

Melting pot or salad bowl: Some evidence from U.S. investments abroad[☆]

Utpal Bhattacharya^{a,*}, Peter Groznik^b

^a*Kelley School of Business, Indiana University, Bloomington, IN 47405, USA*

^b*School of Economics, University of Ljubljana, Ljubljana, Slovenia*

Available online 29 January 2008

Abstract

We estimate the number of foreign-origin persons in the United States classified by their country of origin from census data in 1970, 1980, 1990 and 2000. We find, both in cross-sectional tests and in panel data tests, that the size of the foreign-origin group from a country living in the U.S. is positively correlated with U.S. investments in that country. This national origin bias is strong for direct (FDI) and modest for indirect (equity holdings) investments. The results continue to hold even after controlling for the “fundamentals” hypothesized to affect foreign investments. The other economic geography variables of a country—physical distance from the U.S., race, language and religion—do not seem to affect US investments in that country.

© 2008 Elsevier B.V. All rights reserved.

JEL classification: G11; G15

Keywords: Economic geography; Foreign direct investment; Foreign stockholdings; Home bias; Immigration

[☆] We thank the U.S. Census Bureau for their help with the census data, Richard Portes for planting the seed of this idea in us, and Francis Warnock for sharing his data. Two anonymous referees and the editor, Avanidhar Subrahmanyam, dramatically improved the paper. We are grateful for comments from Mihir Desai, Pankaj Jain, Janya Portniaguina, Jonathan Siegel, Chuck Trzcinka, and from seminar participants at the EFA 2003 Meetings, Harvard Business School, Indiana, New York University, Oklahoma, Tulane and UMass (Amherst). All errors are our own.

*Corresponding author. Tel.: +1 812 855 3413.

E-mail address: ubhattac@indiana.edu (U. Bhattacharya).

1. Introduction

Is the United States a melting pot, where foreign-origin persons immediately discard their old identities and forge new ones?¹ Or is it a salad bowl, where foreign-origin persons keep their separate identities and never assimilate?² The purpose of this paper is to examine this old debate using finance data.

We do this by asking whether, when it comes to investing directly or indirectly in a particular foreign country, foreign-origin persons in the United States consider only the “fundamentals” of the country—an implication of the melting pot hypothesis—or, besides the “fundamentals,” foreign-origin persons also care whether this particular foreign country is their country of origin—an implication of the salad bowl hypothesis. Further, if national origin does affect U.S. investments abroad, we ask whether the effect is in the same direction for direct investments abroad (FDI) as it is for indirect investments abroad (equity investments).

The answer to our above research question should interest financial economists who study the investment decision. As it sheds light on a hitherto ignored effect of immigration policy, the answer should also interest policy makers.

A growing literature has started documenting that investments are affected by their familiarity. Familiarity with an investment opportunity in another country is a multi-dimensional measure. One measure of familiarity is the country of residence.³ A second measure of familiarity is geographical proximity. Physical distance seems to matter in the choice of investments.⁴ A third measure of familiarity is language. This is found in some studies to be an important consideration in the decision to invest.⁵ A fourth measure of

¹The phrase “melting pot” was first coined by David, the protagonist in a play by Israel Zangwill (1914)

²The late Senator Daniel Patrick Moynihan popularized this hypothesis. We could not trace the origin of the phrase “salad bowl”.

³Feldstein and Horioka (1980) document a high positive correlation between a country’s savings and its investment rate, which implies that capital flows to domestic investment opportunities, which may not necessarily be the most profitable nor the best hedge against risk. French and Poterba (1991) show that domestic investors hold a disproportionate amount of domestic stocks in their equity portfolios. Lewis (1999), in a lucid survey of the vast “home bias” literature that developed from these seminal papers, distinguishes between the “consumption home bias” (domestic consumption is more correlated with domestic output than is warranted in a model with optimal risk sharing) and the “equity home bias” (domestic investors hold more domestic equity than is warranted in a model with optimal risk sharing). She, as well as many others (see, for example, Pastor, 2000; Karolyi and Stulz, 2003; Glassman and Riddick, 2001), are skeptical about the usual reasons given to explain home bias: explicit barriers to entry (such as ownership restrictions, transactions costs and taxes) or implicit barriers to entry (such as informational advantages or correlation with human capital.)

⁴Petersen and Rajan (2000) study the effect of distance on lending relationships. Calem and Nakamura (1998) develop and test a spatial model for bank branching. Tesar and Werner (1995) show that U.S. investors favor Canadian stocks. Coval and Moskowitz (1999) document that U.S. investment managers hold a disproportionate amount of stocks of firms that have local headquarters. This does not seem a foolish strategy because they earn substantial abnormal returns in these nearby investments (Coval and Moskowitz, 2001). Huberman (2001) documents that shareholders of a Regional Bell Operating Company tend to live in the area which it serves. Portes and Rey (2005) demonstrate that geography dominates the diversification motive when it comes to equity investments abroad. The economic importance of geography extends to other areas besides investments. See Krugman (1991) for the seminal paper, and Neary (2001) for a comprehensive review of the new economic geography literature.

⁵Grinblatt and Keloharju (2001) show that Finnish investors are more likely to hold, buy, and sell the stocks of Finnish firms that not only are located close to the investor, but also communicate in the same language as the investor. Hau (2001) shows that non-German-speaking traders significantly underperform German-speaking traders when trading the 11 largest blue-chip German stocks.

familiarity is race. The controversial practice of redlining has been a focus of many papers.⁶ A fifth measure of familiarity is religion.⁷ A sixth measure of familiarity is national origin.⁸ As the research question in this paper is not home bias—the examination of the choice among domestic investments and foreign investments—but an examination of the choice between various foreign investments, we focus on the sixth measure of familiarity (national origin), we ignore the first measure of familiarity (country of residence), and we control for the second (distance), third (language), fourth (race) and fifth (religion) measures of familiarity.⁹ To summarize, the specific country-specific familiarity variables that we examine in this paper are: number of foreign-origin persons from the country residing in the US, distance of the country from the US, predominant language of the country, predominant race in the country, and the predominant religion in the country.

The “fundamentals” that determine U.S. direct investments in a particular country, as documented in the literature, are: measures of economic development, the size of the government, corporate tax burden, integration with the rest of the world, infrastructure, capital market development, exchange rate risk, and legal environment. The “fundamentals” that determine U.S. equity holdings in a country, as observed in the literature, are: measures of predicted equity holdings assuming that the international capital asset pricing model holds, stock market development, risk-return tradeoffs, integration with the rest of the world, and legal environment.

The following are the two main results of our paper. First, in 1980, 1990 and 2000, but not in 1970, if we hold our proxies for the “fundamentals” constant, we find that the bigger the size of a foreign-origin group from a particular country living in the United States, the more is the level of U.S. direct investments in that country.¹⁰ None of the other familiarity variables matters in *any* year. A panel data regression correcting for country fixed effects—which essentially is a test for correlations between differences rather than levels—corroborates the above result. This suggests the existence of a national origin bias in foreign direct investments. The second main result of our paper is that, in 1997 and 2001, the only two years for which we have data, if we hold our proxies for the “fundamentals” constant, we find that the bigger the income of the foreign-origin group from a particular country living in the United States—especially the one who were born in the US—the more is the level of U.S. stockholdings in that country. None of the other familiarity variables matters. A panel data regression correcting for country fixed effects gives us the same

⁶For example, see Tootell (1996), for an answer to the question whether racial discrimination occurs in mortgage lending in Boston.

⁷The prohibition of usury has led to the development of Islamic banking. See Alkhalifa (1999) for a comprehensive overview. Previous stigma against usury may have led to less protection of creditor rights in Catholic countries (Stulz and Williamson, 2003).

⁸Berger, Klapper, and Udell (2001) document that foreign-owned banks have difficulty in extending relationship loans to opaque small domestic firms. Berger, DeYoung, Genay, and Udell (2000) show that, except for U.S. banks, domestic banks have higher profit margins than foreign banks in five OECD local markets. Choe, Kho, and Stulz (2001) document that domestic Korean individual investors have better information than foreign investors, as do Bhattacharya, Daouk, Jorgenson, and Kehr (2000) for the case of Mexico.

⁹Age and gender are two familiarity variables that we do not examine. We do not have data on age. Barber and Odean (2001) have documented that sex matters in the investment choice; however, we do not have data on gender.

¹⁰We have data about the income of the foreign-origin group only in the 1990 census. We find that this matters, too, in that year.

result. This suggests the existence of a national origin bias in foreign indirect investments as well, though it is not as strong as in the case of direct investments.

To summarize, two different empirical tests (cross-sectional multi-variate regressions and panel data tests with fixed effects corrections) on each of two different types of data (direct U.S. investments abroad and indirect U.S. investments abroad) give us a similar result: U.S. investments in a foreign country are positively affected by the size of the foreign-origin group from that country living in the United States.

The paper is organized as follows. In Section 2 we discuss how, given our data limitations, we attempt to disentangle the rational from the behavioral motives for investing abroad. Section 3 describes our data. We precisely define each variable and we motivate from past literature or common sense why we are using this variable. The documentation of the sources from which we obtained this variable as well as more details about the variable can be found in the Appendix. The Appendix also contains a table giving descriptive statistics on all the independent variables. Section 4 tests the null hypothesis that only “fundamentals” determine U.S. direct investments abroad. Section 5 tests the null hypothesis that only “fundamentals” determine U.S. stockholdings abroad. We conclude in Section 6. Here we lay out the limitations of our research, and point to further research directions.

2. Rational versus behavioral motives for investing abroad

The Merriam-Webster dictionary defines “familiarity” as a “state of close relationship.” A close relationship with an investment opportunity naturally leads to more information about this investment opportunity. A preference for familiar investments may, therefore, be just a preference for investing in assets that one has more information about. This is rational choice. A close relationship with an investment opportunity may also lead to emotional reactions. A preference for familiar investments in this case may, therefore, be an expression of loyalty. This is a behavioral choice. So we have to test, if we can, whether investments abroad are “rational” or “behavioral.”

A second issue we need to address is that our direct investments data are aggregate investments made by U.S. multinationals, and, therefore, reflect decisions made by U.S. multinationals. Similarly, our indirect investments data are aggregate equity outflows going abroad from U.S. institutional investors, and, therefore, reflect decisions made by U.S. institutional investors. So we have to explain how, if at all, the motives of foreign-origin persons, be they rational or behavioral, may affect the above institutional decisions.

When it comes to indirect foreign investments, we hypothesize that foreign-origin persons influence the decision of U.S. institutional investors (mutual funds, hedge funds, etc) in the following ways. First, foreign-origin person can buy equity directly in some countries, in which case the foreign-origin person and the investor is the same agent. Alternatively, foreign-origin persons can buy single country funds or multi-country funds in which their country of origin has a big weight. Foreign-origin persons may also be important decision makers in the foreign equity desks of institutional investors, in which case the influence of foreign-origin persons would be indirect. As for motives, foreign-origin persons may be partial to their country of origin for rational reasons—foreign-origin persons have an informational advantage over others about their country of origin—or for behavioral reasons—many foreign-origin persons have strong family and

cultural ties to their country of origin. **Portes and Rey (2005)** make a convincing empirical case that the rational reason—information asymmetries—is important.

We test for the rational/behavioral motives for indirect investments by further disaggregating our sample as foreign-born foreign-origin persons (henceforth labeled as “foreign-borns”) or native-born U.S. citizens who claim foreign ancestry (henceforth labeled as “native-borns”). We have this data for 1997 and 2001. Curiously, we find that there is a significant correlation between equity investments and the number of native-borns (weighted by income,) but no correlation between equity investments and the number of foreign-borns (weighted by income). As foreign-borns know more about their country of origin than native-borns, this suggests that the “behavioral” motive is stronger than the “rational” motive when it comes to equity investments.

When it comes to direct foreign investments, we hypothesize that foreign-origin persons influence the decisions of U.S. firms in the following ways. Some foreign-origin persons own the firms themselves, and their decisions are the firms’ decisions. A recent story in the *Wall Street Journal* (March 20, 2007), titled “Foreign-origins Gain Edge Doing Business Back Home,” illustrates how foreign-origin entrepreneurs use their knowledge of the local culture and their personal connections to open doors in their country of origin. These are “rational” motives. The second, and in our view, the predominant means by which foreign-origin persons influence the decisions of firms they work for, is indirect. U.S. multinationals send scouting teams to a foreign country before they invest in that country, and there is anecdotal evidence to suggest that foreign-origin persons from that country are included in the scouting teams. This makes sense because these persons may have a communication and negotiating advantage over others in dealing with natives from their country.¹¹

We can test for the rational/behavioral motives for direct investments by further disaggregating foreign-origin persons as foreign-borns or native-borns as before. However, this data exists only for one census year (2000) out of our four census years. In this year, we find that there is a significant correlation between equity investments and the number and income of native-borns, but no correlation between equity investments and the number or income of foreign-borns. As foreign-borns know more about their country of origin than native-borns, this suggests that the “behavioral” motive is stronger than the “rational” motive for equity investments in 2000.

An alternative hypothesis about why capital flows and immigration could be linked is because of exogenous shocks to the supply and demand for human capital. For example, the explosion of demand for IT professionals in the late 1990s may have led to immigration from India as well as an increase in capital flows to India. In the early 1990s, the ex-communist countries liberalized, and this led to an increase in the supply of immigrants

¹¹We offer three case studies to corroborate this point. The first case is about Coca-Cola in China. Coca-Cola sent a Chinese-American, Steve Chan, to fight Pepsi in China in the 1990s. We quote: “In a culture where power was essentially fluid, *guanxi*, and whether one had *guanxi*, often determined the success or failure of ventures. It was therefore unsurprising that Steve Chan, China Region Manager for Coca-Cola China, spent almost 50% of his time cultivating relations with Chinese officials...“I sang a karaoke with the Vice Minister [of Light Industry], a duet. Karaoke is part of business in China today” (HBS Case 9-795-186). The second case is about Toys “R” Us in Japan. Toys “R” Us sent American-born Robert Nakasone to Japan to negotiate with a potential partner, Den Fujita (HBS Case 9-796-077). The third case is Psi, a consulting and advisory firm based in New York and New Delhi. Its founder is an Indian-American, Manu Bammi, who assisted several U.S. multinationals in developing new ventures, forging strategic alliances, and acquiring companies in India (www.smartanalyst.com).

from these countries coming to the U.S. and capital outflows from the U.S. to these countries. We do have a test for this alternative hypothesis using an instrumental variable: the number and income of native-borns. Though it is possible, as discussed above, that capital flows are linked to the number of foreign-origin persons because of unobservable exogenous supply or demand shocks, the link should be to foreign-borns and not to native-borns. Curiously, we find only a link to native-borns for equity investments and for FDI in the year 2000.

The other ways we try to address endogeneity are standard. First, we try to control for all variables that affect both investment and emigration. So, in all our multi-variate cross-sectional regressions (Tables 3 and 6), we have controlled for all variables like liberalization that we think can affect both sides of the regression equation. The number of control variables considered for the direct investments regression is 26, whereas the number of control variables considered for the indirect investments regression is 15. Second, we employ panel data tests (Tables 4 and 7) with fixed effect corrections. As these panel data tests check for correlations between changes rather than levels, endogeneity is not an issue as long as the omitted variables do not change much over the sample period. Our results for both direct and indirect investments continue to hold after these tests.

3. Data

We list below the salient points about our data items. Details on them are relegated to the Appendix. The Appendix also contains a table that gives descriptive statistics of all the independent variables that we use.

3.1. *Dependent variables*

US agents can invest abroad in two ways. U.S. multinationals can create or expand production capacity abroad (direct investments) or U.S. institutions or U.S. individuals can buy shares of foreign firms (indirect investments). We would like to examine the effect of familiarity variables on both these variables.

Direct investments by U.S. in each country is FDI data obtained for the years 1970, 1980, 1990 and 2000 from the Bureau of Economic Activity, U.S. Department of Commerce. Indirect investments by U.S. in country i is obtained from surveys of institutional investors carried out by the U.S. Treasury Department and the Board of Governors of the Federal Reserve System in 1997 and 2001.

3.2. *Independent variables*

3.2.1. *Familiarity variables*

As we only have data on the level of aggregate U.S. direct or indirect investment in each foreign country, but do not have this data at the disaggregated level of the U.S. individual investor or the U.S. institutional investor, we assume the existence of a representative U.S. investor. This representative U.S. investor has English as his first language, is white, is Christian, and has the blood of many countries flowing in his veins. Given this assumption of the U.S. representative investor, we estimate familiarity variables for each country with respect to the United States as follows.

Physical proximity is the distance in miles from New York City to the country's main stock market.¹² The language variable is 1 if English is one of the listed official languages of the country and 0 otherwise, the race variable is the proportion of whites in the population of the country, the religion variable is the proportion of Christians in the population of the country, and the national origin variable is the number or the total income of foreign-origin persons from a particular foreign country living in the United States.¹³

3.2.2. *Control variables proxying for “fundamentals” determining direct investment*

There is a vast literature using ad hoc linear cross-country regressions to search for the determinants of foreign direct investments. Chakrabarti (2001) surveys the literature and provides a judgment on what works. We follow his lead, and use some other measures. Direct investments should increase with economic development. We use four measures: average GDP in the last three years, average GDP growth in the last three years, average per capita income in the last three years, and average per capita income growth in the last three years. Direct investments should decrease with the size of the government. We use one measure: government consumption as a percentage of GDP. Direct investments should decrease with the corporate tax burden. We use three measures: percentage in total tax revenues from taxes on income, profit and capital gains, percentage in total tax revenues from import duties, and percentage in total tax revenues from export duties. Direct investments should increase with integration with the rest of the world. We use five measures: exports plus imports divided by GDP, which is defined as openness in Bekaert and Harvey (1995), a capital controls index constructed as in Bartolini and Drazen (1997), a liberalization dummy constructed from Bekaert and Harvey (2000), surveys of country credit ratings given by international bankers, and time since liberalization. Direct investments should increase with better infrastructure. We use five measures: air freight, roads freight, percentage of paved roads, telephone lines per 1000, and telephone lines per employee.¹⁴ Direct investments should increase with capital market development. We use six measures: credit to private sector divided by GDP (Beck, Levine, and Loayza, 2000), size of stock market, size of stock market divided by GDP, average size of a listed firm, volume of trade in the stock market, and volume of trade in the stock market divided by size of the stock market. Direct investments should decrease with exchange rate risk. We use one measure: standard deviation of the real exchange rate with respect to the dollar (Goldberg and Kolstad, 1995). Direct investments should increase with the quality of the legal environment. We use one measure: legal origin (La Porta Lopez-de-Silanes, Shleifer, and Vishny, 1998).

¹²We experimented with time zones as an alternative measure of distance. The results, being qualitatively similar, are not reported here.

¹³The question on national origin is a question on the long form of the U.S. census. It need not be answered. We do not know the proportion of U.S. residents who answer this question. This, of course, leads to a selection bias. This selection bias works in our favor. A respondent will answer this question only if he/she feels a strong sense of belonging to the national origin group. This means that our estimates the size of the national origin groups are likely to be biased downwards. We would like to use both numbers as well as total income of foreign-origin persons in all our tests because it could be argued that though the number may be high, their income is low, and so they cannot invest much in their country of origin. Unfortunately, we can obtain income by national origin only for the 1990 census, and for 1997 and 2001 from the Current Population Surveys.

¹⁴A previous version of the paper used rail freight as well. We dropped this variable because our data source (World Bank) does not document this any more.

3.2.3. *Control variables proxying for “fundamentals” determining indirect investment*

There is a vast literature on equity home bias that tries to explain why investors hold a disproportionate amount of home-country equity in their stock portfolios. Lewis (1999) and Karolyi and Stulz (2003) provide comprehensive surveys. We use the measures that have been tried before. Foreign stockholdings should increase with predicted equity holdings. If the world equity portfolio is the market portfolio, then predicted holdings of the U.S. in a country would be the weight of the country's stock market in the total world stock market multiplied by total U.S. stockholdings (Ahearne, Grier, and Warnock, 2004). As a matter of fact, the deviations of actual equity holdings from predicted equity holdings is defined as the home bias. The rest of the measures that follow attempt to explain this deviation. First, foreign stockholdings should increase with stock market development. We use five measures: size of stock market divided by GDP, average size of a listed firm, number of listed firms, volume of trade in the stock market, and volume of trade in the stock market divided by size of the stock market. Second, foreign stockholdings should be affected by risk-return tradeoffs. We use four measures: foreign stockholdings should increase with the ratio of mean stock market dollar return to standard deviation of stock market dollar returns (Ahearne, Grier, and Warnock, 2004) if there is returns chasing behavior; foreign stockholdings should decrease with the correlation of stock market return with U.S. stock market return because of less diversification potential; foreign stockholdings should decrease with the standard deviation of the real exchange rate because purchasing power parity does not hold; foreign stockholdings should increase with last year's return if there is crude returns chasing behavior. Third, foreign stockholdings should increase with integration with the rest of the world. We use three measures: exports plus imports divided by GDP, which is defined as openness in Bekaert and Harvey (1995), a capital controls index constructed as in Bartolini and Drazen (1997), and a foreign ownership restrictions on local equity index constructed as in Edison and Warnock (2003).¹⁵ Fourth, foreign stockholdings should increase with the quality of the legal environment. We use two measures: legal origin (La Porta, Lopez-de-Silanes, Shleifer, and Vishny, 1998) and a corruption index (Transparency International (1997 and 2001)).¹⁶

There is a lot of missing data in our original sample. This would lead to different sample sizes in different tests. To minimize this, we employ the following strict screening procedure to our original sample. For the FDI tests, we only keep those countries for which we have FDI data as well as national origin data in the census years 1970, 1980, 1990 and 2000. This gives us 48 countries. For the portfolio tests, we only keep those countries

¹⁵We do not use the liberalization dummy because in 1997, the only year for which we have data, all countries in our sample were liberalized.

¹⁶A previous version of the paper used a rule of law index (Kaufmann, Kraay, and Zoido-Lobaton, 1999); a shareholder rights index from Bhattacharya and Daouk (2002), who constructed it from data in La Porta, Lopez-de-Silanes, Shleifer, and Vishny, 1998; and a measure of accounting transparency from Bhattacharya, Daouk and Welker (2003). We dropped these three variables because we did not have data on them for 2001. Also, we changed the source of our Corruption Index from Kaufmann, Kraay, and Zoido-Lobaton (1999) to Transparency International, because Transparency International had this data for both 1997 and 2001. Also, since we now have surveys of U.S. foreign equity portfolio holdings for both 1997 and 2001, we dropped the 1997 private data given to us by Froot, O'Connell, and Seasholes (2001), a data set which had given us qualitatively similar results in a previous version of the paper.

for which we have U.S. foreign equity holdings data as well as national origin data in the survey years 1997 and 2001. This gives us 59 countries.

Descriptive statistics of the 48 countries in the FDI sample and the 59 countries in the portfolio sample are given in Table 1.

Table 1

U.S. investments abroad and familiarity variables

Column 1 shows U.S. direct investments (U.S. FDI) in the country measured at historical cost in the year 2000. Source: Bureau of Economic Activity, U.S. Department of Commerce. Column 2 shows U.S. indirect investments (U.S. foreign stockholdings) in the country in the year 2001. Source: U.S. Treasury Department and the Board of Governors of the Federal Reserve System. Column 3 shows the total number of members of the national origin group living in the U.S. in the year 2000. Source: U.S. Bureau of the Census. Column 4 shows physical distance in miles between New York and the city in the foreign country where the main stock exchange is located. Source: <http://www.indo.com/distance/>. Column 5 shows language of country, which takes the value of 1 if the CIA lists English as one of the major languages spoken in the country, and 0 otherwise. Source: CIA World Factbook 2001. Column 6 shows race of country, defined as the percentage of total population of a country that is described by CIA to be of white or Caucasian race. If percentage is not given, but if the CIA lists white or Caucasian to be the majority race, then the country receives 1, and 0 otherwise. Source: CIA World Factbook 2001. Column 7 shows religion of country, defined as the percentage of total population of a country that belongs to any Christian affiliated religion. If percentage is not given, but if the CIA lists Christianity as the majority religion, then the country receives 1, and 0 otherwise. Source: CIA World Factbook 2001.

Country	FDI from U.S. in 2000 (billion USD) (1)	U.S. stock holdings in 2001 (billion USD) (2)	Total immigrants in the U.S. in 2000 (thousands) (3)	Distance from the U.S. (miles) (4)	Language (5)	Race (6)	Religion (7)
Argentina	17.488	0.744	122.156	5253	1	0.97	0.94
Australia	34.838	37.112	61.049	9935	1	0.92	0.89
Austria	2.872	1.204	436.361	4233	0	1	0.83
Bangladesh		0.004	57.412	7880	1	0	0
Belgium	17.973	8.415	225.399	3668	0	0.89	1
Brazil	36.717	21.801	163.608	4784	1	0.55	0.8
Canada	132.472	89.591	481.871	347	1	0.66	0.82
Chile	10.052	1.917	76.836	5102	0	0.95	1
China	1.114	2.370	2865.232	7387	0	0	0.01
Colombia	3.693	0.150	570.150	2480	0	0.2	0.9
Costa Rica	1.716		85.991	2204	1	0.94	0.92
Czech Republic	1.228	0.444	1262.527	4094	0	0.99	0.48
Denmark	0.527	7.533	855.921	3850	1	1	0.98
Ecuador	0.832	0.018	327.320	2822	0	0	0.95
Egypt	1.998	0.340	142.832	5621	1	0.01	0.06
El Salvador		0.002	655.165	2072	0	0.09	0.86
Finland	1.342	51.307	435.484	4123	0	1	0.9
France	42.628	112.205	4872.360	3635	0	1	0.92
Germany	55.508	72.200	30199.866	3862	0	1	0.72
Ghana		0.207	49.944	5120	1	0.015	0.63
Greece	0.795	2.810	942.928	4934	1	0.98	0.98
Guatemala			516.941	2056	0	0.02	1
Honduras	0.399		297.763	2000	0	0.01	1
Hungary	0.192	1.702	1398.724	4376	0	1	0.925
India	2.379	6.897	1899.599	7301	1	0	0.023
Indonesia	8.904	1.526	63.073	10053	1	0	0.08

Table 1 (continued)

Country	FDI from U.S. in 2000 (billion USD) (1)	U.S. stock holdings in 2001 (billion USD) (2)	Total immigrants in the U.S. in 2000 (thousands) (3)	Distance from the U.S. (miles) (4)	Language (5)	Race (6)	Religion (7)
Ireland	35.903	28.374	19282.096	3185	1	1	0.941
Israel	3.735	13.333	96.338	5683	1	1	0.021
Italy	23.484	33.686	12906.708	4026	0	1	1
Jamaica		0.001	681.186	1570	1	0.002	0.653
Japan	57.091	170.714	847.514	6760	0	0	0.007
Jordan		0.061	39.734	5731	1	0	0.06
Kenya		0.005	17.336	7360	1	0	0.66
Latvia		0.013	87.564	6885	1	0	0.49
Lebanon		0.038	440.279	4210	1	0	0.39
Lithuania		0.003	659.992	5620	1	0	0.3
Luxembourg	27.849		26.378	3777	0	1	1
Malaysia	0.791	2.578	18.566	4333	1	0	0
Mexico	39.352	26.279	21607.506	2086	0	0.09	0.95
Morocco		0.037	38.923	2086	0	0.009	0.01
Netherlands	115.429	112.751	2554.833	3652	0	1	0.51
New Zealand	4.271	2.004	16.628	3652	0	0.698	0.535
Norway	4.379	7.906	3241.637	3683	0	1	0.89
Pakistan		0.086	204.309	3683	1	0	0
Peru	0.313	0.452	270.119	3635	0	0.15	0.9
Philippines	3.638	1.344	1848.312	8509	1	0	0.92
Poland	3.884	1.197	1177.112	8509	0	1	0.92
Portugal	2.664	3.819	916.581	3379	0	1	0.94
Romania		0.003	367.310	3379	0	1	0.94
Russia	1.147	4.613	2652.214	4770	0	1	0.82
Singapore	24.133	21.376	3.676	4681	1	0	0.146
Slovakia		0.003	797.764	9534	0	1	0.878
South Africa	3.562	6.714	45.569	4266	1	0.096	0.812
Korea, Rep.	8.968	29.537	1081.091	6885	1	0	0.49
Spain	21.236	32.455	701.604	3591	0	1	0.99
Sweden	25.969	24.274	2436.832	3934	0	1	0.87
Switzerland	55.377		536.503	3939	0	0.94	0.861
Thailand	5.824	1.916	150.283	8664	1	0	0.005
Trinidad & Tobago		1.331	149.741	2209	1	0.006	0.437
Turkey	1.826	2.269	99.190	5026	0	1	0
United Kingdom	230.762	350.014	28222.890	3471	1	0.972	0.9065
Ukraine		0.002	892.922	4676	0	1	1
Uruguay		0.004	18.804	5316	0	0.88	0.68
Venezuela	10.531	0.348	110.690	3635	0	1	0.98

4. Effect of National Origin on U.S. Direct Investments Abroad

Table 2 documents the *unconditional* effect of the familiarity variables and the other control variables on U.S. direct investments abroad in 1970, 1980, 1990 and 2000. We run uni-variate regressions, where the dependent variable is U.S. direct investment, and there is just one independent variable, which could be any of the familiarity or control variables. As the independent variables have widely varying scales, all reported regression coefficients

Table 2

Effect of national origin on U.S. direct investments abroad (results from uni-variate OLS regressions)

The dependent variable is U.S. level of direct investments (U.S. FDI) in a foreign country in billions of dollars. Independent variables are all scaled by their standard deviation. Description and sources of all variables are given in Section 3 and in the Appendix of the paper. There is a blank in 1970 for exchange rate risk, because our sample countries were on fixed exchange rates then; there is a blank in 2000 for liberalization because our sample countries were all liberalized by then; the other blanks exist because of missing data. Standard errors are White-heteroskedasticity corrected. * Significance at the 10% level. ** Significance at the 5% level. *** Significance at the 1% level. *Significance at the 10% level after the Bonferoni correction. ** Significance at the 5% level after the Bonferoni correction. *** Significance at the 1% level after the Bonferoni correction.

Independent variables	1970			1980			1990			2000		
	Coefficient	N	R ²	Coefficient	N	R ²	Coefficient	N	R ²	Coefficient	N	R ²
B1. Familiarity variables												
B11. Physical distance	−0.86	37	0.053	−2.01	37	0.048	−24.97	37	0.018	−4.80	37	0.011
B12. Language	815.24	37	0.052	1.85	37	0.042	2.94	37	0.029	0.75	37	0.028
B13. Race	307.87	37	0.008	1.18	37	0.002	2.48	37	0.022	8.18	37	0.032
B14. Religion	160.62	37	0.002	0.21	37	0.005	−0.40	37	0.0011	−2.06	37	0.002
B15. National origin												
B151. Total	1.08*	37	0.279	10.22***	37	0.210	8.92**	37	0.189	23.74**	37	0.288
B2. Control variables proxying for “fundamentals”												
B21. Economic development												
B211. Average GDP	−0.07	36	0.001	3.71***	37	0.191	7.79*	36	0.214	17.22*	37	0.142
B212. Average GDP growth	−666.31	36	0.033	−1.17	37	0.014	1.38	36	0.009	−3.17	37	0.006
B213. Average per capita income	0.62**	36	0.029	2.69**	37	0.086	5.18*	36	0.094	12.78***	37	0.077
B22. Size of government												
B221. Govt consumption/GDP	823.26	36	0.056	1.58	36	0.029	4.24	36	0.063	−0.88	37	0.002
B23. Corporate tax burden												
B231. Taxes				1.65	36	0.031	3.70	33	0.044	8617.05**	37	0.036
B232. Import duty				−1.94***	37	0.044	−443.26***	32	0.064	−13.43*	24	0.068
B233. Export duty				−0.95	36	0.011	−3.53***	32	0.039	−6.01	24	0.013
B24. Integration with the world												
B241. (Exports + Imports)/GDP	−38.36**	36	0.013	−110.75*	37	0.016	−2.16*	36	0.013	1.97	37	0.004
B242. Capital controls index	−1013.25	37	0.078	−3.40*	37	0.136	−9.84***	36	0.339	−7.79**	37	0.029
B243. Liberalization dummy	0.91*	30	0.055	3.09***	30	0.094	−2.63**	37	0.023	See above	37	0.192
B244. Country credit rating				4.75***	34	0.249	89217.23***	36	0.271	19.91**		

B25. Infrastructure												
B251. Air freight	2.09***	36	0.323	4.85***	36	0.273	10219.94***	37	0.368	22710.02**	36	0.243
B252. Roads freight						0.162	8.44***	24	0.182	16.03*	23	0.086
B253. Percentage of paved roads							5114.52*	37	0.069	14322.51*	37	0.098
B254. Telephone lines per 1000	13912.12	35	0.136	3.71**	37		7007.79***	37	0.173	17174.37**	37	0.141
B255. Telephone lines per empl.				0.71	33	0.006	4410.20***	37	0.068	386.16	37	0.007
B26. Capital market development												
B261. Credit to private sector/GDP							7146.42***	36	0.181	23857.34***	37	0.273
B262. Market cap of stock market	126.56	35	0.001	1.65	37	0.032	6.69	32	0.143	5.09***	37	0.012
B263. Mar. cap of stock mar./GDP							8800.85**	31	0.245	−7.04	37	0.023
B264. Av. mar. cap of listed firm							7.58***	32	0.184	4.99***	37	0.012
B265. Volume of trade							7941.80*	30	0.198	15584.93*	35	0.112
B266. Volume of trade/Mar. Cap							1.66	26	0.007	4.10	34	0.007
B27. Exchange rate risk												
B271. St. dev. of real exch. rate	See above			−1.06	37	0.013	−2.81	34	0.027	−1.58	37	0.001
B28. Legal environment												
B281. Legal origin (Common vs Others)	1377.14	37	0.145	3.98	37	0.188	6.78	37	0.178	17.40	37	0.147

in Table 2 are scaled to reflect the impact of one standard deviation change in the independent variable.

The *only* familiarity variable that is statistically significant (p -values of 10% or less) in *all* four years is total size of the national origin group. As a matter of fact, *none* of the other familiarity variables is significant in *any* year. We have disaggregated data for foreign-origin persons only for 1990. Further analysis on this disaggregated data in 1990, which is not reported in Table 2, reveals that *only* the number of native-borns has a statistically significant impact on FDI.

Of the control variables, the “fundamentals” of a country that statistically significantly (p -values of 10% or less) influence U.S. foreign direct investments in *all* four years are one measure of economic development (average per capita income) and one measure of infrastructure (air freight). Other variables are significant in some years, but not in other years. Unconditionally, countries which are more economically developed, which are more integrated with the rest of the world, which have better infrastructure, and which have more well-developed credit markets, attract more U.S. foreign direct investments.

Comparing the magnitude of the scaled regression coefficients in Table 2, the effect of the size of the national origin group is seen to have an economically significant effect on FDI. Its effect is comparable to the effect of average per capita income of that country.

It would be tempting to emphasize our finding that the only familiarity variable that seems to affect FDI in all the four census years is the size of the national origin group. It could be that our results are driven by the huge number of foreign-origin persons from Germany, Ireland, Italy, Mexico and UK. We dropped these countries. The statistical significance of the size of the foreign-origin group in all four years remains. Second, since we ran so many uni-variate regressions, we could have found statistical significance by chance. Therefore, an adjustment to standard errors needs to be made. There are many ways of doing this, and we choose the harshest one—the Bonferoni correction. Using the Bonferoni correction, we find that the national origin variable is significant in Table 2 only in 1980. Finally, the high R^2 , of the order of 20%, may be misleading. It may be coming about because the national origin is correlated with some omitted effect. It is for this reason that we now turn to multi-variate tests that attempt to control for omitted effects.

Table 3 checks whether the size of the national origin group continues to remain significant after we have controlled for the “fundamental” determinants of foreign direct investments. We first reduce the number of “fundamental” control variables by extracting their major principal components in each census year. We obtain 4, 6, 5 and 6 significant principal components for the years 1970, 1980, 1990 and 2000, respectively. Table 3 gives results of cross-sectional multi-variate regressions for 1970, 1980, 1990 and 2000, where the dependent variable is U.S. direct investment, and the independent variables are the above extracted major principal components from the control variables and the size of the national origin group. We find that the total size of the national origin group statistically significantly influences (p -values of 10% or less) U.S. direct investment in a foreign country in three years, 1980, 1990, and 2000. Though we do not report it in Table 3, we find that after adding the controls, *none* of the other familiarity variables is significant in any year.

We do three more tests in Table 3. We run the cross-sectional tests in each year in Table 3 with and without the size of the national origin group variable, and ask whether the R-squares increase in a statistical sense if we include the national origin variable. This is a F-test. We find that the addition of the national origin variable statistically improves fit for the years 1980 and 1990, but not for the years 1970 and 2000. We also find that the R^2

Table 3

Effect of national origin on U.S. direct investments abroad (results from multi-variate OLS regressions)

The dependent variable is U.S. level of direct investments (U.S. FDI) in a foreign country in billions of dollars. The independent variables are obtained as follows. In the first step, we reduce the number of control variables in Table 2 by extracting major principal components from all the control variables except the familiarity variable. In the second step, we run a multi-variate OLS regression using the extracted principal components and one familiarity variable (Total) as the independent variable. Though we do not report it in the above table, we also ran this multi-variate OLS regression using the other familiarity variables in turn. The results from those regressions show that no other familiarity variable becomes significant after we control for principal components. Standard errors are White-heteroskedasticity corrected. *Significance at the 10% level. ** Significance at the 5% level. *** Significance at the 1% level.

Independent variables		1970		1980		1990		2000	
		Coefficient estimate	p-Value	Coefficient estimate	p-Value	Coefficient estimate	p-Value	Coefficient estimate	p-Value
Constant		1.03**	0.045	6.27***	0.008	6.68**	0.024	20.58**	0.023
Familiarity variable	B151. Total (National Origin)	1.36E–6	0.282	5.75E–7*	0.056	1.81E–6**	0.022	1.95E–6*	0.083
Principal components extracted from the control variables	Principal component one	0.53**	0.032	1.51	0.526	2.38***	0.001	6.56	0.268
	Principal component two	0.29	0.309	0.75**	0.018	1.68**	0.017	–10.36***	0.009
	Principal component three	0.89*	0.040	0.62	0.426	2.26	0.120	0.22	0.268
proxying for “fundamentals” in Table 2	Principal component four	–0.37	0.211	4.51	0.525	–4.94*	0.059	5.13	0.164
	Principal component five			–1.35*	0.056	–4.26*	0.096	2.84	0.598
	Principal component six			–2.56	0.279			2.57	0.631
<i>N</i>		27		26		24		23	
<i>R</i> ²		0.398		0.575		0.7155		0.5131	
<i>R</i> ² with B151. Total excluded		0.298		0.426		0.594		0.386	
<i>F</i> test of inclusion of B151. Total		2.849		4.416**		4.481**		2.892	

for a regression that contains only the “fundamental” principal components is significant, which implies that the “fundamental” principal components do a good job in explaining FDI. Third, we ask whether the effect of the size of the national origin group is economically significant for FDI. The 2000 coefficient in Table 3 is 1.95E-6, which implies that, everything else constant, every additional person in a national origin group is associated with a FDI increase of USD 1,950 in their country of origin. To put this in perspective, if a country in the 25th percentile in foreign-origin persons (Trinidad and Tobago in 2000, with 0.15 million foreign-origin persons in the US) moved up to be in the 75th percentile in foreign-origin persons (Philippines in 2000, with 1.85 million foreign-origin persons in the US), FDI from U.S. would increase by about USD 3.3 billion, everything else constant.

All our tests till now are cross-sectional. We now do a panel data test correcting for country fixed effects. Table 4 gives the results from this test. The dependent variable is U.S. direct investment. In Model A, the independent variables are the independent variables from Table 2 that have no missing observation in either of the four years, 1970, 1980, 1990 and 2000, for a set of countries. This reduces the number of independent variables without any obvious variable selection bias that might affect our results. Table 4, Model A, shows that the effect of the size of the national origin group continues to remain significant. In

Table 4
Effect of National origin on U.S. direct investments abroad (results from panel regression with country-fixed effects)

The dependent variable in both Models A and B is U.S. level of direct investments (U.S. FDI) in a foreign country in billions of dollars. In Model A, the independent variables are all independent variables from Table 2 which have all observations for years, 1970, 1980, 1990, and 2000, for a set of countries. However, we do not include race, religion, language, distance and legal origin variable, because they do not change over time. Average per capita income, Credit to private sector/GDP and Telephone lines per 1000 control variables are very correlated, with Pearson correlation coefficients above 0.7. Therefore, we include only Credit to private sector/GDP in our regression model A. The final sample in Model A contains observations for 33 countries for which we have all control variables for all census years. In Model B, the independent variables are all control variables in Table 2 that were found to be statistically significant in all four years, 1970, 1980, 1990 and 2000. As this selection method may bias estimates, the coefficients and p-values reported above for Model B are estimated with a bootstrap method with re-sampling. The number of repetitions in the bootstrap was 10,000. Standard errors in both models are White-heteroskedasticity corrected. * Significance at the 10% level. ** Significance at the 5% level. *** Significance at the 1% level.

Independent variables	Model A		Model B	
	Coefficient estimate	p-Value	Coefficient estimate	p-Value
B151. Total (national origin)	2.1E-6**	0.043	1.5E-6***	0.000
B211. Average GDP	-4.25E-12	0.962		
B212. Average GDP growth	-0.59	0.128		
B213. Average per capita income			1.07E-4**	0.048
B221. Govt consumption/GDP	0.45	0.162		
B241. (Exports + Imports)/GDP	-1.13*	0.069		
B242. Capital controls index	-6.58	0.975		
B251. Air freight	0.01*	0.071	0.01***	0.000
B261. Credit to private sector/GDP	0.10	0.139		
Number of countries in sample		33		36
R ²		0.5787		0.6478
F test of inclusion of B151. Total		26.585***		19.116***

Model B, the independent variables are the independent variables from Table 2 that were found to be statistically significant in all the four years, 1970, 1980, 1990 and 2000. As this method for reducing the number of independent variables may bias estimates, the coefficients and p-values reported for Model B are estimated with a bootstrap method with re-sampling. Table 4, Model B, shows that the effect of the size of the national origin group continues to remain significant.

We redo both Model A and Model B panel data tests with country fixed effects with and without the size of the national origin group variable, and ask whether the R-squares increase in a statistical sense if we include the national origin variable. We find that the addition of the national origin variable statistically improves fit in both models in Table 4. This provides further evidence that countries with larger emigrant groups in the U.S. attract more U.S. foreign direct investments.

5. Effect of National Origin on U.S. Indirect Investments Abroad

Table 5 documents the *unconditional* effect of the familiarity variables and the other control variables on U.S. indirect investments abroad in 1997 and 2001. We run univariate regressions, where the dependent variable is U.S. foreign stockholdings in 1997 and 2001, and there is just one independent variable, which could be any of the familiarity or control variables. As the independent variables have widely varying scales, all reported regression coefficients in Table 5 are scaled to reflect the impact of one standard deviation change in the independent variable.

None of the familiarity variables is found to be significant in any year, except total income of the national origin group, particularly the total income of the native borns. These two familiarity variables are statistically significant (*p*-values of 10% or less) for *both* 1997 and 2001. The signs of the coefficients tell us that, unconditionally, countries with economically stronger emigrant groups in the United States attract more U.S. shareholdings.

Of the control variables, the “fundamental” that is really important is predicted equity holdings. This seems to explain more than 80% of the variation in actual U.S. stockholdings in a foreign country. However, it does not explain everything, nor is the coefficient on it equal to one, which implies that there are factors other than the ones hypothesized by the international capital asset pricing model that drives actual U.S. stockholdings abroad. What are these factors? From our tests, the other factors that statistically significantly (*p*-values of 10% or less) influence U.S. foreign stockholdings in *both* the years are two measures of stock market development (market capitalization of stock market as a proportion of GDP and volume of trade). The signs of all the coefficients, except for the correlation variable, are as expected. Unconditionally, countries which have more developed stock markets, countries which have a higher ratio of mean stock market return to standard deviation of stock market return, countries which are more integrated with the rest of the world, and countries which have a better legal environment, attract more U.S. shareholdings.

Comparing the magnitude of the scaled regression coefficients in Table 5, the effect of the size times income of the national origin group is seen to have an economically significant effect on equity investments abroad. Its effect is comparable to the effect of predicted equity holdings.

Table 5

Effect of national origin on U.S. indirect investments abroad (results from uni-variate OLS regressions)

The dependent variable is U.S. level of stockholdings in a foreign country (U.S. foreign stockholdings) in billions of dollars. Independent variables are all scaled by their standard deviation. Description and sources of all variables are given in Section 3 and in the Appendix of the paper. Standard errors are White-heteroskedasticity corrected. * Significance at the 10% level. ** Significance at the 5% level. *** Significance at the 1% level. *Significance at the 10% level after the Bonferoni correction. **Significance at the 5% level after the Bonferoni correction. ***Significance at the 1% level after the Bonferoni correction.

Independent Variables	1997			2001		
	Coefficient estimate	N	R ²	Coefficient estimate	N	R ²
B1. Familiarity variables						
B11. Physical distance	−644.71	59	0.028	−705.90	59	0.016
B12. Language	−1.33	59	0.001	1.15	59	0.000
B13. Race	7.10	59	0.035	5.77	59	0.011
B14. Religion	3.24	59	0.007	1.26	59	0.000
B15. National origin						
B151. Total	8.60	59	0.053	7.94	59	0.024
B151a. Total × Income	14.78**	59	0.153	16.17*	59	0.086
B152. Foreign born	3.67	59	0.009	3.34	59	0.003
B152a. Foreign born × Income	5.84	59	0.023	8.05	59	0.021
B153. Native born	9.42	59	0.061	8.89	59	0.026
B153a. Native born × Income	16.12***	59	0.181	17.59**	59	0.104
B3. Control variables proxying for “fundamentals”						
B31. Predicted equity holdings	35.12***	59	0.855	49.53***	59	0.826
B32. Stock market development						
B321. Market cap of stock market/GDP	206765.49**	59	0.296	302336.59**	59	0.302
B322. Average market cap of listed firm	25.64***	59	0.450	2893.28	59	0.281
B323. Number of listed firms	1344.41	59	0.115	2053.69	59	0.142
B324. Volume of trade	154400.62*	59	0.163	270033.26***	59	0.241
B325. Volume of trade/Market cap	257.46	59	0.005	117805.59**	59	0.046
B33. Risk-return						
B331. Mean stock market return/Standard dev. of return	9.64	55	0.061	1272.73*	57	0.009
B332. Correlation of market return with U.S. return	10.96***	55	0.077	−5.47	57	0.009
B333. Standard dev. Of real exchange rate	−3.08	50	0.005	−8.58**	51	0.024
B34. Integration with the world						
B341. (Exports + Imports)/GDP	−170.15	59	0.002	−743.19	53	0.016
B342. Capital controls index	−12.93**	59	0.116	−21.17	59	0.152
B343. Foreign ownership restrictions index	−10.97*	46	0.069	−13.76	43	0.049
B35. Legal environment						
B351. Legal origin (Common vs Others)	3.14	59	0.007	8.08	59	0.022
B352. Corruption index	1637170.21***	43	0.144	2334.26	56	0.175

It could be that our results are driven by the huge number of foreign-origins from Germany, Ireland, Italy, Mexico and UK. We dropped these countries. The statistical significance of the size times income of the foreign-origin group in both the years remains. Again, as we did in the FDI uni-variate tests, we make an adjustment for standard errors for multiple comparisons in Table 5. Using the Bonferoni correction, which is the severest

of all such tests, the statistical significance we found for the total X income variable or the native born X income variable in Table 5 disappears. However, the important tests are the multi-variate tests that attempt to control for omitted effects, and this is where we now turn.

Table 6 checks whether the income of the national origin group continues to remain significant after we have controlled for the “fundamental” determinants of foreign indirect investments. Of the control variables, we keep the predicted equity holdings because it is so important, but further reduce the number of other control variables by extracting principal components from them. We obtain 4 and 5 significant principal components in 1997 and 2001 respectively. Table 6 gives results of cross-sectional multi-variate regressions for the years 1997 and 2001, where the dependent variable is U.S. indirect investment, and the independent variables are the predicted equity holding variable, the above extracted major principal components from the other control variables, and each familiarity variable in turn.

Though we do not report it in Table 6, the effect of the other familiarity variables—race, religion, and language—continue to remain insignificant after we have added the controls, whereas physical distance is significant in only 1997 at the 10% level. The only familiarity variable that statistically significantly influences (p -values of 10% or less) U.S. stockholdings in a foreign country in *both* the two years, 1997 and 2001, is the total income of the native borns. The signs of the coefficients tell us that, controlling for the expected influences, countries with economically stronger native borns in the U.S. attract more U.S. shareholdings. We find that the R^2 for a regression that contains only the “fundamentals” is significant, which implies that the “fundamentals” do a good job in explaining U.S. foreign equity holdings. A F-test in Table 6 shows that the addition of the income of the native borns does not significantly improve the statistical fit of the above cross-sectional regressions in either of the years 1997 and 2001. The 2001 coefficient in Table 6 for the income of the native borns (Native born $\times$ Income) is 7.28E–12, which implies that, everything else constant, every additional dollar earned by a native born is associated with an increase in equity holdings of USD 0.00728 in their country of origin. To put this in perspective, if a country in the 25th percentile in native born income (Czech Republic in 2001, with its native borns in the U.S. earning USD 0.27 billion) moved up to be in the 75th percentile in native born income (Greece in 2001, with its native borns in the U.S. earning USD 2.13 billion), U.S. equity holdings there would increase by about USD 14 million, everything else constant.

Interestingly, the high R^2 in this regression tell us that the independent variables can explain more than four-fifths of the variation in U.S. foreign shareholdings. Much of this high explanatory power comes from the predicted equity holdings variable, telling us that the significant determinant of U.S. stockholdings in a foreign country is the size of that country’s stock market.

We now do a panel data test correcting for country fixed effects. Table 7 gives the results from this panel data test. The dependent variable is U.S. equity holdings abroad. In Model A, the independent variables are the independent variables from Table 5 that have no missing observation in either of the two years, 1997 and 2001, for a set of countries. This reduces the number of independent variables without any obvious variable selection bias that might affect our results. Model A1 shows the effect of the size times income of the total group, whereas Model A2 shows the effect of the size times income of the native-born segment. Model A in Table 7 shows that the effect of the income of the national origin

Table 6

Effect of national origin on U.S. indirect investments abroad (results from multi-variate OLS regressions)

The dependent variable is U.S. level of stockholdings in a foreign country (U.S. foreign stockholdings) in billions of dollars. The independent variables are obtained as follows. In the first step, we reduce the number of control variables from Table 5 by extracting major principal components from all the control variables except the predicted equity holdings and the familiarity variables. In the second step, we run a multi-variate OLS regression using the extracted principal components, the predicted equity holdings variable, and one familiarity variable (Total $\times$ Income or Native $\times$ Income) as the independent variables. Though we do not report it in the above table, we also ran this multi-variate OLS regression using the other familiarity variables. The results from those regressions are: in the 1997 regression model, only the distance becomes significant with p-value of 0.089, while all other familiarity variables are not; in the 2001 regression model, none of the other familiarity variables is significant. Standard errors are White-heteroskedasticity corrected. * Significance at the 10% level. ** Significance at the 5% level. *** Significance at the 1% level.

Independent variables		1997		1997		2001		2001	
		Coefficient estimate	p-Value	Coefficient estimate	p-Value	Coefficient estimate	p-Value	Coefficient estimate	p-Value
Constant		6.39*	0.069	6.44*	0.070	3.65	0.397	3.84	0.372
Familiarity variable	B151a. Total (National Origin) $\times$ Income	3.06E–10	0.418			2.73E–11	0.903		
	B153a. Native born $\times$ Income			3.74E–10*	0.084			7.28E–12*	0.076
B3. Control variables	B31. Predicted equity holdings	1.44E–4***	0.000	1.44E–4***	0.000	1.84E–4***	0.000	1.85E–4***	0.000
Principal components extracted from the other control variables proxying for “fundamentals” in Table 6	Principal component one	6.29***	0.007	6.22***	0.008	8.37***	0.002	8.37***	0.002
	Principal component two	1.61	0.267	1.65	0.259	2.31	0.241	2.36	0.226
	Principal component three	–2.24	0.184	–2.21	0.189	0.049	0.978	0.034	0.985
	Principal component four	–0.42	0.83	–0.44	0.823	–0.32	0.886	–0.37	0.867
	Principal component five					7.63***	0.011	7.68**	0.011
<i>N</i>		34		34		36		36	
<i>R</i> ²		0.905		0.909		0.862		0.864	
<i>R</i> ² with B151a or B153a excluded		0.904		0.904		0.858		0.858	
<i>F</i> test of inclusion of B151a. or B153a.		0.338		1.014		0.686		1.163	

Table 7

Effect of national origin on U.S. indirect investments abroad (results from panel regression with country-fixes effects)

The dependent variable in both Models A and B is U.S. level of indirect investments in a foreign country (U.S. foreign stockholdings) in billions of dollars. In Model A, the independent variables are all independent variables from Table 5 for which we have all observations for both years, 1997 and 2001, for a set of countries. However, we do not include race, religion, language, distance and legal origin variables, because they do not change over time. Total, Total $\times$ Income, Foreign born, Foreign born $\times$ Income, Native born, Native born $\times$ Income variables are highly mutually correlated, with Pearson correlation coefficients above 0.7. Therefore, we include only Total $\times$ Income in regression model A1, and Native born $\times$ Income in regression model A2. We also exclude all three risk-return variables, because they are all highly correlated with Average market cap of listed firm. We also exclude (Exports + Imports)/GDP because it is highly correlated with Capital Controls index, and we exclude Volume of trade/Market cap because of high correlation with Volume of trade. The final sample in Model A contains observations for 29 countries for which we have all control variables for both years, 1997 and 2001. In Model B, the independent variables are all control variables in Table 5 that were found to be statistically significant in both years, 1997 and 2001. As this selection method may bias estimates, the coefficients and p-Values reported above for Model B are estimated with a bootstrap method with re-sampling. The number of repetitions in the bootstrap was 10,000. We run two separate models, Model B1 and Model B2, because Total $\times$ Income and Native born $\times$ Income are highly correlated (Pearson correlation coefficient above 0.9). Standard errors in Models A1, A2, B1 and B2 are White-heteroskedasticity corrected. *Significance at the 10% level. **Significance at the 5% level. ***Significance at the 1% level.

Independent variables	Model A				Model B			
	Model A1		Model A2		Model B1		Model B2	
	Coefficient estimate	p-Value	Coefficient estimate	p-Value	Coefficient estimate	p-Value	Coefficient estimate	p-Value
B151a. Total (National Origin) $\times$ Income	3.748E–10*	0.084			3.714E–10*	0.082		
B153a. Native born $\times$ Income			4.83E–10*	0.087			5.387E–10*	0.064
B31. Predicted equity holdings	1.8E–4***	0.000	1.8E–4***	0.000	1.78E–4***	0.000	1.77E–4***	0.000
B321. Market cap of stock market/GDP	0.25*	0.051	0.25**	0.050	0.16	0.128	0.16	0.136
B322. Average market cap of listed firm	–6.5E–09	0.591	–6.7E–09	0.585				
B323. Number of listed firms	–2.69E–3	0.231	–2.79E–3	0.219				
B324. Volume of trade	0.09	0.143	0.09	0.138	0.04	0.192	0.04	0.184
B342. Capital controls index	–8.05	0.525	–7.71	0.543				
B343. Foreign ownership restrictions index	14.61	0.205	16.86	0.155				
Number of countries in sample	29		29		59		59	
R ²	0.905		0.907		0.917		0.924	
F test of inclusion of B151 a or B153a.	1.853		1.905		1.735		1.841	

group or the income of the native born of the national origin group continues to remain significant. In Model B, the independent variables are the independent variables from Table 5 that were found to be statistically significant in both years, 1997 and 2001. As this method for reducing the number of independent variables may bias estimates, the coefficients and p -values reported for Model B are estimated with a bootstrap method with re-sampling. Model B1 shows the effect of the size times income of the total group, whereas Model B2 shows the effect of the size times income of the native-born segment. Table 7, Model B, shows that the effect of the income of the national origin group or the income of the native born of the national origin group continues to remain significant.

We redo the above panel data test with country fixed effects with and without the income of the national origin group variable or the income of the native born of the national origin group, and ask whether the R^2 increase in a statistical sense if we include these variables. We find that the addition of these variables does not significantly improve fit in any model in Table 7. Table 7 thus provides modest evidence that countries with larger emigrant groups in the U.S. attract more U.S. foreign indirect investments.

6. Concluding remarks

We make an attempt in this paper to explore whether, when it comes to directly investing or indirectly investing in a particular foreign country, foreign-origin persons in the United States consider only the “fundamentals” of the country under consideration—the melting pot hypothesis—or, besides the “fundamentals,” they also care whether this particular foreign country is their country of origin—the salad bowl hypothesis. If national origin affects direct investments abroad, it should affect indirect investments abroad in the same direction as well.

This is what we find. After controlling for the “fundamentals” hypothesized in the literature, we document that U.S. investments in a foreign country are positively affected by the size of the foreign-origin group from that country living in the United States. This national origin bias is strong for direct (FDI) and modest for indirect investments (foreign stock holdings). The other economic geography variables of that country—physical distance, race, language and religion—do not seem to affect U.S. investments in that country. This suggests that the U.S. is perhaps more like a salad bowl than a melting pot when it comes to investing abroad.

Our study has a number of data limitations, which we hope future researchers will rectify. First, as our data are on the aggregate U.S. level, and not at the level of the investor, we cannot explore the possibly different biases exhibited by institutional investors vis-a-vis individual investors. Second, since we have cross-sectional data where the data comes from different sources at different points in time, though we can do robustness checks using different data sources, we have limitations doing panel data tests.

Appendix. Data definitions and data sources

We have data for 64 countries. We do not have data on *all* variables for *all* the 64 countries for *all* years. The data variables are the following.

A. Dependent variables

A1. U.S. *direct investments* in country i is the stock of U.S. direct investments in country i measured at historical cost in the years 1970, 1980, 1990 and 2000 (see Table 8). Source: Bureau of Economic Activity, U.S. Department of Commerce, <http://www.bea.doc.gov>.

A2: U.S. *indirect investments* (US *foreign stockholdings*) in country i is the stock of U.S. stockholdings in country i . This data, available for 1997 and 2001, comes from a survey carried out by the U.S. Treasury Department and the Board of Governors of the Federal Reserve System in 1997 and 2001. Source: <http://www.treas.gov/tic/fpis.html>.

B. Independent variables

B1. Familiarity variables

B11. *Physical distance* of country i is the distance in miles between New York and the city in the foreign country where the main stock exchange is located. Source: <http://www.indo.com/distance/>.

B12. *Language* of country i takes the value of 1 if the CIA lists English as one of the major languages spoken in the country, and 0 otherwise. Source: CIA World Factbook 2001, <http://www.cia.gov/cia/publications/factbook/>.

B13. *Race* of country i is defined as the percentage of total population of a country that is described by CIA to be of white or Caucasian race. If percentage is not given, but if the CIA lists white or Caucasian to be the majority race, then the country receives 1, and 0 otherwise. Source: CIA World Factbook 2001, <http://www.cia.gov/cia/publications/factbook/>.

B14. *Religion* of country i is defined as the percentage of total population of a country that belongs to any Christian affiliated religion. If percentage is not given, but if the CIA lists Christianity as the majority religion, then the country receives 1, and 0 otherwise. Source: CIA World Factbook 2001, <http://www.cia.gov/cia/publications/factbook/>.

B15. National Origin

B151. *Total* is defined as the total number of members of a certain national origin group. For 1970, 1980, 1990, and 2000, this is estimated by the U.S. Bureau of the Census. However, for 1997 and 2001, we estimated these from the Current Population Surveys (CPS) conducted by the U.S. Bureau of the Census. This is our estimation algorithm for the CPS data. We assign a respondent to a national origin group if he or she declared himself/herself as a member of the national origin group (survey code: PENATVTY), or if both parents belonged to the same national origin group, or one of the parents belonged to the national group and the other parent belonged to the USA (survey code for mother: PEMNTVTY and survey code for father: PEFNTVTY). We extrapolate survey results by multiplying them with a survey's second stage weight (PWSSWGT- rake 6 final step weight). Source: U.S. Bureau of the Census and Current Population Survey (<http://www.bls.census.gov/cps/>).

B151a. *Total X Income* is defined as the total number of members of a certain national origin group multiplied by the per capita income of this group. The per capita income was taken to be the mid-point of the range provided by the household divided by the number of members in the household. We have these data for 1990, 1997 and 2001. Source: U.S. Bureau of the Census and Current Population Survey (<http://www.bls.census.gov/cps/>).

Table 8

Independent variables	Panel A: Descriptive statistics of variables explaining U.S. direct investments abroad											
	1970			1980			1990			2000		
	Mean	Med	S.Dev	Mean	Med	S.Dev	Mean	Med	S.Dev	Mean	Med	S.Dev
B1. Familiarity variables												
B11. Physical distance (in miles)	3998	3683	1847	3998	3683	1847	3998	3683	1847	3998	3683	1847
B12. Language	0.38	0	0.49	0.38	0	0.49	0.38	0	0.49	0.38	0	0.49
B13. Race	0.68	0.97	0.43	0.68	0.97	0.43	0.68	0.97	0.43	0.68	0.97	0.43
B14. Religion	0.81	0.91	0.28	0.81	0.91	0.28	0.81	0.91	0.28	0.81	0.91	0.28
B15. National origin												
B151. Total (in thousandths)	637.660	155.156	1076.718	2117.134	223.652	5110.364	3672.310	450.519	8915.413	3737.561	536.503	7913.286
B2. Control variables proxying for “fundamentals”												
B21. Economic development												
B211. Average GDP (in billions of \$)	87.114	12.999	348.271	146.371	58.801	231.824	205.764	85.981	383.716	444.491	166.370	820.231
B212. Average GDP growth	6.03	5.79	2.35	3.73	3.74	2.40	2.91	2.78	1.99	3.21	2.94	1.68
B213. Average per capita income (in dollars)	9112.57	7304.87	7789.36	12047.65	10696.63	9826.30	14287.09	12062.24	12425.27	17834.14	16639.63	14875.93
B22. Size of government												
B221. Govt consumption/GDP(%)	13.56	11.74	5.15	17.05	15.69	6.84	15.90	16.41	6.43	11.66	11.51	3.78
B23. Corporate tax burden												
B231. Taxes (Corp Tax/Total Tax)	35.64	30.29	16.51	35.88	31.51	19.05	34.79	34.23	18.94	31.11	33.00	7.04
B232. Import duty(ImTax/Total Tax)	9.85	6.48	8.15	6.89	5.45	7.22	4.88	1.68	7.08	3.81	1.18	5.25
B233. Export duty (Ex Tax/Total Tax)	0.23	0	0.88	2.19	0	4.70	0.87	0	2.48	0.03	0	0.12
B24. Integration with the world												
B241. (Exports + Imports)/GDP(%)	49.72	42.61	32.31	62.25	52.04	34.88	64.16	59.60	38.88	77.60	83.70	23.67
B242. Capital controls index	0.56	0.67	0.31	0.53	0.67	0.33	0.61	0.67	0.35	0.40	0.33	0.19
B243. Liberalization dummy	0.63	1.00	0.49	0.63	1.00	0.49	0.16	0.00	0.37			
B244. Country credit rating				69.53	72.10	21.20	54.94	63.15	28.85	64.17	75.80	24.31
B25. Infrastructure												
B251. Air freight (million tons-km)	126.41	58.15	156.13	364.08	125.33	486.14	829.60	183.80	1265.63	1542.39	324.73	2251.19
B252. Roads freight (million ton-km)				43800	17450	55420.65	59521.54	25650	76707.92	82182.72	30950	94833.26
B253. Percentage of paved roads							57.96	66.5	34.19	58.53	60.49	29.57
B254. Telephone lines per 1000	114.21	84.00	105.94	182.31	180.63	153.71	278.07	298	201.62	384.90	456.5	223.56
B255. Telephone lines per empl.				87.56	73.52	67.45	115.32	103.33	58.95	211.74	206.26	74.40
B26. Capital market development												
B261. Credit to private sec/GDP (%)	36.09	27.46	22.41	50.59	43.76	29.67	59.94	48.55	44.94	58.34	65.23	34.37
B262. Market cap of stock market (in billions of \$)							179.108	20.893	526.579	10288.549	110.839	59895.059

B263. Mar. cap of sto mar./GDP(%)							29.28	21.49	26.42	33.10	19.6	30.75
B264. Av. mar. cap of listed firm (in billions of \$)							0.270	0.165	0.317	5.722	0.477	28.493
B265. Volume of trade							8.63	4.61	11.28	61.97	37.17	72.20
B266. Volume of trade/Mar. Cap(%)							37.46	30.3	33.18	62.25	60.10	59.84
B27. Exchange rate risk												
B271. St. dev. of real exch. rate				4.78	0.20	11.85	0.06	0.03	0.10	0.04	0.03	0.02
B28. Legal environment												
B281. Legal origin (Common vs Others)	0.22	0.00	0.42	0.22	0.00	0.42	0.19	0.00	0.40	0.19	0.00	0.40
Independent variables						Panel B: Descriptive statistics of variables explaining U.S. indirect investments abroad						
						1997			2001			
						Mean	Med.	S.Dev	Mean	Med.	S.Dev	
B1. Familiarity variables												
B11. Physical distance (in miles)				5059	4376	2199	5059	4376	2199			
B12. Language				0.46	0	0.50	0.46	0	0.50			
B13. Race				0.56	0.89	0.47	0.56	0.89	0.47			
B14. Religion				0.63	0.83	0.39	0.63	0.83	0.39			
B15. National origin												
B151. Total (in thousandths)				273.956	74.426	685.446	412.805	124.350	1099.254			
B151a. Total × Income (in billions of \$)				3.763	1.252	6.483	6.569	2.008	11.886			
B152. Foreign born (in thousandths)				55.062	19.017	118.307	98.403	31.781	210.978			
B152a. Foreign born × Income (in billions of \$)				0.780	0.286	1.337	1.688	0.644	2.650			
B153. Native born (in thousandths)				218.894	46.586	580.216	314.402	76.670	902.699			
B153a. Native born × Income (in billions of \$)				2.983	0.782	5.447	4.881	1.390	9.667			
B3. Control variables proxying for “fundamentals”												
B31. Predicted equity holdings per country (in millions of \$)				86.885	16.967	200.344	105.176	18.402	229.492			
B32. Stock market development												
B321. Market cap of stock market/GDP(%)				41.01	31.5	35.95	45.59	32.17	41.69			
B322. Average market cap of listed firm(in billions of \$)				0.344	0.160	0.441	0.366	0.149	0.529			
B323. Number of listed firms				519.64	235	883.29	557.90	230	976.76			
B324. Volume of trade(% of GDP)				22.99	14.35	26.75	37.20	9.93	56.26			
B325. Volume of trade/Market cap(%)				57.99	52.73	47.44	57.72	37.4	69.08			
B33. Risk-return												
B331. Mean stock market return/Standard dev. of return				0.15	0.17	0.34	0.00	0.03	0.17			
B332. Correlation of market return with U.S. return (in billions of \$)				0.28	0.32	0.26	−0.04	−0.06	0.18			

Table 8 (continued)

Independent variables	Panel B: Descriptive statistics of variables explaining U.S. indirect investments abroad					
	1997			2001		
	Mean	Med.	S.Dev	Mean	Med.	S.Dev
B333. Standard dev. of real exchange rate	0.03	0.02	0.03	5.49	5.43	4.14
B34. Integration with the world						
B341. (Exports + Imports)/GDP(%)	72.94	60.79	39.14	79.04	68.04	41.41
B342. Capital controls index	0.38	0.33	0.25	0.13	0.17	0.07
B343. Foreign ownership restrictions index	0.13	0.01	0.23	0.10	0.00	0.20
B35. Legal environment						
B351. Legal origin (Common vs Others)	0.27	0.00	0.45	0.27	0.00	0.45
B352. Corruption index	5.68	5.2	2.51	5.18	4.5	2.44

B152. Foreign born is defined as the total number of members of a certain national origin group that were born abroad, and became U.S. citizens by naturalization. For 1997 and 2001, we estimated these from the Current Population Surveys (CPS) from the variable PRCITSHP (value 4 defines Foreign Born U.S. citizens)-United States citizenship group. For 1990 this is estimated by the U.S. Bureau of the Census. Source: U.S. Bureau of the Census and Current Population Survey (<http://www.bls.census.gov/cps/>).

B152a. Foreign born X Income is defined as the total number of members of a certain national origin group who are foreign born as defined in B152 multiplied by the per capita income of this group. We have these data for 1990, 1997 and 2001. The per capita income is estimated as in B151a. Source: U.S. Bureau of the Census and Current Population Survey (<http://www.bls.census.gov/cps/>).

B153. Native born is defined as the total number of members of a certain national origin group that were born in the US, or the U.S. provinces or outlying areas, or abroad to parents who are U.S. citizens. For 1997 and 2001, we estimated these from the Current Population Surveys (CPS) from the variable PRCITSHP (values 1, 2, and 3 define Native born U.S. citizens)-United States citizenship group. For 1990 this is estimated by the U.S. Bureau of the Census. Source: U.S. Bureau of the Census and Current Population Survey (<http://www.bls.census.gov/cps/>).

B153a. Native born X Income is defined as the total number of members of a certain national origin group who were native born as defined in B153. multiplied by the per capita income of this group. We have these data for 1990, 1997 and 2001. The per capita income is estimated as in B151a. Source: U.S. Bureau of the Census and Current Population Survey (<http://www.bls.census.gov/cps/>).

B2. Control variables proxying for “fundamentals” determining direct investment

B21. Economic development

B211. Average GDP is defined as the average of GDP levels measured in current U.S. dollars for the three years prior to and including 1970, 1980, 1990, and 2000. Source: World Bank, <http://www.worldbank.org/data/>.

B212. Average GDP growth is defined as the average of GDP annual growth for the three years prior to and including 1970, 1980, 1990, and 2000. Source: World Bank, <http://www.worldbank.org/data/>.

B213. Average per capita income is defined as the average per capita income levels measured in constant 1995 U.S. dollars for the three years prior to and including 1970, 1980, 1990, and 2000. Source: World Bank, <http://www.worldbank.org/data/>.

B214. Average GDP per capita growth is defined as the average growth of annual real GDP per capita for the three years prior to and including 1970, 1980, 1990, and 2000. Source: World Bank, <http://www.worldbank.org/data/>.

B22. Size of the government

B221. Government consumption/GDP is the general government final consumption expenditure as a percentage of GDP in the years 1970, 1980, 1990 and 2000. Source: World Bank, <http://www.worldbank.org/data/>.

B23. Corporate tax burden

B231. Taxes are defined as the percentage in total tax revenues from taxes on income, profit and capital gains in years 1980, 1990 and 2000. We do not use this variable for 1970

because we only have data for 17 countries. Source: World Bank, <http://www.worldbank.org/data/>.

B232. *Import duty* is defined as the percentage of import duties in total taxes in years 1980, 1990 and 2000. We do not use this variable for 1970 because we only have data for 17 countries. Source: World Bank, <http://www.worldbank.org/data/>.

B233. *Export duty* is defined as the percentage of export duties in total taxes in years 1980, 1990 and 2000. We do not use this variable for 1970 because we only have data for 17 countries. Source: World Bank, <http://www.worldbank.org/data/>.

B24. Integration with the world

B241. $(\text{Export} + \text{Imports}) / \text{GDP}$ is defined as the sum of imports and exports expressed as a percentage of GDP in 1970, 1980, 1990 and 2000. Source: World Bank, <http://www.worldbank.org/data/>.

B242. *Capital controls index* is defined as the average of three dummies. The first dummy equals 1 if a country imposes any restriction on capital account transactions, and 0 otherwise; the second dummy equals 1 if a country uses multiple exchange rates, and 0 otherwise; the third dummy equals 1 if the country enforces surrender requirements for export proceeds, and 0 otherwise. The average for 1969 and 1970 was used to construct the value for 1970, the average for years 1979 and 1980 was used to construct the value for 1980, the average for years 1989 and 1990 was used to construct the value for 1990, and the average for 1999 and 2000 was used to construct the value for 2000. Source: IMF's *Annual Report on Exchange Arrangements and Exchange Restrictions*, 1969, 1970, 1979, 1980, 1989, 1990, 1999 and 2000.

B243. *Liberalization dummy* equals 1 at time t ($t = 1970, 1980, 1990, 2000$) if the official liberalization date of this country is at or before t . Source: Bekaert and Harvey (2000).

B244. *Country credit rating* comes from the Institutional Investor credit ratings that are based on a survey of leading international bankers who are asked to rate each country on a scale from 0 to 100 (where 100 represents maximum creditworthiness). Institutional Investor averages these ratings, providing greater weights to respondents with greater worldwide exposure and more sophisticated country analysis systems. We take the values from the September issue of the Institutional Investor issued a year prior to 1980, 1990 and 2000. We do not have data for this variable for year 1970.

B245. *Time since liberalization* is measured as: $1 - e^{-\text{years since the official liberalization date}}$. We construct this variable only for year 2000.

B25. Infrastructure

B251. *Air freight* is defined as the average air freight in millions tons per kilometer for the three years prior to and including 1970, 1980, 1990, and 2000. Source: World Bank, <http://www.worldbank.org/data/>.

B252. *Roads freight* is defined as the average goods transported in million ton-km for the three years prior to and including 1990 and 2000. We do not have data for this variable for years 1970 and 1980. Source: World Bank, <http://www.worldbank.org/data/>.

B253. *Percentage of paved roads* is defined as the average percentage of roads in each country that are paved for the three years prior to and including 1990 and 2000. We do not have data for this variable for years 1970 and 1980. Source: World Bank, <http://www.worldbank.org/data/>.

B254. *Telephone lines per 1000* is defined as the average of telephone mainlines per 1000 people for three years prior to and including 1970, 1980, 1990, and 2000. Source: World Bank, <http://www.worldbank.org/data/>.

B255. Telephone lines per employee is defined as the average of telephone mainlines per employee for three years prior to and including 1980, 1990, and 2000. We do not have data for this variable for year 1970. Source: World Bank, <http://www.worldbank.org/data/>.

B26. Capital market development

B261. Credit to private sector/GDP is defined as the total credit approved by financial intermediaries to the private sector as percentage of GDP during 1970, 1980, 1990 and 2000. Source: World Bank, <http://www.worldbank.org/data/>.

B262. Market cap of stock market is defined as the market capitalization of listed companies measured in current USD for years 1990 and 2000. We do not have data for this variable for years 1970 and 1980. Source: World Bank, <http://www.worldbank.org/data/> and <http://www.fibv.com>.

B263. Market cap of stock market/GDP is defined as market capitalization of listed companies measured as percentage of GDP for years 1990 and 2000. We do not have data for this variable for years 1970 and 1980. Source: World Bank, <http://www.worldbank.org/data/> and <http://www.fibv.com>.

B264. Average market cap of listed firms is estimated as the ratio between total market capitalization of listed companies measured in current US\$ for years 1990 and 2000, and total number of listed stocks in those years. We do not have data for this variable for years 1970 and 1980. Source: World Bank, <http://www.worldbank.org/data/> and <http://www.fibv.com>.

B265. Volume of trade is defined as stock exchange dollar volume traded as a percentage of GDP for years 1990 and 2000. We do not have data for this variable for years 1970 and 1980. Source: World Bank, <http://www.worldbank.org/data/> and <http://www.fibv.com>.

B266. Volume of trade/Market cap is calculated as the ratio between stock exchange dollar volume traded and its dollar market capitalization for years 1990 and 2000. We do not have data for this variable for years 1970 and 1980. Source: World Bank, <http://www.worldbank.org/data/> and <http://www.fibv.com>.

B27. Exchange rate risk

B271. Standard dev. of real exchange rate is estimated as the standard deviation of real exchange rates. We use monthly nominal exchange rates and monthly consumer price index levels for three years prior to and including 1980, 1990, and 2000 in real exchange calculations. We do not use this variable for 1970, because of the global prevalence of fixed exchange rate regimes at that time. Source: IMF, *International Finance Statistics*, 2001.

B28. Legal environment

B281. Legal origin (Common vs. Others) is a dummy variable that equals 1 for countries with common law legal environments, and 0 for other legal systems. Source: CIA World Factbook, 2001, <http://www.cia.gov/cia/publications/factbook/>.

B3. Control variables proxying for “fundamentals” determining indirect investment (foreign stockholdings)

All data here are for years 1997 and 2001 because the U.S. Treasury Department and the Board of Governors of the Federal Reserve System carried out surveys on U.S. foreign stock holdings at the end of 1997 and at the end of 2001.

B31. Predicted equity holdings are defined as U.S. holdings in country *i* if the world equity portfolio was the market portfolio. In this case, predicted equity holdings of a U.S. investor would be the weight of stock market capitalization of country *i* in the total world

stock market capitalization multiplied by total U.S. equity holdings. Total U.S. equity holdings are calculated as the sum of foreign equity holdings and domestic equity holdings held by U.S. citizens. Domestic equity holdings of U.S. citizens is U.S. stock market capitalization minus the U.S. equities owned by foreigners. Sources for the stock market capitalization of various countries: IFC, The Emerging Stock Markets Factbook (2000), Standard & Poor's Emerging Stock Markets Factbook (2002) and <http://www.fibv.com>. Source for equity holdings of U.S. investors abroad was <http://www.treas.gov/tic/fpis.html>, and source for equity holdings of foreigners for 1997 in U.S. was <http://www.treas.gov/tic/shl94sum.html> and for 2001 in <http://www.treas.gov/tic/shl2000r.pdf>.

B32. Stock market development

B321. *Market cap of stock market/GDP* is the same as B263 for 1997 and 2001.

B322. *Average market cap of listed firm* is the same as B264 for 1997 and 2001.

B323. *Number of listed firms* is defined as the total number of listed firms in country i in 1997 and 2001. Source: World Bank, <http://www.worldbank.org/data/>.

B324. *Volume of trade* is the same as B265 for 1997 and 2001.

B325. *Volume of trade / Market cap* is the same as B266 for 1997 and 2001.

B33. Risk-Return

B331. *Mean stock market return/Standard dev. of return* is estimated as the ratio of average monthly dollar dividend-adjusted return of stock market index of country i and its standard deviation for the January, 1994 to December 1997 period for the year 1997, and January, 1999 to December 2001 period for the year 2001. The minimum required number of months with data for each market is 12. Sources: All indices are from Morgan Stanley Capital International (<http://www.msci.com>), except for the data for 12 emerging markets, that come from IFC, *Emerging Stock Markets Factbook* (1998) and Standard & Poor's *Emerging Stock Markets Factbook* (2002).

B332. *Correlation of stock market return with U.S. return* is defined as the correlation coefficient between dollar monthly returns in USA and foreign markets in the period three years before and including 1997 and 2001. A minimum number of 12 observations is required to calculate a correlation coefficient. Sources: All indices are from Morgan Stanley Capital International (<http://www.msci.com>), except for the data for 12 emerging markets that come from IFC, *Emerging Stock Markets Factbook* (1998) and Standard & Poor's *Emerging Stock Markets Factbook* (2002).

B333. *Standard dev. of real exchange rate* is the same as B271 for 1997 and 2001.

B334. *Last year's return* is estimated as the average monthly dollar dividend-adjusted return of stock market index of country i for 12 months prior to the year 1997 and 2001. Sources: All indices are from Morgan Stanley Capital International (<http://www.msci.com>), except for the data for 12 emerging markets that come from IFC, *Emerging Stock Markets Factbook* (1998) and Standard & Poor's *Emerging Stock Markets Factbook* (2002).

B34. Integration with the world

B341. *(Exports+Imports)/GDP* is the same as B241 for 1997 and 2001.

B342. *Capital controls index* is the same as B242 for 1997 and 2001.

B343. *Foreign ownership restrictions index* is defined as one minus the ratio of the market capitalization of the IFC's Investable Index to the market capitalization of the IFC's Global Index. If the country was not included in IFC emerging market coverage, it received a value of 0. This variable is constructed as in Edison and Warnock (2003) and, for additional countries, from the IFC, *Emerging Stock Markets Factbook* (1998) and

Standard & Poor's Emerging Stock Markets Factbook (2002). The data are for 1997 and 2001.

B35. Legal environment

B351. Legal origin (Common vs. Others) is the same as B281 for 1997 and 2001.

B352. Corruption index—Corruption Perception Index (CPI) for years 1997 and 2001. Score for a country relates to perceptions of the degree of corruption as seen by business people, academics and risk analysts, and ranges between 10 (highly clean) and 0 (highly corrupt). Source: Transparency International, <http://www.transparency.org>.

References

- Ahearn, A.G., Grier, W.L., Warnock, F., 2004. Information costs and home bias: an analysis of U.S. holdings of foreign equities. *Journal of International Economics* 62, 313–336.
- Alkhalifa, A., 1999. Islamic banking. Ph.D. dissertation, University of Miami.
- Annual Report on Exchange Arrangements and Exchange Restrictions, 1979, 19980, 1989, 1990, 1999, 2000. International Monetary Fund, Washington, DC.
- Barber, B., Odean, T., 2001. Boys will be boys: gender, overconfidence, and common stock investment. *Quarterly Journal of Economics* 116, 261–292.
- Bartolini, L., Drazan, A., 1997. When liberal policies reflect external shocks, what do we learn? *Journal of International Economics* 42, 249–273.
- Beck, T., Levine, R., Loayza, N., 2000. Finance and the sources of growth. *Journal of Financial Economics* 58, 261–300.
- Bekaert, G., Harvey, C., 1995. Time varying world market integration. *Journal of Finance* 50, 403–444.
- Bekaert, G., Harvey, C., 2000. Foreign speculators and emerging equity markets. *Journal of Finance* 55, 565–613.
- Berger, A., De Young, R., Genay, H., Udell, G., 2000. Globalization of financial institutions: Evidence from cross-border banking performance. Working paper, Federal Reserve Board (Washington, D.C.)
- Berger, A., Klapper, L., Udell, G., 2001. The ability of banks to lend to informationally opaque small businesses. *Journal of Banking and Finance* 25, 2127–2167.
- Bhattacharya, U., Daouk, H., Jorgenson, B., Kehr, C.-H., 2000. When an event is not an event: the curious case of an emerging market. *Journal of Financial Economics* 55, 69–101.
- Bhattacharya, U., Daouk, H., 2002. The world price of insider trading. *Journal of Finance* 57, 75–108.
- Bhattacharya, U., Daouk, H., Welker, M., 2003. The world price of earnings opacity. *Accounting Review* 78, 641–678.
- Calem, P., Nakamura, L., 1998. Branch banking and the geography of bank pricing. *Review of Economics and Statistics* 80, 600–610.
- Chakrabarti, A., 2001. The determinants of foreign direct investment: Sensitivity analyses of cross-country regressions. *Kyklos* 54, 89–113.
- Choe, H., Kho, B.-C., Stulz, R., 2001. Do domestic investors have more valuable information about individual stocks than foreign investors? Working paper, National Bureau of Economic Research (Cambridge, MA) WP No 8073.
- Coval, J., Moskowitz, T., 1999. Home bias at home: local equity preference in domestic portfolios. *Journal of Finance* 54, 2045–2073.
- Coval, J., Moskowitz, T., 2001. The geography of investment: informed trading and asset prices. *Journal of Political Economy* 109, 811–841.
- Edison, H., Warnock, F., 2003. A simple measure of the intensity of capital controls. *Journal of Empirical Finance* 10, 81–103.
- Emerging Stock Markets Factbook, 1998, 2000. International Finance Corporation, Washington, DC.
- Feldstein, M., Horioka, C., 1980. Domestic saving and international capital flows. *Economic Journal* 90, 314–329.
- French, K., Poterba, J., 1991. Investor diversification and international equity markets. *American Economic Review* 81, 222–226.
- Froot, K., O'Connell, P., Seasholes, M., 2001. The portfolio flows of international investors. *Journal of Financial Economics* 59, 151–193.
- Glassman, D., Riddick, L., 2001. What causes home asset bias and how should it be measured? *Journal of Empirical Finance* 8, 35–54.

- Goldberg, L., Kolstad, C., 1995. Foreign direct investment, exchange rate variability, and demand uncertainty. *International Economic Review* 36, 855–873.
- Grinblatt, M., Keloharju, M., 2001. How distance, language, and culture influence stockholdings and trades. *Journal of Finance* 56, 1053–1073.
- Hau, H., 2001. Location matters: an examination of trading profits. *Journal of Finance* 56, 1959–1983.
- Huberman, G., 2001. Familiarity breeds investment. *Review of Financial Studies* 14, 659–680.
- International Finance Statistics, 2001. International Monetary Fund, Washington, DC.
- Kaufmann, D., Kraay, A., Zoido-Lobaton, P., 1999. Governance matters. Working paper, World Bank Policy Research Department WP No. 2196.
- Krugman, P., 1991. Increasing returns and economic geography. *Journal of Political Economy* 99, 483–499.
- Karolyi, A., Stulz, R., 2003. Are financial assets priced locally or globally? In: Constantinides, G., Harris, M., Stulz, R. (Eds.), *Handbook of the Economics of Finance* 1B. Elsevier North-Holland, Amsterdam (Chapter 16).
- La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R., 1998. Law and finance. *Journal of Political Economy* 106, 1113–1155.
- Lewis, K., 1999. Trying to explain home bias in equities and consumption. *Journal of Economic Literature* 37, 571–608.
- Neary, P., 2001. Of hype and hyperbolas: Introducing the new economic geography. *Journal of Economic Literature* 39, 536–561.
- Pastor, L., 2000. Portfolio selection and asset pricing models. *Journal of Finance* 55, 179–223.
- Petersen, M., Rajan, R., 2000. Does distance still matter? The information revolution in small business lending. Working paper, Northwestern University.
- Portes, R., Rey, H., 2005. The determinants of cross-border equity flows. *Journal of International Economics* 65, 269–296.
- Stulz, R., Williamson, R., 2003. Culture, openness, and finance. *Journal of Financial Economics* 70, 313–349.
- Tesar, L., Werner, I., 1995. Home bias and high turnover. *Journal of International Money and Finance* 14, 467–492.
- Tootell, G., 1996. Redlining in Boston: Do mortgage lenders discriminate against neighborhoods? *Quarterly Journal of Economics* 111, 1049–1079.
- Zangwill, I., 1914. *The melting pot: Drama in four acts*. Macmillan, New York.