, combining a large number of variables and indicators (e.g. stock market capitalization to GDP, stocks traded to GDP, or stock market turnover ratio). Financial development of a country proxies for the quality, accessibility, and trustworthiness of hard data. We hypothesize that markets with better-developed financial markets make contract enforcement easier, even for M&A projects. Consequently, we expect social connectedness will play a smaller role in countries with well-developed financial markets.[14]

Column (4) of Table 5 presents the results. The coefficient of SCI without taking the financial market development into account is much higher and statistically significant, yet its importance significantly decreases in financial market development. A one standard deviation increase in financial market development in the destination country reduces the effect of social connectedness on FDI by roughly 30%. Thus, the better-developed financial markets are the less important is social connectedness. Further investigation in Table A13 reveals that this effect is similar for low and high M&A periods.

World Governance Indicators for Institutional Quality: We next draw on data from the Worldwide Governance Indicators (WGI). This index collects data on institutional quality for a large set of countries over the time period from 1996-2021 and groups them into six different dimensions: *Control of Corruption, Rule of Law, Voice and Accountability, Political Stability and Lack of Violence, Government Effectiveness* and *Regulatory Quality.* The first

[14]The heterogeneities with regard to financial market development are also in line with an information asymmetry interpretation as discussed in section 4.1.

two dimensions are related to contract and law enforcement and the degree to which extent people trust and believe in the authorities that are in force to govern societies' interactions. *Voice and Accountability, Political Stability and Lack of Violence* express the determination of authorities and the level of political distress associated with violent conflicts and terrorism. To what extent the governing authorities are able to implement policies and commit themselves to these is captured by *Government Effectiveness*. The indicator also covers infrastructure, bureaucracy and transportation. Precise information on the collection and composition of all dimensions can be found in Kaufmann et al. (2006). They report that the standard normal units, which are the values that we use for the analysis here, range from approximately -2.5 to 2.5, with higher values indicating better quality institutions.

All dimensions of the WGI relate to areas that serve as an important basis for cross-border investments. If social connections are able to substitute for a lack of high institutional quality, we expect the interaction term of the SCI and the respective dimension of the WGI to show a negative influence on FDI. We estimate the influence of four WGI dimensions (Control of Corruption, Rule of Law, Political Stability and Government Effectiveness) by including them for $Z_{i,t}$ and $Z_{j,t}$ in Equation 3. Intuitively the quality of institutions of the destination country should have a larger influence as it has a direct effect on the investment as the institutional hypothesis is based on institutions in the destination country.

Table 6 presents the results. Compared to the baseline result, all four dimensions show a higher elasticity of FDI with respect to the SCI. This is not directly comparable to the baseline coefficient of 0.16, as it expresses the elasticity for the case that the institutional dimensions take a value of zero. Since the mean for all variables of the WGI is positive, a higher elasticity is in line with the baseline result. Next to that, all WGI dimensions report a statistically significant effect on the institutional quality of the destination country. The coefficients are not only statistically significant but also economically meaningful in size and reveal a very large influence of the SCI in combination with institutional quality.

– Table 6 around here –

The first two columns which correspond to the dimensions *Control of Corruption* and *Rule of Law*. The effect for the destination country of FDI shows that a one standard deviation increase in the institutional variable leads to a drop in the elasticity of FDI regarding the SCI of 0.14 and 0.15. This equals a drop of almost 100% of the elasticity from the baseline regression. In other words, the positive effect of social connections on FDI vanishes completely for countries with good institutional scores in these two dimensions. On the other hand, the effect of social connectedness on FDI is also equally stronger in countries, where these dimensions are below average, or even below zero.

Note that for *Control of Corruption* there is also a small negative effect with regards to the originating country. This is an interesting finding as it proves that social ties influence FDI via a channel that is independent of the destination countries' characteristics. We suggest two possible explanations: On the one hand, the initiative can come from investors themselves. Those who deal with corruption in their home country to a greater degree than others may know better how to use social relationships to their advantage. That corruption norms from the home country can affect the actions of people in foreign countries is shown by Fisman and Miguel (2007). On the other hand, the effect could be driven by the destination country's demand for FDI. Since FDI is an important factor for economic development, many countries actively compete for investors from abroad. Social connections might play a more important role in acquiring FDI if corruption is more prevalent in the originating country.

Political Stability (Column (3)) has a similar, but somewhat smaller effect: Investing in politically stable countries does not necessitate strong social connections. High political instability however significantly increases the importance of social connectedness. The effects may be somewhat smaller, due to the fact that political stability might matter only in its extreme expressions. *Government Effectiveness* (Column (4)) also reduces the elasticity of SCI with regard to FDI by around 0.13 in case of the increase by one standard deviation in the institu-

tional measure. This change also largely removes the baseline effect of social connectedness if institutions are good, but also similarly amplifies its importance, when they are bad. We show the results of interactions with all other WGI indicators in Table A14 and interactions from other sources in Table A15, which further confirm our findings.

To illustrate the influence of institutional quality more clearly, we choose two countries that differ in their rule of law and calculate its effect as an example. In 2016, Germany shows a value of 1.62, and Argentina a value of -0.39. Using the estimated elasticity of the SCI only from Column (2) of Table 6, which is 0.27, Germany has an elasticity of 0[15]. This means that, based on the WGI measure of their rule of law, social connections have a very limited influence on FDI inflows to Germany. In the case of Argentina, the picture is different. The elasticity of FDI regarding the SCI in combination with the WGI measure for the rule of law amounts to 0.33[16]. This means that a doubling of social connectedness with Argentina would imply an increase in FDI by 33%.

Individual Rights and Participation: The WGI dimensions focus predominantly on authorities, the government, and the political system. Often institutions also affect individuals' rights and their possibility to actively and freely participate in society. We thus use indicators taken from the Varieties of Democracy Institute to check whether restricted participation in society alters the value of social connectedness for FDI. All indicators vary between zero and one. We focus on the "exclusion of socio-economic groups" in Table 6 to illustrate the overall effect and provide full estimates of all indicators in Tables A15 and A16.

As for the WGI indicators, the interaction of social connectedness and the index of exclusion of socio-economic groups is statistically significant and economically important. Column (5) shows the results. It yields a coefficient of 0.38, statistically significant on a 1-%level. Economically, the effect is also meaningful: An increase of one standard deviation results in an increase of 0.09. Since the mean of 0.24 is almost equal to the standard deviation of 0.23 and the stand-alone

[15] $0.268 + 1.62 \times (-0.160) = 0.008$
[16] $0.268 + (-0.39) \times (-0.160) = 0.33$

effect of the SCI is almost zero in Column (5), FDI has an elasticity of 0.09 for countries at the mean. However, the value of social ties increases drastically, as the elasticity goes up to 0.37 for countries with the maximum value of almost one for the exclusion of socio-economic groups. This result shows, that countries with bad institutions which result in reduced participation in society for socio-economic groups are more reliant on social connections to attract FDI. It suggests that discrimination presents an investment barrier, which makes social connectedness more important. This is an interesting finding, which highlights that discrimination can be inefficient for investment and can cause a reduction in diversification due to few investors having sufficient social ties to overcome this friction.

To sum up, the interaction of the SCI with dimensions measuring institutional quality shows that social connectedness can influence FDI by substituting for good institutions. The effect is very broadly distributed across different types of institutions. Nevertheless, it is evident that social connectedness compensates better for variables that have a direct influence on entrepreneurial action. Corruption plays a special role because it has an influence not only in the destination but also in the origin country of the FDI, although the effect in the latter is much smaller.

4.3 Uncertainty and Social Connectedness

Global events over the last two decades have highlighted the importance of uncertainty in investment and growth (see Cascaldi-Garcia et al. (2020) for an overview). Uncertainty has become the focus of policymakers and researchers alike. More recently, possible measurements of uncertainty have also arisen for the first time in the literature, mostly based on textual analysis of various kinds (Jurado et al., 2015; Segal et al., 2015; Baker et al., 2016; Bekaert et al., 2022). Since uncertainty reflects a barrier to investment which is likely to affect foreign direct investment, we ask the question of whether social connectedness can help to overcome this investment barrier created by increasing uncertainty.

World Uncertainty Index: We begin by interacting social connectedness with the World Uncertainty Index (WUI), which has been shown to measure uncertainty that negatively affects FDI (Choi et al., 2021). The index reflects the relative frequencies of the word "uncertainty" in EIU country reports in each country (Ahir et al., 2022). It also provides an aggregated global index of uncertainty. Compared to other indices, the WUI has the largest and most diverse sample of the recently developed text based measures. Note that in these regressions we estimate fixed-effects OLS instead of our baseline gravity model, because we expect uncertainty to largely affect the intensive, but not the extensive margin of FDI.

We present the results of the interactions in Table 7.

– Table 7 around here –

Column (1) demonstrates the familiar importance of social connectedness for FDI. Social connectedness becomes more important, as the destination country's uncertainty increases. A one standard deviation increase in the WUI increases the effect of log(SCI) on FDI by about 13% of the baseline, statistically significant at the 10% level. The importance of social connections also increase in uncertainty in the originating country, but the interaction is not statistically significant. The results become much stronger when considering high-information need periods (Column (2)). During periods when FDI is mostly dominated by greenfield investment, a one standard deviation increase in uncertainty is associated with an increase in the importance of SCI for FDI by roughly 25%.

But even more important than country-specific uncertainty is global uncertainty. The idea that global uncertainty plays an out-sized role for foreign direct investment similarly arises in Choi et al. (2021). Columns (3) and (4) demonstrate that social connections matter significantly more when globally uncertainty is high, even more so when greenfield investment is predominant. A one standard deviation increase in global uncertainty is connected to an increase in the importance of SCI for FDI by roughly 50% of the baseline effect. The importance of social connections with regard to global uncertainty is also much more pronounced than country-specific

uncertainty. Columns (5) and (6) repeat the exercise using the GDP-weighted global uncertainty index and arrive at similar findings. A one standard deviation increase in uncertainty is related to an increase in the importance of SCI for FDI which is similar in high-information need periods. We also confirm these results using the Geopolitical Risk index produced by Caldara and Iacoviello (2022) in Table A17.

Stock Market Uncertainty: Next, we turn to uncertainty on international financial markets and uncover an interesting pattern with regard to the importance of social connections for FDI. Financial market uncertainty is often measured using volatility indices (Kim and Kung, 2017). We thus interact the SCI with the VIX index (in logs), to see if social connectedness is more important when financial market uncertainty is high. We display the results in Table 8.

– Table 8 around here –

The results are initially surprising. In Column (1), we estimate the interaction on the full sample. Higher volatility as measured by the VIX is associated with a decrease in importance of SCI for FDI. The net effect is still positive: Social connectedness matters for FDI independent of the uncertainty in financial markets, but it does not help to overcome financial market uncertainty. The results become more intuitive when decomposing the sample into low M&A and high M&A periods. In low M&A periods, the interaction of SCI and log(VIX) is significantly negative (Column(2)). We suggest that when greenfield investment is prominent, social connections can contain a negative investment signal for investing through financial markets, leading to an overall negative effect. Put differently, in high-uncertainty financial environments investors benefit from social connections, by refraining to invest in the current period, perhaps by shifting to greenfield investment which might take more time to realize. Yet, when financial markets are needed, i.e. when M&A's are prominent, Column (3) suggests that social connections can help to overcome financial market uncertainty, as the interaction is positive in high M&A periods.

Summing up, social connectedness can play an important role in overcoming global and country-specific uncertainty. This has very important implications should levels of uncertainty continue to rise globally. In this case, establishing inter-country connections will become increasingly important, both for governments and private industries. For example, our evidence suggests that policies that increase international interactions, such as (temporary) migration or even short-term exchange programs may have increasingly large returns. It also suggests that geographically diverse teams might become increasingly important due to their large pool of international social connections.

5 Conclusion

This paper shows that social connectedness, proxied for by Facebook's Social Connectedness Index, is an important determinant of foreign direct investment. We estimate a gravity equation with a large set of economic and cultural covariates, classic gravity controls and fixed effects using a panel data set of FDI flows by the UNCTAD. The main result shows a statistically an economically significant effect of social connections on FDI. A 1% higher social connectedness is associated with an increase in FDI by 0.16%, and statistically eliminates previously important determinants such as distance between countries. To improve identification we use bilateral migration from 1960 to 2000 and genetic distance as instruments. This approach confirms the original finding and suggest a causal effect on social connections on FDI. Our results are in line with an emerging literature highlighting the economic effects of social relationships, such as bank lending and trade and suggest that social connections are a highly important determinant in foreign direct investment.

We highlight three channels through which social connections enables FDI: Overcoming high information requirements, overcoming bad institutions and overcoming global uncertainty. We show first that social connectedness is most important for high-information-need investments, such as greenfield investments and where circumstances of language and culture make information acquisition more difficult. We then show that social connections mitigate the hampering

effect of bad institutions like corruption and bad contract enforcement. Importantly, social connections also overcome the prohibiting effect of institutions that limit individuals' rights such as the discrimination of socio-economic groups. Lastly we demonstrate that social connectedness helps to overcome investment barriers posed by increased global and country-specific uncertainty.

Our findings are not only important to better understand the cross-country determinants of FDI, but are also highly valuable for policy attempting to promote FDI. While many other key determinants like religion or legal origins are almost impossible to change, it is comparably easy to establish and promote social contacts across countries. Our results suggest that there might be significant economic value in promoting cross-country social exchange, especially in a world where global uncertainty is on the rise.

References

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5): 1369–1401.

Agrawal, A., Kapur, D., and McHale, J. (2008). How do spatial and social proximity influence knowledge flows? Evidence from patent data. *Journal of Urban Economics*, 64(2): 258–269.

Ahearne, A. G., Griever, W. L., and Warnock, F. E. (2004). Information costs and home bias: An analysis of us holdings of foreign equities. *Journal of International Economics*, 62(2): 313–336.

Ahir, H., Bloom, N., and Furceri, D. (2022). The world uncertainty index. Technical report, National Bureau of Economic Research.

Alfaro, L. and Chen, M. X. (2018). Transportation cost and the geography of foreign investment. In *Handbook of International Trade and Transportation*. Edward Elgar Publishing.

Ali, F. A., Fiess, N., and MacDonald, R. (2010). Do institutions matter for foreign direct investment? *Open Economies Review*, 21(2): 201–219.

Allcott, H., Braghieri, L., Eichmeyer, S., and Gentzkow, M. (2020). The welfare effects of social media. *American Economic Review*, 110(3): 629–76.

Anderson, J. E. and Marcouiller, D. (2002). Insecurity and the Pattern of Trade: An Empirical Investigation. *The Review of Economics and Statistics*, 84(2): 342–352.

Avom, D., Njangang, H., Nawo, L., et al. (2020). World economic policy uncertainty and foreign direct investment. *Economics Bulletin*, 40(2): 1457–1464.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., and Wong, A. (2018). Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives*, 32(3): 259–280.

Bailey, M., Dávila, E., Kuchler, T., and Stroebel, J. (2019). House price beliefs and mortgage leverage choice. *The Review of Economic Studies*, 86(6): 2403–2452.

Bailey, M., Gupta, A., Hillenbrand, S., Kuchler, T., Richmond, R., and Stroebel, J. (2021). International trade and social connectedness. *Journal of International Economics*, 129: 103418.

Baker, M., Foley, C. F., and Wurgler, J. (2009). Multinationals as Arbitrageurs: The Effect of Stock Market Valuations on Foreign Direct Investment. *The Review of Financial Studies*, 22(1): 337–369.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4): 1593–1636.

Balakina, O., Bäckman, C., Hackethal, A., Hanspal, T., and Lammer, D. M. (2022). Good peers, good apples? Peer effects in portfolio quality. Technical report, Leibniz Institute for Financial Research SAFE.

Bekaert, G., Engstrom, E. C., and Xu, N. R. (2022). The time variation in risk appetite and uncertainty. *Management Science*, 68(6): 3975–4004.

Berger, A., Busse, M., Nunnenkamp, P., and Roy, M. (2013). Do trade and investment agreements lead to more FDI? Accounting for key provisions inside the black box. *International Economics and Economic Policy*, 10(2): 247–275.

Bergstrand, J. H. and Egger, P. (2007). A knowledge-and-physical-capital model of international trade flows, foreign direct investment, and multinational enterprises. *Journal of International Economics*, 73(2): 278–308.

Berkowitz, D., Moenius, J., and Pistor, K. (2006). Trade, Law, and Product Complexity. *The Review of Economics and Statistics*, 88(2): 363–373.

Biswas, R. (2002). Determinants of Foreign Direct Investment. *Review of Development Economics*, 6(3): 492–504.

Blonigen, B. A. and Piger, J. (2014). Determinants of foreign direct investment. *Canadian Journal of Economics*, 47(3): 775–812.

Brown, J. R., Ivković, Z., Smith, P. A., and Weisbenner, S. (2008). Neighbors matter: Causal community effects and stock market participation. *The Journal of Finance*, 63(3): 1509–1531.

Buchanan, B. G., Le, Q. V., and Rishi, M. (2012). Foreign direct investment and institutional quality: Some empirical evidence. *International Review of Financial Analysis*, 21: 81–89.

Burchardi, K. B., Chaney, T., and Hassan, T. A. (2019). Migrants, ancestors, and foreign investments. *The Review of Economic Studies*, 86(4): 1448–1486.

Büthe, T. and Milner, H. V. (2008). The politics of foreign direct investment into developing countries: Increasing FDI through international trade agreements? *American Journal of Political Science*, 52(4): 741–762.

Caldara, D. and Iacoviello, M. (2022). Measuring geopolitical risk. *American Economic Review*, 112(4): 1194–1225.

Cascaldi-Garcia, D., Sarisoy, C., Londono, J. M., Rogers, J. H., Datta, D., RT Ferreira, T., Grishchenko, O. V., Jahan-Parvar, M. R., Loria, F., Ma, S., et al. (2020). What is certain about uncertainty? *Journal of Economic Literature, forthcoming*.

Chandrasekhar, A. G., Kinnan, C., and Larreguy, H. (2018). Social networks as contract enforcement: Evidence from a lab experiment in the field. *American Economic Journal: Applied Economics*, 10(4): 43–78.

Choi, S., Furceri, D., and Yoon, C. (2021). Policy uncertainty and foreign direct investment. *Review of International Economics*, 29(2): 195–227.

Combes, P.-P., Lafourcade, M., and Mayer, T. (2005). The trade-creating effects of business and social networks: Evidence from France. *Journal of International Economics*, 66(1): 1–29.

Correia, S., Guimarães, P., and Zylkin, T. (2020). Fast poisson estimation with high-dimensional fixed effects. *The Stata Journal*, 20(1): 95–115.

Coval, J. D. and Moskowitz, T. J. (1999). Home bias at home: Local equity preference in domestic portfolios. *The Journal of Finance*, 54(6): 2045–2073.

Cuadros, A., Martín-Montaner, J., and Paniagua, J. (2019). Migration and fdi: The role of job skills. *International Review of Economics & Finance*, 59: 318–332.

Damgaard, J., Elkjaer, T., and Johannesen, N. (2019). *What is real and what is not in the global FDI network?* International Monetary Fund.

Daude, C. and Fratzscher, M. (2008). The pecking order of cross-border investment. *Journal of International Economics*, 74(1): 94–119.

Daude, C. and Stein, E. (2007). The quality of institutions and foreign direct investment. *Economics & Politics*, 19(3): 317–344.

Davies, R. B., Desbordes, R., and Ray, A. (2018). Greenfield versus merger and acquisition fdi: Same wine, different bottles? *Canadian Journal of Economics*, 51(4): 1151–1190.

Dunning, J. H. (1988). The eclectic paradigm of international production: A restatement and some possible extensions. *Journal of International Business Studies*, 19(1): 1–31.

Economides, G. and Egger, P. H. (2009). The role of institutions in economic outcomes: editorial introduction. *European Journal of Political Economy*, 3(25): 277–279.

Egger, P. and Pfaffermayr, M. (2004). Distance, trade and FDI: A hausman–taylor sur approach. *Journal of Applied Econometrics*, 19(2): 227–246.

Facebook, I. (2020). Facebook reports third quarter 2020 results. `https://s21.q4cdn.com/399680738/files/doc_financials/2020/q3/FB-09.30.2020-Exhibit-99.1.pdf`.

Fisman, R. and Miguel, E. (2007). Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political Economy*, 115(6): 1020–1048.

Freiburg, M. and Grichnik, D. (2012). Institutional investments in private equity funds: Social ties and the reduction of information asymmetry. *Venture Capital*, 14(1): 1–26.

Georgarakos, D., Haliassos, M., and Pasini, G. (2014). Household debt and social interactions. *The Review of Financial Studies*, 27(5): 1404–1433.

Greif, A. (1989). Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders. *The Journal of Economic History*, 49(4): 857–882.

Guiso, L., Sapienza, P., and Zingales, L. (2008). Social capital as good culture. *Journal of the European Economic Association*, 6(2-3): 295–320.

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

Hellmanzik, C. and Schmitz, M. (2017). Taking gravity online: The role of virtual proximity in international finance. *Journal of International Money and Finance*, 77: 164–179.

Hong, H., Kubik, J. D., and Stein, J. C. (2005). Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. *The Journal of Finance*, 60(6): 2801–2824.

Hsieh, H.-C., Boarelli, S., and Vu, T. H. C. (2019). The effects of economic policy uncertainty on outward foreign direct investment. *International Review of Economics & Finance*, 64: 377–392.

Jardet, C., Jude, C., and Chinn, M. D. (2022). Foreign direct investment under uncertainty: evidence from a large panel of countries. Technical report, National Bureau of Economic Research.

Javorcik, B. S., Özden, Ç., Spatareanu, M., and Neagu, C. (2011). Migrant networks and foreign direct investment. *Journal of Development Economics*, 94(2): 231–241.

Julio, B. and Yook, Y. (2016). Policy uncertainty, irreversibility, and cross-border flows of capital. *Journal of International Economics*, 103: 13–26.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring Uncertainty. *American Economic Review*, 105(3): 1177–1216.

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2006). Governance matters v: Aggregate and individual governance indicators for 1996-2005. World Bank Policy Research Working Paper 4012.

Kim, H. and Kung, H. (2017). The asset redeployability channel: How uncertainty affects corporate investment. *The Review of Financial Studies*, 30(1): 245–280.

Kox, H. L. and Rojas-Romagosa, H. (2020). How trade and investment agreements affect bilateral foreign direct investment: Results from a structural gravity model. *The World Economy*, 43(12): 3203–3242.

Kuchler, T., Li, Y., Peng, L., Stroebel, J., and Zhou, D. (2022a). Social proximity to capital: Implications for investors and firms. *The Review of Financial Studies*, 35(6): 2743–2789.

Kuchler, T., Russel, D., and Stroebel, J. (2022b). JUE insight: The geographic spread of COVID-19 correlates with the structure of social networks as measured by facebook. *Journal of Urban Economics*, 127:103314.

Kuchler, T. and Stroebel, J. (2021). Social finance. *Annual Review of Financial Economics*, 13: 37–55.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. (1999). The quality of government. *The Journal of Law, Economics, and Organization*, 15(1): 222–279.

Leblang, D. (2010). Familiarity breeds investment: Diaspora networks and international investment. *American Political Science Review*, 104(3): 584–600.

Lin, W. and Wooldridge, J. M. (2019). Testing and correcting for endogeneity in nonlinear unobserved effects models. In *Panel Data Econometrics*, pages 21–43. Elsevier.

Ma, D., Zhang, C., Hui, Y., and Xu, B. (2022). Economic uncertainty spillover and social networks. *Journal of Business Research*, 145: 454–467.

Nguyen, C. P. and Lee, G. S. (2021). Uncertainty, financial development, and fdi inflows: Global evidence. *Economic Modelling*, 99: 105473.

Nunn, N. (2007). Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade. *The Quarterly Journal of Economics*, 122(2): 569–600.

Okawa, Y. and Van Wincoop, E. (2012). Gravity in international finance. *Journal of International Economics*, 87(2): 205–215.

Ouimet, P. and Tate, G. (2020). Learning from coworkers: Peer effects on individual investment decisions. *The Journal of Finance*, 75(1): 133–172.

Özden, Ç., Parsons, C. R., Schiff, M., and Walmsley, T. L. (2011). Where on earth is everybody? The evolution of global bilateral migration 1960–2000. *The World Bank Economic Review*, 25(1): 12–56.

Paniagua, J., Korzynski, P., and Mas-Tur, A. (2017). Crossing borders with social media: Online social networks and FDI. *European Management Journal*, 35(3): 314–326.

Paul, J. and Feliciano-Cestero, M. M. (2021). Five decades of research on foreign direct investment by MNEs: An overview and research agenda. *Journal of Business Research*, 124: 800–812.

Pool, V. K., Stoffman, N., and Yonker, S. E. (2015). The people in your neighborhood: Social interactions and mutual fund portfolios. *The Journal of Finance*, 70(6): 2679–2732.

Portes, R. and Rey, H. (2005). The determinants of cross-border equity flows. *Journal of International Economics*, 65(2): 269–296.

Ranjan, P. and Lee, J. Y. (2007). Contract enforcement and international trade. *Economics & Politics*, 19(2): 191–218.

Ratha, D. K., De, S., Kim, E. J., Plaza, S., Seshan, G. K., Shaw, W., and Yameogo, N. D. (2019). Leveraging economic migration for development: A briefing for the World Bank board. Technical report, The World Bank.

Rauch, J. E. (2001). Business and Social Networks in International Trade. *Journal of Economic Literature*, 39(4): 1177–1203.

Rehbein, O. and Rother, S. (2020). The role of social networks in bank lending. ECONtribute Discussion Paper 033, Bonn and Cologne.

Rodrik, D. (2000). Institutions for high-quality growth: What they are and how to acquire them. *Studies in Comparative International Development*, 35(3): 3–31.

Rodrik, D., Subramanian, A., and Trebbi, F. (2004). Institutions rule: the primacy of institutions over geography and integration in economic development. *Journal of economic growth*, 9(2): 131–165.

Schiavo, S. (2007). Common currencies and FDI flows. *Oxford Economic Papers*, 59(3): 536–560.

Segal, G., Shaliastovich, I., and Yaron, A. (2015). Good and bad uncertainty: Macroeconomic and financial market implications. *Journal of Financial Economics*, 117(2): 369–397.

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

Spolaore, E. and Wacziarg, R. (2018). Ancestry and development: New evidence. *Journal of Applied Econometrics*, 33(5): 748–762.

UNCTAD (2019). *World Investment Report 2019: Special Economic Zones*. United Nations.

UNCTAD (2020). *World Investment Report 2020: International production beyond the pandemic*. United Nations.

Waldinger, R. (2015). *The cross-border connection: Immigrants, emigrants, and their homelands*. Harvard University Press.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, 50(2): 420–445.

Figures

Figure 1: Foreign Direct Investment vs. Social Connectedness

This figure shows binscatter plots of Inward FDI and the Social Connectedness Index (SCI). Panel A regresses the logarithm of FDI on the logarithm of SCI, while Panel B shows the same specification including all gravity variables as control variables. Both panels incorporate country and year fixed effects. By taking the logarithm of FDI all country pairs with zero FDI are dropped. The remaining data set contains 26,041 observations and represents the intensive margin of FDI.

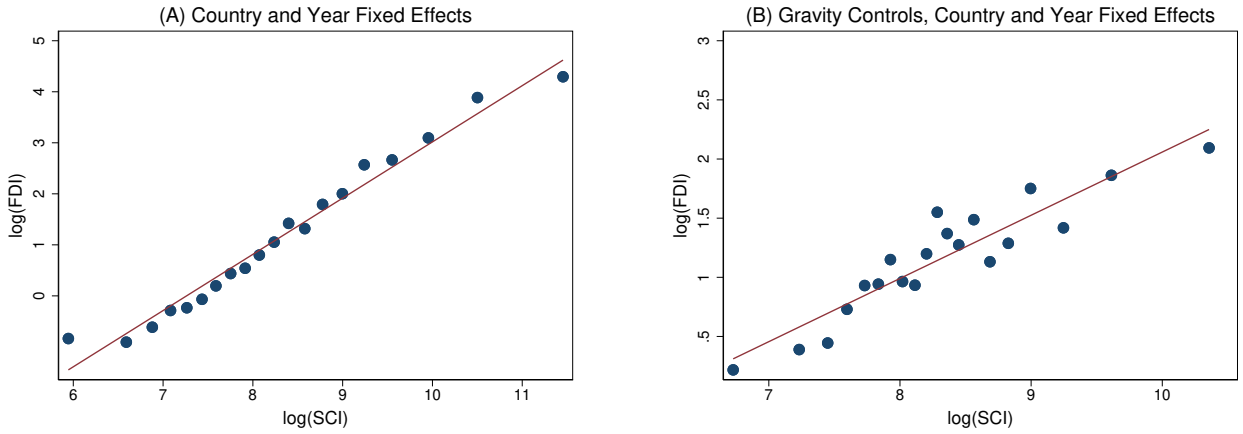
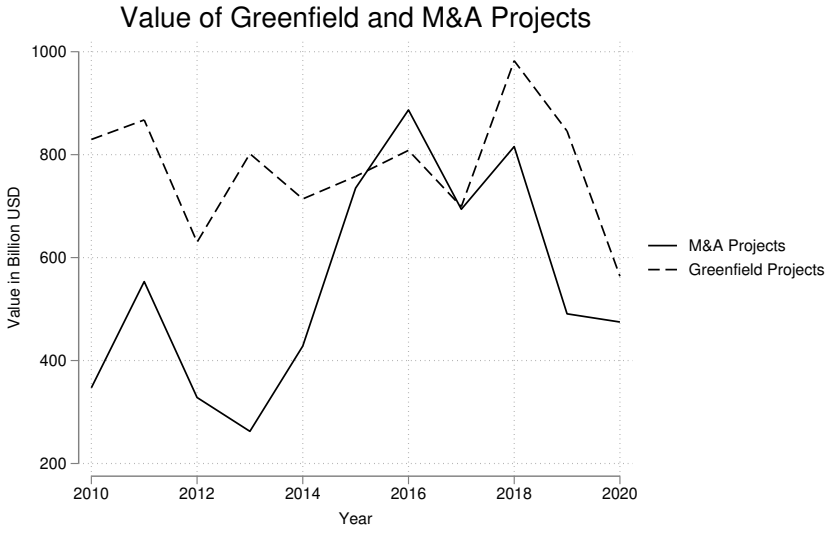


Figure 2: Greenfield Projects vs. M&A Projects

This figure illustrates the value of greenfield and M&A-projects from 2010 to 2020 in billion USD. Data is taken from Table 9 and Table 15 of the Annex of the UNCTAD World Investment Report 2020. The dashed line presents the greenfield projects while the solid line refers to M&A deals.



Tables

Table 1: Descriptive Statistics

This table displays descriptive statistics for the main data set. All variables are at the country pair level. For example, *Inward FDI in million USD* captures FDI of investors from the origin countries to the destination countries. With the exception of Inward FDI in million USD, independent variables are symmetric, i.e. show the same value for the country pair (i, j) and (j, i). The data includes 151 origin countries and 109 destination countries as an unbalanced panel from 2010 until 2019 with a total number of country pairs equal to 5,904. A detailed description of all variables can be found in Table A1.

	Mean	Median	SD	Min	Max	Obs
Inward FDI in Million USD	309.37	0.78	2,298.69	0.00	125,256.92	40,343
log(SCI)	8.26	8.11	1.63	4.19	12.77	40,343
log(Distance)	8.27	8.47	0.99	4.88	9.89	40,343
Common Border	0.14	0.00	0.35	0.00	1.00	40,343
Common Official Language	0.12	0.00	0.33	0.00	1.00	40,343
Common Colonizer	0.05	0.00	0.21	0.00	1.00	40,343
Colonial Relationship after 1945	0.02	0.00	0.12	0.00	1.00	40,343
log(Trade Volume)	19.66	19.84	2.68	5.78	27.11	40,343
log(GDP Difference)	26.81	26.87	1.88	14.39	30.61	40,343
log(GDP Growth Difference)	0.62	0.84	1.16	-8.51	3.42	40,343
Regional Trade Agreement	0.42	0.00	0.49	0.00	1.00	40,343
Common Currency	0.06	0.00	0.24	0.00	1.00	40,343
Shared Religion Index	0.44	0.49	0.30	0.00	0.99	40,343
Political Distance	0.19	0.11	0.25	0.00	1.00	40,343
Common Legal Origin	0.37	0.00	0.48	0.00	1.00	40,343
Common Ethno. Language	0.14	0.00	0.35	0.00	1.00	40,343
log(Avg. Migration 1960-2000) (WB)	6.11	6.17	3.01	0.00	14.45	39,441
Genetic Distance	0.02	0.02	0.02	0.00	0.08	37,429
log(Dest. Days to Start a Business)	2.40	2.40	0.87	-0.69	5.16	40,305
log(Orig. Days to Start a Business)	2.37	2.40	0.94	-0.69	6.54	40,248
log(Dest. Procedures to Start a Business)	1.80	1.79	0.46	0.00	2.77	40,305
log(Orig. Procedures to Start a Business)	1.77	1.79	0.51	0.00	2.89	40,248
Dest. Costs to Start a Business (GNI share)	0.10	0.05	0.17	0.00	1.58	40,305
Orig. Costs to Start a Business (GNI share)	0.13	0.05	0.26	0.00	7.35	40,248
Dest. Financial Market Development	0.37	0.33	0.30	0.00	0.97	40,343
Orig. Financial Market Development	0.42	0.43	0.30	0.00	0.97	40,255
Dest. Control of Corruption	0.38	0.16	0.99	-1.67	2.41	40,343
Orig. Control of Corruption	0.54	0.37	1.09	-1.82	2.41	40,343
Dest. Rule of Law	0.47	0.42	0.94	-1.59	2.13	40,343
Orig. Rule of Law	0.57	0.51	1.01	-1.85	2.13	40,343
Dest. Political Stability	0.16	0.36	0.83	-2.81	1.64	40,343
Orig. Political Stability	0.20	0.39	0.87	-2.81	1.64	40,343
Dest. Government Effectiveness	0.56	0.52	0.85	-1.64	2.24	40,343
Orig. Government Effectiveness	0.63	0.59	0.96	-2.08	2.24	40,343
Dest. Exclusion of Socio-economic Groups	0.24	0.14	0.23	0.01	0.96	40,343
Orig. Exclusion of Socio-Economic Groups	0.25	0.16	0.23	0.01	0.96	40,343
Dest. WUI	0.22	0.21	0.13	0.03	0.50	37,871
Orig. WUI	0.22	0.20	0.14	0.03	0.54	37,713
Global WUI	0.22	0.23	0.04	0.15	0.28	40,343
Global GDP-Weighted WUI	0.24	0.23	0.06	0.17	0.41	40,343
log(VIX Index)	2.79	2.76	0.21	2.41	3.19	40,343

Table 2: Baseline Result

This table shows the results from estimating regression 2 using a PPML estimator. The dependent variable is Inward FDI from country i to country j in year t. The main explanatory variable is the logarithm of SCI. The model controls for gravity covariates (log of distance, a common border dummy, a common official language dummy, a dummy for a common colonizer post 1945, and a dummy for a colonial relationship post 1945), economic covariates (log of the total trade volume, log of the GDP difference, log of the GDP growth rate difference, a dummy for a regional trade agreement in place and a dummy indicating a shared currency) and cultural covariates (a shared religion index, a measure for political distance, a dummy for a common legal origin and a dummy indicating whether both countries share a common ethnological language). In Column (3) the log of average bilateral migration stocks between 1960 and 2000 from the Wold Bank is used to instrument for the SCI, Column (2) exploits the genetic distance from Spolaore and Wacziarg (2018) as an instrument for the SCI. Both specifications are estimated using a control function approach as presented in Equation 5. The first stage results estimated by Equation 4 are presented in Table A5 in the appendix. The coefficient of the first stage as well as the p-value and F-value for the instruments are displayed at the bottom of this Table. The coefficient of $log(SCI)$ represents the share of influence of the SCI on FDI that is explained by the instrument, while the coefficient of the residuals expresses the bias that the SCI bears. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable			Inward FDI		
Instrumental variable				Migration 1960-2000	Genetic Distance
	(1)	(2)	(3)	(4)	(5)
log(SCI)		0.544***	0.157***	0.251*	0.522**
		(0.040)	(0.053)	(0.139)	(0.262)
Residuals				-0.123	-0.384
				(0.153)	(0.272)
log(Trade Volume)			0.495***	0.465***	0.398***
			(0.068)	(0.083)	(0.085)
log(GDP Difference)			-0.103***	-0.102***	-0.105***
			(0.038)	(0.038)	(0.039)
log(GDP Growth Difference)			-0.064**	-0.063**	-0.067***
			(0.025)	(0.025)	(0.026)
Regional Trade Agreement			0.198	0.192	0.134
			(0.125)	(0.124)	(0.138)
Common Currency			0.041	0.057	0.096
			(0.188)	(0.187)	(0.192)
Shared Religion Index			0.847***	0.767**	0.268
			(0.288)	(0.318)	(0.493)
Political Distance			-0.532**	-0.510*	-0.531**
			(0.263)	(0.268)	(0.267)
Common Legal Origin			0.208**	0.196**	0.153*
			(0.086)	(0.087)	(0.088)
Common Ethno. Language			-0.114	-0.152	-0.222
			(0.201)	(0.195)	(0.194)
log(Distance)			0.065	0.124	0.309
			(0.107)	(0.125)	(0.226)
Common Border			-0.491***	-0.501***	-0.502***
			(0.167)	(0.167)	(0.167)
Common Official Language			0.228	0.179	0.009
			(0.217)	(0.241)	(0.287)
Common Colonizer			-0.215	-0.260	-0.490
			(0.472)	(0.470)	(0.590)
Colonial Relationship after 1945			0.197	0.119	-0.087
			(0.206)	(0.227)	(0.312)
Origin-country × Year FE	Yes	Yes	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.723	0.760	0.779	0.779	0.777
Instrument (1st stage)				0.190	-17.763
p-value (1st stage)				0.000	0.000
F-value (1st stage)				829.2	641.0
Observations	40,272	40,272	40,272	39,371	37,360

Table 3: Information Hypothesis - Greenfield Investment vs. M&A

This table shows the results of estimating Regression 2 with the sample split into two periods. Column (1) includes the years 2010-2014 and 2019. On average, according to the UNCTAD World Investment Report 2020, M&A accounts for a fraction of one third of all FDI in this period. Column (2) includes all years from 2015 to 2018. During this period the share of M&A accounts for almost a half of all FDI. Column (3) is a replication of the baseline result from Table 2. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	FDI		
Sample split	Low M&A	High M&A	Full sample
	(1)	(2)	(3)
log(SCI)	0.235***	0.067	0.157***
	(0.055)	(0.075)	(0.053)
log(Trade Volume)	0.462***	0.538***	0.495***
	(0.064)	(0.103)	(0.068)
log(GDP Difference)	-0.070**	-0.145***	-0.103***
	(0.034)	(0.056)	(0.038)
log(GDP Growth Difference)	-0.046	-0.122**	-0.064**
	(0.028)	(0.053)	(0.025)
Regional Trade Agreement	0.096	0.327*	0.198
	(0.113)	(0.187)	(0.125)
Common Currency	0.449**	-0.466*	0.041
	(0.199)	(0.256)	(0.188)
Shared Religion Index	0.585*	1.174***	0.847***
	(0.311)	(0.355)	(0.288)
Political Distance	-0.571**	-0.444	-0.532**
	(0.242)	(0.364)	(0.263)
Common Legal Origin	0.153*	0.282**	0.208**
	(0.088)	(0.116)	(0.086)
Common Ethno. Language	-0.156	-0.099	-0.114
	(0.198)	(0.252)	(0.201)
log(Distance)	0.095	0.024	0.065
	(0.087)	(0.165)	(0.107)
Common Border	-0.461***	-0.532**	-0.491***
	(0.133)	(0.258)	(0.167)
Common Official Language	0.289	0.126	0.228
	(0.209)	(0.275)	(0.217)
Common Colonizer	-0.291	-0.105	-0.215
	(0.370)	(0.652)	(0.472)
Colonial Relationship after 1945	0.167	0.277	0.197
	(0.197)	(0.283)	(0.206)
Origin-country × Year FE	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes
Pseudo R^2	0.774	0.789	0.779
Observations	23,106	17,166	40,272

Table 4: Information Hypothesis - Trade, Economic and Cultural Proximity

This table shows the results for Equation 3, estimating the influence of the SCI interacted with different quantiles for the bilateral trade volume, dummies for regional trade agreements, a common official language, a common legal origin and a measure for political distance between two countries. Descriptive statistics can be found in Table 1, detailed descriptions of all variables are presented in Table A1. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	FDI				
	(1)	(2)	(3)	(4)	(5)
log(SCI)	0.162***	0.269***	0.224***	0.105*	0.226***
	(0.053)	(0.061)	(0.055)	(0.059)	(0.054)
log(SCI) × Trade Volume Quantile 1	0.143***				
	(0.052)				
log(SCI) × Trade Volume Quantile 2	-0.000				
	(0.029)				
log(SCI) × Trade Volume Quantile 3	-0.025				
	(0.022)				
log(SCI) × Trade Volume Quantile 4	0.001				
	(0.014)				
log(SCI) × Regional Trade Agreement		-0.194***			
		(0.064)			
log(SCI) × Common Legal Origin			-0.121**		
			(0.052)		
log(SCI) × Political Distance				0.294**	
				(0.130)	
log(SCI) × Common Official Language					-0.271***
					(0.084)
Full Set of Controls	Yes	Yes	Yes	Yes	Yes
Origin-country × Year FE	Yes	Yes	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.781	0.780	0.780	0.780	0.781
Observations	40,272	40,272	40,272	40,272	40,272

Table 5: Institution Hypothesis - Ease of Starting a Business and Financial Markets

This table presents the results of Equation 3, estimating the influence of the SCI on FDI by interacting the SCI with variables for the time, procedures and costs that are required to start a business in the destination country as well as the destination country's financial market development. To check for robustness, interaction terms for the origin countries are also included. Time and procedures are simple count data, entering the regression in logarithms. Cost is originally measured as percentage of gross national income and rescaled by dividing by 100. The index for the financial market development ranges between 0 and 1 with higher values corresponding to more developed financial markets. Variables for columns (1) to (3) are winsorized at the 99th percentile to account for outliers and data is taken from the CEPII database. Data for the financial market development is taken from the IMF Financial Development Index Database. Descriptive statistics can be found in Table A12, detailed descriptions of all variables are presented in Table A1. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	FDI			
	(1)	(2)	(3)	(4)
log(SCI)	-0.216	-0.279	0.068	0.437***
	(0.138)	(0.174)	(0.063)	(0.091)
log(SCI) × log(Dest. Days to Start a Business)	0.125***			
	(0.039)			
log(SCI) × log(Orig. Days to Start a Business)	0.026			
	(0.031)			
log(SCI) × log(Dest. Procedures to Start a Business)		0.205***		
		(0.067)		
log(SCI) × log(Orig. Procedures to Start a Business)		0.034		
		(0.056)		
log(SCI) × Dest. Costs to Start a Business (GNI share)			0.673***	
			(0.189)	
log(SCI) × Orig. Costs to Start a Business (GNI share)			0.217	
			(0.220)	
log(SCI) × Dest. Financial Market Development				-0.502***
				(0.140)
log(SCI) × Orig. Financial Market Development				-0.129
				(0.127)
Full Set of Controls	Yes	Yes	Yes	Yes
Origin-country × Year FE	Yes	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes	Yes
Pseudo R^2	0.782	0.781	0.781	0.781
Observations	40,140	40,140	40,140	40,184

Table 6: Institution Hypothesis - World Governance and Exclusion Indicators

This table shows the results for Equation 3, estimating the influence of the SCI interacted with different measures for the quality of institutions in the destination and origin country on the FDI. Data is taken from the Worldwide Governance Indicators (WGI) project. Estimates for the institutional quality vary approximately between -2.5 (weak) and 2.5 (strong). Precise information on the collection and composition of all dimensions can be found in Kaufmann et al. (2006). Data for the exclusion of socio-economic groups is taken from the Varieties of Democracy Institute and ranges between 0 (weak) and 1 (strong). Descriptive statistics can be found in Table A12, detailed descriptions of all variables are presented in Table A1. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	FDI				
	(1)	(2)	(3)	(4)	(5)
log(SCI)	0.275***	0.268***	0.186***	0.285***	0.038
	(0.062)	(0.066)	(0.061)	(0.069)	(0.072)
log(SCI) × Dest. Control of Corruption	-0.137***				
	(0.032)				
log(SCI) × Orig. Control of Corruption	-0.062*				
	(0.034)				
log(SCI) × Dest. Rule of Law		-0.160***			
		(0.036)			
log(SCI) × Orig. Rule of Law		-0.046			
		(0.041)			
log(SCI) × Dest. Political Stability			-0.095**		
			(0.041)		
log(SCI) × Orig. Political Stability			-0.041		
			(0.046)		
log(SCI) × Dest. Government Effectiveness				-0.158***	
				(0.040)	
log(SCI) × Orig. Government Effectiveness				-0.046	
				(0.042)	
log(SCI) × Dest. Exclusion of Socio-economic Groups					0.381***
					(0.140)
log(SCI) × Orig. Exclusion of Socio-Economic Groups					0.041
					(0.219)
Full Set of Controls	Yes	Yes	Yes	Yes	Yes
Origin-country × Year FE	Yes	Yes	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.782	0.782	0.780	0.781	0.780
Observations	40,272	40,272	40,272	40,272	40,272

Table 7: Uncertainty Hypothesis - World Uncertainty Index

This table shows the results for Equation 3, estimating the influence of the SCI interacted with the World Uncertainty Index (WUI) estimated by OLS. For each version of the WUI the table show a specification with the full sample and the sample with low-M&A years only. Descriptive statistics can be found in Table A12, detailed descriptions of all variables are presented in Table A1. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	log(FDI)					
Sample split	Full sample	Low M&A	Full sample	Low M&A	Full sample	Low M&A
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)	0.209***	0.114*	-0.212**	-0.379***	0.0898	0.0532
	(3.57)	(1.86)	(-2.52)	(-3.58)	(1.40)	(0.79)
log(SCI) × Dest. WUI	0.201*	0.440***				
	(1.69)	(3.13)				
log(SCI) × Orig. WUI	0.156	0.223				
	(1.22)	(1.47)				
log(SCI) × Global WUI			2.241***	3.021***		
			(6.49)	(6.30)		
log(SCI) × Global GDP-Weighted WUI					0.830***	0.826***
					(4.02)	(3.91)
Full Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes
Origin-country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.642	0.639	0.643	0.640	0.642	0.639
Observations	24,609	14,183	24,609	14,183	24,609	14,183

Table 8: Uncertainty Hypothesis - Stock Market Uncertainty

This table shows the results for Equation 3, estimating the influence of the SCI interacted with the CBOE VIX Index estimated by OLS. Column (1) shows the estimates for the full sample, Column (2) includes low-M&A years only and Column (3) high-M&A years only. Descriptive statistics can be found in Table A12, detailed descriptions of all variables are presented in Table A1. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses.
Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	log(FDI)		
Sample split	Full sample	Low M&A	High M&A
	(1)	(2)	(3)
log(SCI)	1.236***	1.730***	-0.0783
	(7.37)	(7.78)	(-0.32)
log(SCI) × log(VIX Index)	-0.328***	-0.507***	0.168*
	(-5.86)	(-6.82)	(1.89)
Full Set of Controls	Yes	Yes	Yes
Origin-country × Year FE	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes
R^2	0.645	0.643	0.644
Observations	27,895	15,960	11,935

Appendix

A1 Implementation: Control Function Approach

The problems that occur when estimating a gravity equation with OLS, highlighted by Silva and Tenreyro (2006), also apply to standard IV strategies like a two stage least squares estimator. Next to that there is no efficient PPML estimator that allows to instrument for a variable and to control for a large set of fixed effects. To overcome these problems, we use a control function approach which is well explained by Wooldridge (2015) and Lin and Wooldridge (2019). This approach relies on two steps: In the first one we regress the SCI on the instrument, the gravity, economic and cultural controls and include the full set of fixed effects, using a simple OLS regression:

$$\log\left(\mathrm{SCI}_{i,j,t}\right) = \beta_1 \cdot \log\left(\mathrm{IV}_{i,j}\right) + \beta_2 \cdot \mathrm{G}_{i,j,t} + \beta_3 \cdot \mathrm{A}_{i,j,t} + \delta_{i,t} + \delta_{j,t} + \eta_{i,j,t} \tag{4}$$

Since this equation is estimated with OLS, the dependent variable SCI is included in a logarithmized term, IV represents the instrumental variable and also enters logarithmized[17]. All other notations are similar to the baseline model. To preserve consistency in the subsequent step, the error term is represented by $\eta_{i,j,t}$. Based on the results of Equation 4, the residuals are estimated and denoted by $Res_{i,j,t}$. These are subsequently included as another control variable in the PPML estimation, similar to the baseline model:

$$\begin{aligned}\mathrm{FDI}_{i,j,t} = \exp[&\beta_1 \cdot \log\left(\mathrm{SCI}_{i,j}\right) + \beta_2 \cdot \mathrm{Res}_{i,j,t} + \beta_3 \cdot \mathrm{G}_{i,j,t} + \beta_4 \cdot \mathrm{A}_{i,j,t} \\ &+ \delta_{i,t} + \delta_{j,t}] \cdot \epsilon_{i,j,t}\end{aligned} \tag{5}$$

This allows to decompose the effect of SCI into the part which can be explained by the instrument and a part that cannot. The coefficient of the SCI represents the elasticity of FDI using

[17] To prevent the loss of all country pairs that exhibit zero migration stocks, we add one before taking the logarithm.

the respective instrument while the coefficient of the residual displays the remaining effect of the SCI in explaining FDI. The closer the latter coefficient is to zero, the smaller is the over- or underestimation of the role of social relationships as a determinant of FDI.

A2 Social Connectedness, Migration and Foreign Direct Investment

Compared to the trade literature less is known about the effect of migration on capital flows like FDI. Javorcik et al. (2011) and Cuadros et al. (2019) find a positive influence of migration on FDI, especially for high-skilled migrants. Based on a larger dataset, Leblang (2010) also shows that migrants promote FDI. He explicitly highlights a channel of formal or information contacts of migrants to their home countries. Burchardi et al. (2019) point out a positive effect of migration flows to US counties over 130 years on today's cross-border investments of US firms. We establish a link between social connectedness and FDI and acknowledge that migration is a major determinant of present social ties (Bailey et al., 2018). To show that migration does not directly influence FDI but the effect materializes largely through social connectedness, we perform the estimation of Equation 2 in a slightly changed way. Table A4 presents the results. The first three columns regress different covariates for migration and our full vector of controls on FDI without including the SCI. In Column (4) to (6) the specifications are estimated again, this time including the SCI. To test the effect of migration we include three different measures from different sources, covering different migration movements. Starting with the most recent migration, Column (1) controls for the average stock of migrants from the partner country in both, the destination and the origin country of investment. Data is extracted from census rounds with an interval of five years between 1990 and 2010 by the UN Population Division. Column (2) shows results for older average migration stock data covering five census rounds from 1960 until 2000, compiled by the World Bank. In order to identify even more distant migratory movements we use data on genetic distance from Spolaore and Wacziarg (2018) in Column (3). The genetic distance is a measure that builds on the time elapsed between a destination and origin country of investment having last common ancestors. Using gene mutations this data covers far more movements of people around the globe than any data set can do by simply counting people moving between countries. While the estimate for migration 1990 until 2010 in Column (1) is almost zero and not significant, the other two measures show estimates significant on a 5%-level and a 10%-level, small with 0.05 in Column (2) and large

with −8.27 in Column (3) which yields an elasticity of −1 regarding FDI[18]. When we include the SCI, the coefficient in Column (4) is still insignificant, in Column (5) and Column (6) the estimates turn insignificant and drop down to 0.02 and −6.83, respectively. However, all elasticities of the SCI with respect to FDI are significant and comparable to our baseline result ranging between 0.13 and 0.18. This result suggests that migration is not the only driver of FDI and that regardless on which control for migration is included, social connectedness still plays a significant role in explaining FDI.

A3 Appendix Tables and Figures

Table A1: Variable Definitions and Data Sources

Variable Name	Description
Main dependent variable	
Inward FDI	Inward foreign direct investment flow in million USD from origin country i to destination country j in year t; source: UNCTAD WIR 2020
Main explanatory variable	
Social Connectedness Index	Relative probability of a user in country i being friends with a user in country j on Facebook taking values between 1 and 1 billion, winsorized at the 99th percentile; source: Humanitarian Data Exchange
Gravity control variables	
Distance	Population weighted distance measured in km between country i and country j; source: CEPII
Common Border	Binary variable equal to one if the origin country i and destination country j share the same border; source: CEPII
Common Official Language	Binary variable equal to one if the origin country i and destination country j share the same official language; source: CEPII
Common Colonizer	Binary variably equal to one if the origin country i and destination country j had a common colonizer after 1945; source: CEPII

[18]Elasticity is collocated by $exp(-8.27) - 1 = -1$.

Table A1: Variable Definitions and Data Sources(continued)

Variable Name	Description
Colonial Relationship after 1945	Binary variable equal to one if the origin country i and destination country j had a colonial relationship after 1945; source: CEPII
Economic control variables	
Avg. Trade Volume	Total average trade volume per country pair in million USD, i.e., country i's imports from country j added up with country j's imports from country i measured as cost, freight, insurance (CIF); averaged over 3 years prior to FDI in t; source: IMF
Avg. GDP Difference	Absolute value of the difference between the average GDP in current (2020) USD of origin country i and destination country j; averaged over three years prior to the FDI in t; source: The World Bank
Avg. GDP Growth Difference	Absolute value of the difference between the average GDP growth rate in percentage points of origin country i and destination country j; averaged over three years prior to the FDI in t; source: The World Bank
Regional Trade Agreement	Binary variable equal to one if origin country i and destination country j have a regional trade agreement in place in year t. Values are three years lagged, as all other economic covariates; source: CEPII
Common Currency	Binary variable equal to one if origin country i and destination country j have the same currency; source: DataHub
Cultural Control Variables	
Shared Religion Index	Index reaching from 0 to 1 indicating to what extend origin country i and destination country j share the same religion; data for 2010; source: PEW Research Center
Political Distance	Index reaching from 0 to 1 indicating proximity of the political systems in origin country i and destination country j in year t, constructed as euclidean distance; source: Varieties of Democracy Institute
Common Legal Origin	Binary variable equal to one if origin country i and destination country j have the same legal origin CEPII
Common Ethnological Language	Binary variable equal to one if origin country i and destination country j share a common language spoken by at least 9% of the population; source: CEPII
Instruments for the SCI	

Table A1: Variable Definitions and Data Sources(continued)

Variable Name	Description
Genetic Distance	Measure building on the time elapsed since the last common ancestors of people living in origin country i and destination country j; source Spolaore and Wacziarg (2018)
Avg. Migration Stock	The (average) value of the country pair's total migration stocks i.e., number of migrants from country i living in country j plus the number of migrants from country j living in country i. For 5 census rounds from 1960-2000, source: The World Bank and for 7 census rounds from 1990-2019, source: United Nations Population Division
Unilateral Variables Institutional Channel	
Dest. Days to Start a Business	Unilateral measure of the days that are required to start a business in the destination country of the investment, winsorized at the 99th percentile; source: CEPII
Orig. Days to Start a Business	Unilateral measure of the days that are required to start a business in the origin country of the investment, winsorized at the 99th percentile; source: CEPII
Dest. Procedures to Start a Business	Unilateral measure of the procedures that are required to start a business in the destination country of the investment, winsorized at the 99th percentile; source: CEPII
Orig. Procedures to Start a Business	Unilateral measure of the procedures that are required to start a business in the origin country of the investment, winsorized at the 99th percentile; source: CEPII
Dest. Costs to Start a Business	Unilateral measure of the costs as percent of the gross national income (GNI) that occur when starting a business in the destination country of the investment, winsorized at the 99th percentile; source: CEPII
Orig. Costs to Start a Business	Unilateral measure of the costs as percent of the gross national income (GNI) that occur when starting a business in the origin country of the investment, winsorized at the 99th percentile; source: CEPII
Dest. Financial Market Development	Unilateral measure for the development of financial markets in the destination country of the investment, taking values between 0 (low) and 1 (high); source: IMF Financial Development Index

Table A1: Variable Definitions and Data Sources(continued)

Variable Name	Description
Orig. Financial Market Development	Unilateral measure for the development of financial markets in the origin country of the investment, taking values between 0 (low) and 1 (high); source: IMF Financial Development Index
Dest. Control of Corruption	Measure of the strength and effectiveness of a country's policy and institutional framework to prevent and combat corruption. Standard normal estimates take values that vary approximately between -2.5 (weak) and 2.5 (strong); source:Worldwide Governance Indicators (WGI) project
Orig. Control of Corruption	Measure of the strength and effectiveness of a country's policy and institutional framework to prevent and combat corruption. Standard normal estimates take values that vary approximately between -2.5 (weak) and 2.5 (strong); source:Worldwide Governance Indicators (WGI) project
Dest. Rule of Law	Measure of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence. Standard normal estimates take values that vary approximately between -2.5 (weak) and 2.5 (strong); source:Worldwide Governance Indicators (WGI) project
Orig. Rule of Law	Measure of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence. Standard normal estimates take values that vary approximately between -2.5 (weak) and 2.5 (strong); source:Worldwide Governance Indicators (WGI) project
Dest. Political Stability	Measure of the perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism. Standard normal estimates take values that vary approximately between -2.5 (weak) and 2.5 (strong); source:Worldwide Governance Indicators (WGI) project
Orig. Political Stability	Measure of the perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism. Standard normal estimates take values that vary approximately between -2.5 (weak) and 2.5 (strong); source:Worldwide Governance Indicators (WGI) project

Table A1: Variable Definitions and Data Sources(continued)

Variable Name	Description
Dest. Government Effectiveness	Measure of the perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.Standard normal estimates take values that vary approximately between -2.5 (weak) and 2.5 (strong); source:Worldwide Governance Indicators (WGI) project
Orig. Government Effectiveness	Measure of the perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.Standard normal estimates take values that vary approximately between -2.5 (weak) and 2.5 (strong); source:Worldwide Governance Indicators (WGI) project
Dest. Exclusion of Socio-economic Groups	Measure of the extend to which individuals are denied access to services or participation in governed spaces based on their identity or belonging to a particular group. Index takes values between 0 (low) and 1 (high); source: Varieties of Democracy Institute
Orig. Exclusion of Socio-economic Groups	Measure of the extend to which individuals are denied access to services or participation in governed spaces based on their identity or belonging to a particular group. Index takes values between 0 (low) and 1 (high); source: Varieties of Democracy Institute
Uncertainty Channel	
Dest. World Uncertainty Index	Index measuring uncertainty in the destination country of the investment computed by counting the percent of word "uncertain" (or its variant) in the Economist Intelligence Unit country reports. The WUI is then rescaled by multiplying by 1,000,000. A higher number means higher uncertainty and vice versa; source: World Uncertainty Index

Table A1: Variable Definitions and Data Sources(continued)

Variable Name	**Description**
Orig. World Uncertainty Index	Index measuring uncertainty in the origin country of the investment computed by counting the percent of word "uncertain" (or its variant) in the Economist Intelligence Unit country reports. The WUI is rescaled to take values between 0 and 1. A higher number means higher uncertainty and vice versa; source: World Uncertainty Index
Global World Uncertainty Index	Index measuring global uncertainty computed by counting the percent of word "uncertain" (or its variant) in the Economist Intelligence Unit country reports. The WUI is then rescaled by multiplying by 1,000,000. A higher number means higher uncertainty and vice versa; source: World Uncertainty Index
Global GDP-Weighted World Uncertainty Index	GDP-Weighted index measuring global uncertainty computed by counting the percent of word "uncertain" (or its variant) in the Economist Intelligence Unit country reports. The WUI is rescaled to take values between 0 and 1. A higher number means higher uncertainty and vice versa; source: World Uncertainty Index
VIX	Index that expresses the stock market's expectation of volatility of the U.S. equity index S&P 500 introduced by the Chicago Board Options Exchange (CBOE). Higher values correspond to higher uncertainty, daily closing ; source: VIX Index

Table A2: Correlation matrix

This table displays the correlation between all bilateral control variables and instrumental variables and the log of SCI and the log of distance.

	SCI	Distance
Social Connectedness Index	1.00	-0.29
Distance	-0.29	1.00
Common Border	0.26	-0.26
Common Official Language	0.29	-0.03
Common Colonizer	0.25	-0.15
Colonial Relationship after 1945	0.09	0.02
Trade Volume	0.03	-0.09
GDP Difference	-0.09	0.24
Avg. GDP Growth Difference	-0.06	0.05
Regional Trade Agreement	0.24	-0.52
Common Currency	0.13	-0.28
Shared Religion Index	0.17	-0.13
Political Distance	-0.14	0.10
Common Legal Origin	0.18	-0.11
Common Ethno. Language	0.25	-0.02
Genetic Distance	-0.19	0.53
log(Avg. Migration 1990-2010) (UN)	0.24	-0.17
log(Avg. Migration 1960-2000) (WB)	0.24	-0.14

Table A3: Gravity Regressions - Predicting Separately

This table shows the results of estimating the effect of the SCI and all gravity, economic and cultural variables that are included as controls in Equation 2 on the dependent variable Inward FDI from country i to country j separately. Thereby the same PPML estimator as in Equation 2 is used. Column (1) represents a specification with fixed effects only, similar to Column (1) of Table 2. Columns (2)-(16) include all other controls one-by-one in decreasing order of explanatory power measured by the within-R^2. Descriptive statistics can be found in Table 1, detailed descriptions of all variables are presented in Table A1. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

	Dependent Variable: Inward FDI															
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
log(Trade Volume)		0.598***														
log(SCI)			0.544***													
log(Distance)				-0.610***												
Regional Trade Agreement					0.902***											
Common Border						0.893***										
Shared Religion Index							2.388***									
Political Distance								-1.768***								
Common Official Language									0.764***							
Common Ethno. Language										0.740***						
Common Legal Origin											0.544***					
Common Colonizer												2.032***				
Common Currency													0.798***			
log(GDP Difference)														-0.140***		
log(GDP Growth Difference)															-0.147***	
Colonial Relationship after 1945																-0.038
Origin-country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.723	0.767	0.760	0.756	0.742	0.734	0.733	0.733	0.733	0.733	0.732	0.729	0.728	0.725	0.725	0.723
Observations	40,272	40,272	40,272	40,272	40,272	40,272	40,272	40,272	40,272	40,272	40,272	40,272	40,272	40,272	40,272	40,272

Table A4: Social Connectedness and Foreign Direct Investment - Migration

This table presents results of estimating baseline Equation 2 while adding different controls for migratory movements in the past. Thereby, in Columns (1)-(3), the log of the SCI is not included, before we repeat the specifications with the log of the SCI in Columns (4)-(6). Columns (1) and (4) include the log of average migration stocks in five census rounds from 1990 until 2010. Columns (2) and (5) include the log of average migration stocks in five census rounds from 1960 until 2000. Finally, Columns (3) and (6) include a control for genetic distance taken from Spolaore and Wacziarg (2018). Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	FDI					
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)				0.183***	0.128**	0.138**
				(0.059)	(0.058)	(0.057)
log(Avg. Migration 1990-2010) (UN)	0.024			-0.019		
	(0.022)			(0.023)		
log(Avg. Migration 1960-2000) (WB)		0.053**			0.023	
		(0.026)			(0.029)	
Genetic Distance			-8.270*			-6.828
			(4.749)			(4.838)
log(Trade Volume)	0.524***	0.506***	0.531***	0.498***	0.486***	0.492***
	(0.068)	(0.068)	(0.066)	(0.069)	(0.070)	(0.071)
log(GDP Difference)	-0.106***	-0.106***	-0.109***	-0.104***	-0.102***	-0.104***
	(0.038)	(0.038)	(0.040)	(0.038)	(0.038)	(0.039)
log(GDP Growth Difference)	-0.065**	-0.066**	-0.065**	-0.062**	-0.065**	-0.065**
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Regional Trade Agreement	0.188	0.221*	0.186	0.188	0.217*	0.211*
	(0.128)	(0.129)	(0.125)	(0.128)	(0.131)	(0.126)
Common Currency	0.052	0.058	0.055	0.039	0.048	0.049
	(0.188)	(0.188)	(0.187)	(0.188)	(0.188)	(0.187)
Shared Religion Index	1.068***	1.037***	0.848***	0.822***	0.873***	0.678**
	(0.281)	(0.279)	(0.313)	(0.284)	(0.288)	(0.316)
Political Distance	-0.481*	-0.459*	-0.450*	-0.531**	-0.512*	-0.510*
	(0.264)	(0.266)	(0.265)	(0.263)	(0.267)	(0.263)
Common Legal Origin	0.207**	0.211**	0.215**	0.209**	0.210**	0.214**
	(0.085)	(0.085)	(0.085)	(0.086)	(0.086)	(0.086)
Common Ethno. Language	-0.047	-0.065	0.020	-0.111	-0.118	-0.072
	(0.194)	(0.189)	(0.194)	(0.201)	(0.199)	(0.206)
Gravity Controls	Yes	Yes	Yes	Yes	Yes	Yes
Origin-country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.777	0.778	0.776	0.778	0.779	0.777
Observations	39,102	39,371	37,360	39,102	39,371	37,360

Table A5: First Stage Regressions for the Control Function Approach

This table presents the results for Equation 4, the first stage OLS regressions for Columns (4) and (5) of Table 2. All specifications regress the logarithmized SCI on the full set of controls and the respective instrument which is the log of total bilateral migration in Column (1) and genetic distance in Column (2). Descriptive statistics for all variables are presented in Table 1, detailed descriptions are presented in Table A1. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	log(SCI)	
	(1)	(2)
log(Avg. Migration 1960-2000) (WB)	0.190***	
	(0.000)	
Genetic Distance		-17.763***
		(0.000)
log(Trade Volume)	0.173***	0.244***
	(0.000)	(0.000)
log(GDP Difference)	0.000	0.003
	(0.973)	(0.778)
log(GDP Growth Difference)	-0.013**	0.006
	(0.031)	(0.352)
Regional Trade Agreement	0.207***	0.201***
	(0.000)	(0.000)
Common Currency	-0.070	-0.123**
	(0.200)	(0.038)
Shared Religion Index	0.863***	1.066***
	(0.000)	(0.000)
Political Distance	-0.020	0.055
	(0.736)	(0.375)
Common Legal Origin	0.114***	0.158***
	(0.000)	(0.000)
Common Ethno. Language	0.274***	0.393***
	(0.000)	(0.000)
Gravity Controls	Yes	Yes
Origin-country × Year FE	Yes	Yes
Destination-country × Year FE	Yes	Yes
Adjusted R^2	0.811	0.792
Observations	39,407	37,396

Table A6: Instrumental Variable Approach - OLS

This table presents results of an instrumental variable approach using simple OLS. In Column (1) the log of average bilateral migration stocks between 1960 and 2000 from the Wold Bank is used to instrument for the SCI and Column (2) uses genetic distance from Spolaore and Wacziarg (2018) as an instrument for the SCI. As we loose all observations with zero FDI when taking the logarithm of FDI, we present specifications where we add one before taking the logarithm to be able to keep observations yielding zero FDI in Columns (3) and (4). The coefficient of the first stage as well as the p-value for both instruments are also included to show, that the instruments are relevant. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	log(FDI)		log(1+FDI)	
Instrument	Bilateral Migration	Genetic Distance	Bilateral Migration	Genetic Distance
	(1)	(2)	(3)	(4)
log(SCI)	0.483***	0.450***	0.221***	0.449***
	(0.078)	(0.147)	(0.039)	(0.108)
log(Trade Volume)	0.357***	0.354***	0.098***	0.039
	(0.034)	(0.049)	(0.015)	(0.029)
log(GDP Difference)	-0.263***	-0.264***	-0.166***	-0.164***
	(0.037)	(0.037)	(0.016)	(0.017)
log(GDP Growth Difference)	0.017	0.005	0.022*	0.023*
	(0.027)	(0.027)	(0.013)	(0.013)
Regional Trade Agreement	-0.088	-0.110	0.027	-0.022
	(0.081)	(0.087)	(0.042)	(0.049)
Common Currency	0.004	0.099	-0.235***	-0.231***
	(0.184)	(0.177)	(0.072)	(0.074)
Shared Religion Index	0.738***	0.695***	0.359***	0.084
	(0.133)	(0.215)	(0.067)	(0.133)
Political Distance	-0.374**	-0.401**	-0.118	-0.130*
	(0.161)	(0.162)	(0.074)	(0.075)
Common Legal Origin	0.060	0.101	0.147***	0.122***
	(0.065)	(0.064)	(0.029)	(0.034)
Common Ethno. Language	-0.135	-0.164	0.194***	0.115
	(0.128)	(0.141)	(0.065)	(0.080)
log(Distance)	-0.718***	-0.706***	-0.359***	-0.204**
	(0.084)	(0.119)	(0.040)	(0.079)
Common Border	-0.401***	-0.235**	-0.015	0.009
	(0.111)	(0.097)	(0.078)	(0.086)
Common Official Language	0.803***	0.809***	0.066	-0.078
	(0.134)	(0.164)	(0.064)	(0.085)
Common Colonizer	0.131	0.169	-0.051	-0.171*
	(0.138)	(0.173)	(0.062)	(0.093)
Colonial Relationship after 1945	0.343	0.406*	0.269**	0.138
	(0.230)	(0.244)	(0.106)	(0.135)
Origin-country FE	Yes	Yes	Yes	Yes
Destination-country FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.225	0.223	0.136	0.119
Instrument (1st stage)	0.190	-17.763	0.190	-17.763
p-value (1st stage)	0.000	0.000	0.000	0.000
F-value (1st stage)	904.1	227.9	999.9	204.6
Observations	27,201	25,862	39,407	37,396

Table A7: Control Function Approach with Different Sets of Migration Data

This table estimates Equation 5 with different sets of migration data. Column (1) represents the control function approach with data for the 1960 census round only. The following column adds the subsequent census rounds decade by decade and use the average of the total migration flows between two countries as instrument for the SCI. Hence Column (2) uses the average of the 1960 and 1970 census rounds, Column (3) uses the average of the census rounds from 1960, 1070 and 1980 and so on. Column (5) is a replication of Column (4) of Table 2 since the average of all census rounds represents the IV which is used there. The first stage estimates for this table are presented in Table A8. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	FDI				
Insturmental variable	Migration averaged up to				
	1960	1970	1980	1990	2000
	(1)	(2)	(3)	(4)	(5)
log(SCI)	0.937***	0.678***	0.496**	0.342**	0.251*
	(0.246)	(0.225)	(0.197)	(0.167)	(0.139)
Residuals	-0.869***	-0.601**	-0.402*	-0.229	-0.123
	(0.262)	(0.242)	(0.214)	(0.185)	(0.153)
log(Trade Volume)	0.277***	0.349***	0.399***	0.441***	0.465***
	(0.099)	(0.096)	(0.092)	(0.087)	(0.083)
log(GDP Difference)	-0.095***	-0.099***	-0.100***	-0.103***	-0.102***
	(0.037)	(0.037)	(0.037)	(0.037)	(0.038)
log(GDP Growth Difference)	-0.060**	-0.061**	-0.063**	-0.063**	-0.063**
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Regional Trade Agreement	0.089	0.145	0.165	0.180	0.192
	(0.124)	(0.124)	(0.124)	(0.124)	(0.124)
Common Currency	0.144	0.126	0.103	0.071	0.057
	(0.188)	(0.190)	(0.189)	(0.188)	(0.187)
Shared Religion Index	0.030	0.271	0.485	0.664**	0.767**
	(0.387)	(0.375)	(0.359)	(0.337)	(0.318)
Political Distance	-0.477*	-0.479*	-0.500*	-0.506*	-0.510*
	(0.264)	(0.269)	(0.264)	(0.266)	(0.268)
Common Legal Origin	0.114	0.140	0.160*	0.181**	0.196**
	(0.091)	(0.090)	(0.089)	(0.088)	(0.087)
Common Ethno. Language	-0.363**	-0.295	-0.239	-0.187	-0.152
	(0.180)	(0.184)	(0.189)	(0.193)	(0.195)
Gravity Controls	Yes	Yes	Yes	Yes	Yes
Origin-country × Year FE	Yes	Yes	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.780	0.779	0.779	0.779	0.779
Observations	39,371	39,371	39,371	39,371	39,371

Table A8: First Stage Regressions for Table A7

This table presents the results of the first stage regressions for Table A7. Dependent variable is the logarithm of the SCI which is regressed on the different instruments and the full set of controls. In Column (1) the migration data from the 1960 census round only is used as an instrument. The following columns add the subsequent census rounds decade by decade and use the average of the total migration flows between two countries as instrument for the SCI. Hence Column (2) uses the average of the 1960 and 1970 census rounds, Column (3) uses the average of the census rounds from 1960, 1070 and 1980 and so on. Column (5) is a replication of Table A5 Column (1) since the average of all census rounds represents the IV which is used there. Descriptive statistics for all variables are presented in Table 1, detailed descriptions of all variables are presented in Table A1. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	log(SCI)				
	(1)	(2)	(3)	(4)	(5)
log(Avg. Migration in 1960) (WB)	0.092***				
	(0.007)				
log(Avg. Migration 1960-1970) (WB)		0.111***			
		(0.007)			
log(Avg. Migration 1960-1980) (WB)			0.127***		
			(0.007)		
log(Avg. Migration 1960-1990) (WB)				0.156***	
				(0.008)	
log(Avg. Migration 1960-2000) (WB)					0.190***
					(0.008)
log(Trade Volume)	0.217***	0.209***	0.204***	0.187***	0.173***
	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)
log(GDP Difference)	-0.011	-0.011	-0.009	-0.001	0.000
	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)
log(GDP Growth Difference)	-0.002	-0.004	-0.006	-0.010	-0.013**
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Regional Trade Agreement	0.244***	0.215***	0.204***	0.216***	0.207***
	(0.035)	(0.034)	(0.034)	(0.034)	(0.033)
Common Currency	-0.123**	-0.111*	-0.122**	-0.082	-0.070
	(0.058)	(0.058)	(0.057)	(0.056)	(0.055)
Shared Religion Index	0.983***	0.960***	0.955***	0.924***	0.863***
	(0.066)	(0.065)	(0.065)	(0.064)	(0.063)
Political Distance	0.080	0.090	0.083	0.036	-0.020
	(0.060)	(0.060)	(0.060)	(0.060)	(0.059)
Common Legal Origin	0.121***	0.122***	0.119***	0.118***	0.114***
	(0.028)	(0.028)	(0.028)	(0.027)	(0.026)
Common Ethno. Language	0.351***	0.339***	0.335***	0.315***	0.274***
	(0.067)	(0.067)	(0.066)	(0.063)	(0.062)
Gravity Controls	Yes	Yes	Yes	Yes	Yes
Origin-country × Year FE	Yes	Yes	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.790	0.794	0.797	0.804	0.811
Observations	39,407	39,407	39,407	39,407	39,407

Table A9: Social Connectedness and FDI - Excluding Tax Havens

This table represents the results of Regression 2 estimating the influence of the SCI on Inward FDI, similar to Table 2 while excluding locations that are known to be tax havens. These are the Netherlands, Hong Kong and Singapore. Descriptive statistics for the entire set of FDI data and all covariates can be found in Table 1, detailed descriptions of all variables are presented in Table A1. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable:	log(FDI)		
	(1)	(2)	(3)
log(SCI)		0.589***	0.130***
		(0.039)	(0.049)
log(Trade Volume)			0.629***
			(0.058)
log(GDP Difference)			-0.126***
			(0.029)
log(GDP Growth Difference)			-0.020
			(0.025)
Regional Trade Agreement			0.051
			(0.084)
Common Currency			0.332**
			(0.158)
Shared Religion Index			1.265***
			(0.258)
Political Distance			-0.449*
			(0.262)
Common Legal Origin			0.196**
			(0.079)
Common Ethno. Language			-0.390**
			(0.177)
log(Distance)			0.135*
			(0.078)
Common Border			-0.439***
			(0.138)
Common Official Language			0.614***
			(0.183)
Common Colonizer			0.045
			(0.290)
Colonial Relationship after 1945			0.228
			(0.175)
Origin-country × Year FE	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes
Pseudo R^2	0.746	0.784	0.809
Observations	37,074	37,074	37,074

Table A10: Social Connectedness and FDI - OLS

This table represents the results of Regression 2 estimating the influence of the SCI on FDI by OLS. Since for this estimator the dependent variable FDI needs to be included in logarithms, all observations exhibiting zero FDI are dropped. It therefore shows the effect of the SCI on the intensive margin of FDI. Descriptive statistics for the entire set of FDI data and all covariates can be found in Table 1, detailed descriptions of all variables are presented in Table A1. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable:	log(FDI)			log(FDI+1)		
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)		1.086***	0.322***		0.471***	0.109***
		(0.026)	(0.044)		(0.013)	(0.021)
log(Trade Volume)			0.406***			0.128***
			(0.040)			(0.016)
log(GDP Difference)			-0.267***			-0.164***
			(0.035)			(0.016)
log(GDP Growth Difference)			0.020			0.020
			(0.026)			(0.013)
Regional Trade Agreement			-0.078			0.052
			(0.117)			(0.055)
Common Currency			-0.017			-0.263***
			(0.222)			(0.086)
Shared Religion Index			0.920***			0.495***
			(0.186)			(0.090)
Political Distance			-0.404*			-0.109
			(0.210)			(0.098)
Common Legal Origin			0.057			0.149***
			(0.086)			(0.039)
Common Ethno. Language			-0.072			0.210*
			(0.230)			(0.117)
log(Distance)			-0.828***			-0.440***
			(0.092)			(0.042)
Common Border			-0.370**			0.006
			(0.173)			(0.101)
Common Official Language			0.886***			0.141
			(0.252)			(0.125)
Common Colonizer			0.240			0.032
			(0.214)			(0.094)
Colonial Relationship after 1945			0.521			0.388**
			(0.317)			(0.155)
Origin-country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.538	0.621	0.645	0.371	0.435	0.459
Observations	27,895	27,895	27,895	40,310	40,310	40,310

Table A11: Social Connectedness and Investment Stocks

This table shows the results of estimating the influence of the SCI on Foreign Direct Investment Stocks and Portfolio Investment Stocks in 2019 as dependent variable using PPML. The data is taken from Coordinated Direct Investment Survey of the IMF and the Coordinated Portfolio Investment Survey of the IMF. Missing data for the year 2019 is filled with data for 2018, if available. Negative values are set to 0. Descriptive statistics for the independent variables are presented in Table 1. Regressions include a full set of origin country and destination country fixed effects. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable:	Foreign Direct Investment Stock in 2019			Portfolio Investment Stock in 2019		
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)		0.525***	0.288***		0.564***	0.174*
		(0.091)	(0.071)		(0.083)	(0.093)
log(Trade volume)			0.271**			0.177***
			(0.110)			(0.067)
log(GDP differential)			-0.124***			-0.027
			(0.040)			(0.030)
log(GDP growth differential)			0.016			-0.013
			(0.047)			(0.042)
Regional trade agreement			0.193			0.229**
			(0.156)			(0.089)
Common Currency			-0.061			0.769***
			(0.173)			(0.228)
Shared religion index			1.935***			0.506
			(0.553)			(0.600)
Political distance			-0.002			-0.002
			(0.005)			(0.007)
Common legal origin			-0.020			0.176
			(0.149)			(0.120)
Common Ethno. Language			-0.053			0.121
			(0.183)			(0.209)
log(Distance)			0.068			0.040
			(0.144)			(0.072)
Common Border			-0.122			0.242
			(0.161)			(0.187)
Common Official Language			0.077			-0.119
			(0.167)			(0.165)
Common Colonizer			0.725			1.035
			(0.694)			(0.696)
Colonial Relationship after 1945			0.246			-0.127
			(0.185)			(0.208)
Origin-country FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-country FE	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.897	0.920	0.928	0.931	0.957	0.968
Observations	7,695	7,695	7,665	7,337	7,337	7,305

Table A12: Descriptive Statistics for Unilateral Controls

This table shows summary statistics for the unilateral variables used to identify the channels through which social connectedness influences FDI. Since the panel is unbalanced and the variables unilateral, statistics are presented for the set of origin and destination countries separately. A detailed description of all variables can be found in Table A1.

	Mean	Median	SD	Min	Max	Obs
log(Avg. Migration 1990-2010) (UN)	6.30	6.64	3.43	0.00	15.33	39,172
log(Avg. Migration in 1960) (WB)	4.88	4.82	3.29	0.00	13.93	39,441
log(Avg. Migration 1960-1970) (WB)	5.18	5.18	3.22	0.00	13.95	39,441
log(Avg. Migration 1960-1980) (WB)	5.45	5.52	3.19	0.00	13.96	39,441
log(Avg. Migration 1960-1990) (WB)	5.77	5.87	3.11	0.00	14.22	39,441
log(Avg. Migration 1960-2000) (WB)	6.11	6.17	3.01	0.00	14.45	39,441
Dest. Financial Market Development	0.37	0.33	0.30	0.00	0.97	40,343
Orig. Financial Market Development	0.42	0.43	0.30	0.00	0.97	40,255
Dest. Control of Corruption	0.38	0.16	0.99	-1.67	2.41	40,343
Orig. Control of Corruption	0.54	0.37	1.09	-1.82	2.41	40,343
Dest. Rule of Law	0.47	0.42	0.94	-1.59	2.13	40,343
Orig. Rule of Law	0.57	0.51	1.01	-1.85	2.13	40,343
Dest. Voice & Accountability	0.45	0.60	0.84	-2.07	1.74	40,343
Orig. Voice & Accountability	0.41	0.56	0.95	-2.12	1.74	40,343
Dest. Political Stability	0.16	0.36	0.83	-2.81	1.64	40,343
Orig. Political Stability	0.20	0.39	0.87	-2.81	1.64	40,343
Dest. Government Effectiveness	0.56	0.52	0.85	-1.64	2.24	40,343
Orig. Government Effectiveness	0.63	0.59	0.96	-2.08	2.24	40,343
Dest. Regulatory Quality	0.62	0.64	0.82	-2.24	2.26	40,343
Orig. Regulatory Quality	0.63	0.68	0.94	-2.35	2.26	40,343
Dest. Access to Justice	0.79	0.89	0.21	0.18	1.00	40,343
Orig. Access to Justice	0.79	0.89	0.22	0.09	1.00	40,343
Dest. Property Rights	0.85	0.91	0.13	0.23	0.97	40,343
Orig. Property Rights	0.82	0.90	0.15	0.12	0.97	40,343
Dest. Exclusion of Socio-economic Groups	0.24	0.14	0.23	0.01	0.96	40,343
Orig. Exclusion of Socio-Economic Groups	0.25	0.16	0.23	0.01	0.96	40,343
Dest. Corruption Perception Index	0.53	0.49	0.19	0.15	0.92	32,981
Orig. Corruption Perception Index	0.55	0.52	0.21	0.14	0.92	32,804
Political corruption index	0.36	0.28	0.30	0.00	0.95	40,343
Political corruption index	0.33	0.22	0.30	0.00	0.97	40,343
Dest. Rule of Law (V-DEM)	0.72	0.83	0.28	0.03	1.00	40,343
Orig. Rule of Law (V-DEM)	0.73	0.84	0.28	0.03	1.00	40,343
Dest. Rule of Law (FH)	10.23	11.00	4.15	0.00	16.00	36,724
Orig. Rule of Law (FH)	10.17	11.00	4.60	0.00	16.00	36,223
log(Geopolitical Risk Index)	4.54	4.59	0.10	4.40	4.68	40,343

Table A13: Information Hypothesis - Financial Market Development

This table presents the results of Equation 3, estimating the influence of the SCI interacted with the financial market development in the destination and in the origin country on the FDI. The sample is split into the two time periods, one with a low share of M&A projects (2010-2014 and 2019) in Cloumn (1) and one with a high share of M&A projects (2015-2018) in Column (2). Data is taken from the IMF Financial Development Index Database. Descriptive statistics can be found in Table A12, detailed descriptions of all variables are presented in Table A1. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	FDI		
Sample split	Full Sample	low M&A	high M&A
	(1)	(2)	(3)
log(SCI)	0.437***	0.593***	0.229*
	(0.091)	(0.094)	(0.131)
log(SCI) × Dest. Financial Market Development	-0.502***	-0.452***	-0.531**
	(0.140)	(0.119)	(0.207)
log(SCI) × Orig. Financial Market Development	-0.129	-0.284**	0.084
	(0.127)	(0.120)	(0.177)
Full Set of Controls	Yes	Yes	Yes
Origin-country × Year FE	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes
Pseudo R^2	0.781	0.776	0.790
Observations	40,184	23,059	17,125

Table A14: Institutional Hypothesis - World Governance Indicators

This table shows the presents a robustness check for the results of Table 6, estimating the influence of the SCI interacted with different measures for the quality of institutions in the destination and origin country on the FDI. Data is taken from the Worldwide Governance Indicators (WGI) project. Estimates for the institutional quality vary approximately between -2.5 (weak) and 2.5 (strong). Precise information on the collection and composition of all dimensions can be found in Kaufmann et al. (2006). Descriptive statistics can be found in Table A12, detailed descriptions of all variables are presented in Table A1. Regressions include the full set of bilateral controls as well as a full set of interactions between origin country and destination country fixed effects with year dummies. Standard errors are clustered on the same level as fixed effects are defined and are depicted in parentheses. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	FDI					
	(1)	(2)	(3)	(4)	(5)	(6)
log(SCI)	0.275***	0.268***	0.201***	0.186***	0.285***	0.304***
	(0.062)	(0.066)	(0.060)	(0.061)	(0.069)	(0.068)
log(SCI) × Dest. Control of Corruption	-0.137***					
	(0.032)					
log(SCI) × Orig. Control of Corruption	-0.062*					
	(0.034)					
log(SCI) × Dest. Rule of Law		-0.160***				
		(0.036)				
log(SCI) × Orig. Rule of Law		-0.046				
		(0.041)				
log(SCI) × Dest. Voice & Accountability			-0.086**			
			(0.040)			
log(SCI) × Orig. Voice & Accountability			-0.016			
			(0.044)			
log(SCI) × Dest. Political Stability				-0.095**		
				(0.041)		
log(SCI) × Orig. Political Stability				-0.041		
				(0.046)		
log(SCI) × Dest. Government Effectiveness					-0.158***	
					(0.040)	
log(SCI) × Orig. Government Effectiveness					-0.046	
					(0.042)	
log(SCI) × Dest. Regulatory Quality						-0.163***
						(0.041)
log(SCI) × Orig. Regulatory Quality						-0.051
						(0.045)
Full Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes
Origin-country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.782	0.782	0.780	0.780	0.781	0.782
Observations	40,272	40,272	40,272	40,272	40,272	40,272

Table A15: Institutional Hypothesis - Individual Rights and Participation

This table shows the presents a robustness check for the results of Table 6, estimating the influence of the SCI interacted with different measures for the quality of institutions in the destination and origin country on the FDI. Data is taken from the Varieties of Democracy Institute. Descriptive statistics can be found in Table A12, detailed descriptions of all variables are presented in Table A1. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	FDI		
	(1)	(2)	(3)
log(SCI)	0.478***	0.583**	0.038
	(0.173)	(0.262)	(0.072)
log(SCI) × Dest. Access to Justice	-0.354***		
	(0.129)		
log(SCI) × Orig. Access to Justice	-0.068		
	(0.193)		
log(SCI) × Dest. Property Rights		-0.624**	
		(0.249)	
log(SCI) × Orig. Property Rights		0.103	
		(0.276)	
log(SCI) × Dest. Exclusion of Socio-economic Groups			0.381***
			(0.140)
log(SCI) × Orig. Exclusion of Socio-Economic Groups			0.041
			(0.219)
Full Set of Controls	Yes	Yes	Yes
Origin-country × Year FE	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes
Pseudo R^2	0.780	0.780	0.780
Observations	40,272	40,272	40,272

Table A16: Institutional Hypothesis - Alternatives for WGI Measures

This table shows the presents a robustness check for the results of Table 6. We interact the SCI with four different measures for institutional quality, two corruption and two rule of law measures. Thereby we follow the Equation 3. We include the interaction term for the destination and origin country on the FDI. Data is taken from the Varieties of Democracy Institute, the Corruption Oerception Index and the Freedom House Organisation. The first three columns include indices with values between zero and one, the rule of law measure from Freedom House ranges between zero and 16. While high values for the political corruption index represent a high level of corruption all other variables display high levels of institutional quality with higher numbers. Descriptive statistics can be found in Table A12, detailed descriptions of all variables are presented in Table A1. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	log(FDI)			
	(1)	(2)	(3)	(4)
log(SCI)	0.777***	-0.052	0.479***	0.347**
	(0.147)	(0.078)	(0.141)	(0.138)
log(SCI) × Dest. Corruption Perception Index	-0.795***			
	(0.181)			
log(SCI) × Orig. Corruption Perception Index	-0.301			
	(0.196)			
log(SCI) × Political corruption index		0.437***		
		(0.110)		
log(SCI) × Political corruption index		0.222		
		(0.137)		
log(SCI) × Dest. Rule of Law (V-DEM)			-0.352***	
			(0.115)	
log(SCI) × Orig. Rule of Law (V-DEM)			-0.092	
			(0.142)	
log(SCI) × Dest. Rule of Law (FH)				-0.020**
				(0.009)
log(SCI) × Orig. Rule of Law (FH)				-0.003
				(0.009)
Full Set of Controls	Yes	Yes	Yes	Yes
Origin-country × Year FE	Yes	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes	Yes
Pseudo R^2	0.785	0.781	0.780	0.787
Observations	32,724	40,272	40,272	36,112

Table A17: Uncertainty Hypothesis - Geopolitical Risk Index

This table shows the presents a robustness check for the results of Table 7, estimating the influence of the SCI interacted with the geopolitical risk index, estimated by OLS. Dependent variable is the log of FDI. Data is taken from the Geopolitical Risk Index. Descriptive statistics can be found in Table A12, detailed descriptions of all variables are presented in Table A1. Regressions include a full set of interactions between origin country and destination country fixed effects with year dummies. Observations that are fully explained by the fixed effects are dropped before the estimation. Standard errors are clustered by the origin country and the destination country and are depicted in parentheses. Significance levels: *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Dependent variable	log(FDI)		
	(1)	(2)	(3)
log(SCI)	-2.395***	-2.320***	1.338
	(-4.14)	(-2.66)	(0.77)
log(SCI) × log(Geopolitical Risk Index)	0.597***	0.580***	-0.208
	(4.67)	(2.98)	(-0.55)
Full Set of Controls	Yes	Yes	Yes
Origin-country × Year FE	Yes	Yes	Yes
Destination-country × Year FE	Yes	Yes	Yes
Pseudo R^2	0.621	0.615	0.622
Observations	27,895	15,960	11,935