The Impact of Linguistic Distance and Financial Reporting Readability on Foreign Holdings of U.S. Stocks

Kristian D. Allee, *University of Arkansas*

Lisa S. Anderson, *University of Arkansas*

Michael J. Crawley, *University of Arkansas*†

December 2019

ABSTRACT

Using a sample of 75 countries, we show that foreigners invest less in U.S. stocks when they are from countries with greater linguistic distances and when U.S. financial reports are more difficult to read. Our results suggest that linguistic distance and linguistic complexity in financial reports act as information processing frictions for foreign investors, even in the U.S. market where foreign investors should have the greatest ability, resources, and incentives to overcome language translation and readability issues. Additionally, we show that foreigners invest more in U.S. Treasuries and consume more when facing greater linguistic distance and financial reporting readability difficulties. These results are consistent with a "substitution effect" where foreigners increase their consumption and holdings of U.S. Treasuries because these alternatives avoid the linguistic distance and readability frictions associated with analyzing financial statements and investing in U.S. equities.

Keywords: Linguistic distance; financial reporting readability; foreign investors; U.S. stocks; U.S. Treasuries; consumption

JEL Classifications: M41, F21, G15

†Corresponding author.

1. Introduction

Foreign investment is an essential component of United States (U.S.) security markets (Stulz 1999). Approximately one-third of U.S. stocks are owned by foreign investors, and foreign investment can enable growth and lower U.S. firms' cost of equity (Bump 2017; Henry 2000a, b; Lizardo and Mollick 2009). However, investors face economic frictions and other impediments when investing abroad. These frictions can include (1) legal barriers (e.g. capital controls), (2) information asymmetries, and (3) behavioral factors (French and Poterba 1991; Kang and Stulz 1997; Karolyi and Stulz 2003; Chan et al. 2005; Beugeldsdijk and Frijns 2010).

We examine whether linguistic distance and the readability of U.S. financial reports are frictions that affect foreign investment in U.S. stocks. Linguistic distance and financial reporting readability are related constructs in an international setting because both may affect foreign investors' information processing costs, and in turn, influence the amount of U.S. stocks they are willing to hold. Linguistic distance reflects how different one language is from another (Isphording and Otten 2013). Prior studies examining the economic consequences of language differences generally use binary variables which treat all foreign languages the same (Demirbag et al. 2007). In contrast, we use the "language acquisition approach" from the linguistics literature (Chiswick and Miller 2005) to create a continuous measure of the linguistic distance between English and 43 other languages spoken in a sample of 75 countries. For example, our linguistic distance measure reflects the notion that the difference between English and Dutch (both alphabetic languages that share grammatical structures and Germanic ancestry) is smaller than the difference between English and Japanese.

Moreover, we allow the economic impact of linguistic distance to vary depending on the readability of U.S. financial reports. Financial reporting readability refers to the ease with which

a reader can process written text (Bonsall et al. 2017). While there has been considerable recent research on financial statement readability (Li 2008; Miller 2010; Bonsall et al. 2017), the prior research does not consider the impact of readability on foreign investors. We create a proxy for aggregate U.S. financial reporting readability based on the "Bog Index" created by Bonsall et al. (2017) and allow financial reporting readability to interact with linguistic distance. For example, a Dutch speaking investor facing unreadable U.S. financial reports may alter their investment behavior differently than a Japanese speaking investor facing the same reports.

We demonstrate that linguistic distance and financial reporting readability are both negatively associated with foreign holdings of U.S. stocks. That is, foreign investors invest less in U.S. stocks when they are from countries with greater linguistic distances and when U.S. financial reports are more difficult to read. Furthermore, we also show that linguistic distance and financial reporting readability interact. Specifically, foreign investors from high linguistic distance countries facing unreadable financial reports reduce their investments less than foreign investors from low language distance countries facing the same reports. This suggests that foreign holdings of U.S. stocks are concave with respect to total frictions and that foreign investors are reluctant to reduce their investments in U.S. stocks below a certain floor.

Our results are robust to (1) the inclusion of controls for a country's geographic, demographic, macroeconomic, tax, financial reporting, legal, and cultural environment, (2) the inclusion of controls for U.S. equity and U.S. Treasury returns, oil prices, and the level of voluntary disclosure by U.S. firms, (3) the use of an alternative linguistic distance measure, (4) excluding English-speaking countries and countries where English is an official language, (5) the use of monthly or annual data, (6) excluding offshore banking centers, (7) separate analyses of large and small economies, (8) estimating changes specifications to reduce the potential for non-

stationarity or unobserved time-invariant heterogeneity, and (9) estimating our models excluding one country at a time to rule out significant influence by any single country in our sample.

If foreigners invest less in U.S. stocks in response to linguistic distance and difficulties reading financial reports, then they are potentially allocating a significant amount of capital elsewhere. For example, foreign investors could invest in their home country, invest in other foreign countries, invest in alternative U.S. securities, or consume rather than invest. Based on available data, we explore two of these possibilities. First, we examine U.S. Treasuries as an alternative investment vehicle. U.S. Treasuries allow foreigners to access U.S. financial markets without facing the same information processing frictions. Investing in U.S. Treasuries is less linguistically complex because U.S. Treasuries do not have financial statements to analyze. We find that linguistic distance and financial reporting readability are positively associated with foreign holdings of U.S. Treasuries, and the interaction of linguistic distance and financial reporting readability is negatively associated with foreign holdings of U.S. Treasuries.

Second, we examine foreigners' consumption behavior. We find that linguistic distance and financial reporting readability are positively associated with consumption, and the interaction of linguistic distance and financial reporting readability is negatively associated with consumption. Our results for U.S. Treasuries and consumption are mirror images of the U.S. stock results and are consistent with a "substitution effect." In other words, foreigners who want to invest in U.S. stocks, but are sufficiently deterred by the linguistic distance and financial reporting readability frictions, appear to (1) seek the relative ease of investing in U.S. Treasuries and/or (2) simply invest less and consume more in their home country.

Our study is most closely related to the literature which examines language as an investment friction. Lundholm et al. (2018) measure the "Frenchness" of firms located in the

Canadian province of Quebec by counting the relative number of French versus English language documents resulting from a Google search of the company name. The authors find that U.S. equity ownership is inversely correlated with a firm's Frenchness. Grinblatt and Keloharju (2001) find that Finnish investors prefer to invest in Finnish companies that publish their annual reports in Finnish rather than Swedish. Collectively, these studies suggest that investing in firms that communicate in a different language is costly.

We make significant contributions to the literature examining language as an information processing friction. First, we are the first to show that linguistic distance and financial reporting readability influence foreign holdings of U.S. stocks and U.S. Treasuries. Demonstrating an association between linguistic distance and financial reporting readability on foreign investment in the U.S. is important because the U.S. is the largest and most competitive capital market in the world with a strong information environment and no capital controls (Bradshaw et al. 2004; Ahearne et al. 2004; Stulz and Wasserfallen 1995). Moreover, English is the dominant *lingua franca* ("language of business") around the world (Fredriksson et al. 2006; Neely 2012). Thus, our results suggest that linguistic distance and financial reporting readability persist as significant frictions, even in the market where foreign investors should have the greatest ability, resources, and incentives to overcome translation and readability issues.

Additionally, we use a unique data set to investigate previously unanswerable questions. Examining foreign investment in U.S. firms is difficult due to data limitations. Most studies use institutional investors' Form 13F filings submitted to the Securities and Exchange Commission (SEC) (e.g. Choi et al. 2017). However, using 13F filings poses several problems. First, Form 13F only details an institution's holdings of U.S. equities (not U.S. Treasuries or other assets). Second, foreign institutions are only required to file if the institution (1) has positions exceeding

$100 million, and (2) uses a "means or instrumentality of United States interstate commerce in the course of their business" (e.g. the U.S. mail).[1] Thus, 13F data fails to capture many foreign institutions with significant U.S. equity holdings. Third, Form 13F identifies the country in which an institution is domiciled but not the nationality of the underlying investors.[2]

In contrast to prior studies, we use a novel data set from the U.S. Treasury. The International Investment and Trade in Services Survey Act requires U.S. and foreign financial institutions to file reports with the U.S. Federal Reserve. Treasury International Capital (TIC) Form Securities Long-Term (SLT) requires institutions to aggregate the ownership of U.S. stocks and U.S. Treasuries by the country of citizenship of its account holders.[3] Unlike 13F filings, monthly TIC SLT reports do not specify the individual stocks that investors hold, and individual institutional reports are not publicly available. However, the U.S. Treasury publishes aggregate ownership of U.S. stocks and U.S. Treasuries for all investors from each country on a monthly basis. Using U.S. Treasury data has multiple advantages in our setting. First, our data includes holdings of both U.S. equities and U.S. Treasuries which allows us to examine whether foreign investors make tradeoff decisions about asset allocation and information processing costs. Second, our data covers all foreign investors (individuals, institutions of all sizes, and governments). Third, investors' nationalities are well specified for a large sample of countries.

---

[1] See Section 13(f)(1) of the Securities Exchange Act of 1934 and SEC implementation guidance at https://www.sec.gov/divisions/investment/13ffaq.htm.

[2] For example, see Deutsche Bank's 13F filing for the quarter ending December 31, 2015 at https://www.sec.gov/Archives/edgar/data/948046/000094804616000122/0000948046-16-000122-index.htm. The firm had positions in individual stocks (e.g. 26 million shares of Apple, 32 million shares of Microsoft, etc.) totaling $145 billion under management. However, the 13F filing does not specify whether the stocks are ultimately owned by German account holders, Swiss account holders, French account holders, etc.

[3] Continuing with the prior example, Deutsche Bank would have to summarize the $145 billion of U.S. equities under management at December 31, 2015 by the country of citizenship of its account holders. For example, Deutsche Bank's TIC Form SLT may report collective German ownership of $90 billion, Swiss ownership of $40 billion, and French ownership of $15 billion. See Section 3 for further description of the TIC system.

We also make multiple contributions to the broader "home bias" literature which shows that investors tend to overweight (underweight) domestic (foreign) stocks (see Ardalan 2019 for a recent review). Our results suggest that foreign investors hold fewer U.S. stocks when linguistic distance and difficulties reading U.S. financial reports are greater. These results are consistent with rational explanations for an equity home bias arising from differences in language (Brennan and Cao 1997; Coval and Moskowitz 1999). However, our results also suggest that linguistic distance and financial reporting readability frictions are positively associated with foreign U.S. Treasury holdings. These results are the *opposite* of a home bias and are unique because investment frictions generally reduce demand. Our results suggest that foreign investors use U.S. Treasuries as a substitute for U.S. stocks when language distance and readability costs are sufficiently high. Therefore, foreign investors' exposure to the U.S. market could be closer to optimal levels as specified by theoretical home bias models once their combined holdings of U.S. equities and U.S. Treasuries are considered.

Second, our results suggest that linguistic distance and financial reporting readability can alter the relative attractiveness of U.S. stocks, U.S. Treasuries, and consumption. Altering the demand for Treasuries may change Treasury yields and the risk-free rate for all firms. Similarly, consumption-based asset pricing models suggest that changes in the marginal utilities of investment and consumption may alter firms' betas (Cochrane 2009). Thus, our results suggest that linguistic distance, financial reporting readability, and the interaction between the two have the potential to influence firms' cost of capital and overall macroeconomic growth.

Finally, our results may be of interest to regulators. Aside from requiring firms to issue disclosures in foreign languages, policy makers can do little about linguistic distances. However, regulators may be able to influence the readability of U.S. financial reports. The SEC has long

been concerned that firms fail to make their filings readable and understandable (Firtel 1999). The SEC issued plain English guidelines in 1998 (SEC 1998) and has considered using a readability measure to assess compliance (Loughran and McDonald 2014). Our results suggest that a lack of readability imposes a unique information externality on foreign investors (a large but overlooked population in the literature) and that foreign investors are heterogeneous with respect to translation costs. Regulators may wish to more explicitly consider the ramifications of financial reporting readability for this economically important community, particularly given recent trends towards less readable reports and more foreign investment. Lastly, our results suggest that the readability of U.S. firms' financial reports can influence demand for U.S. Treasuries that do not have financial statements. Thus, regulations designed to alter the readability of U.S. financial reports could induce externalities in alternative asset markets.

This study is organized as follows. Section 2 motivates our hypotheses. Section 3 describes our data and sample. Section 4 describes our methodology. Section 5 presents the results, and Section 6 presents additional analyses on the substitution effect. Section 7 concludes.

2. Background and Hypothesis Development

Linguistic Distance

As described above, prior research in accounting and finance shows that language differences can act as a market friction (Lundholm et al. 2018; Grinblatt and Keloharju 2001). Relatedly, Baik et al. (2013) show that increases in foreign institutional ownership of U.S. stocks are negatively correlated with future returns. This result is consistent with language and cultural differences creating a "liability of foreignness" that impairs foreign investors' ability to predict returns. Lundholm et al. (2014) show that foreign firms with U.S. stock exchange listings write more readable financial reports than their U.S. counterparts to lessen the psychological distance

between the firm and prospective U.S. investors. This result is consistent with experimental evidence which shows that concrete language increases investors' willingness to invest, particularly when investors are psychologically distant from the firm (Elliot et al. 2015).

We contribute to the literature examining language as an investment friction by investigating the association between linguistic distance and foreign investors' demand for U.S. stocks. At the construct level, linguistics researchers generally agree that linguistic distance represents how different one language is from another. However, languages are complex (e.g. they differ with respect to vocabulary, grammar, syntax, written form, etc.), and researchers do not agree on the best way to measure linguistic distance (Chiswick and Miller 2005; Isphording and Otten 2013). There are at least three methods to measure linguistic distance. First, some studies adopt an anthropologic approach (e.g. Guiso et al. 2009). The underlying theory is that all languages descend from a small number of languages, and language diversity is the result of prehistoric *Homo sapien* migration "out of Africa." According to this view, languages can be represented in the form of a tree (like genealogical trees). Languages that share more (fewer) branches are deemed to have a lower (greater) linguistic distance (Ginsburgh and Weber 2016).

Second, other studies use a phonetic approach to measure linguistic distance (e.g. Isphording 2014; Isphording and Otten 2013). Researchers identify words having the same meaning in two different languages (e.g. "hello" and "hola") and use software to measure how much the pronunciations differ. Languages with a greater (lower) average pronunciation similarity across a set of words are deemed to have a lower (greater) linguistic distance.

Finally, Chiswick and Miller (2005) use a language acquisition approach. The idea is to measure how difficult it is to learn to speak another language. Conveniently, the U.S. Department of State collects data on how difficult it is for native English speakers to learn other

languages. The U.S. Department of State's School of Language Studies (SLS) administers language training protocols in over 70 languages to native English-speaking U.S. government employees (e.g. diplomats). At the end of each protocol, the SLS administers standardized tests to assess the employee's reading and speaking proficiency. Higher (lower) average test scores indicate a lower (greater) linguistic distance. Data for a sample of SLS test subjects was initially published by Hart-Gonzalez and Lindemann (1993) and first used as a measure of linguistic distance in the academic literature by Chiswick and Miller (2005).

We use the Chiswick and Miller (2005) measure of linguistic distance as our primary measure for three reasons. First, the language acquisition approach is simple and atheoretic. The approach does not rely on assumptions about the functional form of language trees, the validity of theories about prehistoric migration patterns, or the appropriate way to measure differences in pronunciation. Second, the Chiswick and Miller (2005) measure is the only proxy measuring how difficult it was for humans to process and respond to communication in another language. Third, the Chiswick and Miller (2005) measure is publicly available for many languages.

One criticism of the Chiswick and Miller (2005) measure is that linguistic distances only exist relative to English (Isphording 2014), but this restriction does not apply in our study because we only require a proxy for the distance between English and other languages. Another concern about the Chiswick and Miller (2005) measure is the lack of symmetry (i.e. the difficulty for a native English speaker to learn Spanish may not necessarily be the same as the difficulty for a native Spanish speaker to learn English). Hence, we use an alternate symmetric measure of linguistic distance using the phonetic approach as a robustness check (see Section 5).

We expect linguistic distance to increase foreign investors' information processing costs, and in turn, reduce the amount of U.S. stocks they are willing to hold. However, there are

multiple reasons why we may fail to find evidence supporting this prediction. First, the U.S. is the largest and most competitive capital market in the world with a strong information environment and no capital controls (Bradshaw et al. 2004; Ahearne et al. 2004; Stulz and Wasserfallen 1995). Second, English is the dominant language of business around the world (Fredriksson et al. 2006; Neely 2012). Third, tens of millions of foreign citizens have studied at universities where instruction is delivered in English, including many who study business and go on to be investors (Tietze 2004). Overall, foreign investors investing in the U.S. are likely to be relatively sophisticated and may have the ability, resources, and economic incentives to overcome translation issues. Therefore, we state our first hypothesis in null form:

HYPOTHESIS 1. *Linguistic distance has no impact on foreign holdings of U.S. stocks.*

Financial Reporting Readability

Firm financial reports are an important element of the U.S. information environment (Kothari 2001), and both foreign and domestic investors attempting to process the information within financial statements face a non-trivial task. Attempts to quantify the difficulty of processing written English text date back to the Gunning (1952) Fog Index. However, the use of readability measures in accounting and finance research was limited until relatively recently. Li (2008) shows that longer and less readable 10-K filings are associated with lower profitability and earnings persistence, consistent with managers manipulating the readability of financial reports to obfuscate poor performance. Subsequent research shows that such opportunistic behavior can affect the decision usefulness of financial statements for investors (You and Zhang 2009; Miller 2010; Rennekamp 2012; Lawrence 2013), analysts (Lehavy et al. 2011; Bozanic and Thevenot 2015), and credit rating agencies (Bonsall and Miller 2017).

The empirical proxies for financial reporting readability have evolved over time. Most early studies use the Gunning (1952) Fog Index which measures average sentence length and the

percentage of complex words (i.e. words with more than three syllables). However, Loughran and McDonald (2014) note that financial reports contain a large percentage of words with more than three syllables that are not difficult to understand for most financial statement users (e.g. depreciation, liability). As such, the authors propose the file size of the firm's electronic Form 10-K filing as an alternative readability measure. More recently, Bonsall et al. (2017) use computational linguistics software to create a "Bog Index" which captures plain English attributes of disclosure (e.g. avoiding passive voice, weak verbs, complex words, and jargon).

We use the Bonsall et al. (2017) Bog Index as the starting point for our measure of aggregate U.S. financial reporting readability for several reasons. First, the Bog Index uses a dictionary which avoids overstating word difficulty based on syllable counts. Second, the Bog Index captures the most dimensions of the SEC's plain English guidelines. Third, Bonsall et al. (2017) use an experiment and validate that human subjects find high Bog score documents harder to read (controlling for word count, file size, formatting, and Fog score). Fourth, Bog scores for a large sample of 10-K filings over multiple years are publicly available.

We expect financial reporting readability (i.e., financial reports that are more difficult to read) to increase foreign investors' information processing costs, and in turn, reduce the amount of U.S. stocks they are willing to hold. However, there are multiple reasons why we may fail to find evidence supporting this prediction. First, Guay et al. (2016) find that firms with less readable financial reports engage in more voluntary disclosure which suggests managers mitigate the negative effects of unreadable reports. Second, foreigners investing in the U.S. are likely to be relatively sophisticated and may have the ability, resources, and incentives to minimize the impact of readability frictions. Therefore, we state our second hypothesis in null form:

> HYPOTHESIS 2. *Financial reporting readability has no impact on foreign holdings of U.S. stocks.*

Linguistic Distance and Financial Reporting Readability

Linguistic distance and financial reporting readability may interact and influence the amount of U.S. stocks that foreign investors hold. On one hand, a foreign investor from a country with a high linguistic distance facing unreadable U.S. financial reports may further reduce her U.S. stock holdings due to the compounding of multiple frictions (i.e. the interaction could amplify the main effects of linguistic distance and financial reporting readability).

On the other hand, there are at least two reasons why the interaction could attenuate the main effects of linguistic distance and financial reporting readability. First, foreign holdings of U.S. stocks may be concave with respect to total frictions. For example, consider a native French speaker contemplating investing in U.S. stocks. The linguistic distance between French and English is relatively small. Both languages use the Latin alphabet, both use similar grammatical structures, and both share hundreds of "cognates" (words that have a common etymological origin and which look and/or are pronounced alike). The French speaker may be able to process highly readable U.S. financial reports, but she may significantly reduce her holdings if U.S. financial reports are sufficiently difficult to process. In contrast, consider the case of a native Mandarin Chinese speaker. The linguistic distance between Mandarin and English is large. Mandarin is a "logographic" language that uses symbols to represent words (in contrast to English which uses an alphabetic system where letters represent sounds). Mandarin and English also differ with respect to phonology (e.g. tone) and grammar (e.g. verb tense, punctuation, and the use of articles). Processing U.S. financial reports may be difficult for the Mandarin speaker regardless of their readability, and thus the Mandarin speaker may reduce his U.S. stock holdings less than the French speaker when facing less readable reports.

Second, if a foreign investor with a high linguistic distance facing unreadable U.S. financial reports invests less in U.S. stocks, then where might she direct those funds? Investing in her home country will reduce her global diversification and prevent her from enjoying the levels of investor protection, financial reporting quality, and financial market development available in the U.S. Additionally, she's unlikely to escape language translation and financial reporting readability issues if she invests in another foreign country. For example, an Egyptian investor will also face a high language distance if she decides to invest those funds in German stocks, and German financial reports may be difficult to read as well. Consequently, U.S. stocks may still be her first best option for at least a minimum portion of her portfolio.

Ultimately, the impact of the interaction between linguistic distance and financial reporting readability is an empirical question, and we state our final hypothesis in null form:

> HYPOTHESIS 3. *The interaction of linguistic distance and financial reporting readability has no impact on foreign holdings of U.S. stocks.*

3. Data and Sample

Foreign Ownership of U.S. Stocks

The U.S. Treasury International Capital reporting system monitors ownership of U.S. assets by foreign investors and helps policy makers monitor systemic risks in the financial system (Brandner et al. 2012). Commercial banks, securities dealers, and other financial institutions are required to file various reports with district U.S. Federal Reserve Banks. Reporting is mandatory for U.S. entities and foreign entities pursuant to the International Investment and Trade in Services Survey Act (22 U.S.C. 3101 et seq.), and penalties can be imposed for non-compliance (U.S. Treasury 2016). TIC Form SLT requires each reporting entity to aggregate the ownership of U.S. stocks (at nominal current market value) by the country of citizenship of its account holders by month. An institution's monthly report does not include the

identities of its account holders or any other account-level information (e.g. which individual stocks they hold), and individual reports are not available to the public. However, reported values for all institutions with foreign holders of U.S. stocks are aggregated by country. The resulting total ownership in U.S. stocks for investors from country i at the end of month j ($CS_{i,j}$) is available continuously beginning in December 2011.[4] We deflate holdings by the annual price index from U.S. Bureau of Economic Analysis (BEA) National Income and Product Account (NIPA) Table 1.1.4 to convert CS values to millions of real 2010 U.S. dollars.[5]

Linguistic Distance

TIC data on foreign holdings of U.S. stocks are at the country level. Thus, we require a measure of linguistic distance by country. We begin with the measure of linguistic distance between English and 43 other languages based on how difficult it was for a sample of English-speaking U.S. State Department employees to learn those languages (Chiswick and Miller 2005; Hart-Gonzalez and Lindemann 1993). The individual linguistic distance scores are publicly available, see Table 1 of Chiswick and Miller (2005). We create a mapping from languages to countries by weighting the linguistic distance of the languages spoken within a country by the percentage of the population speaking each language.[6] Finally, we subtract each value from 4 so that the resulting time-invariant measure of linguistic distance for country i (LD_i) varies from 0 to 3 with higher values denoting greater linguistic distance.

Financial Reporting Readability

---

[4] See http://ticdata.treasury.gov/Publish/fslt-dec2013.pdf for the official TIC Form SLT from the U.S. Treasury. See Table 1D in Section B.2 at https://www.treasury.gov/resource-center/data-chart-center/tic/Pages/ticsec2.aspx for aggregated U.S. stock holdings by country-month.
[5] See BEA NIPA Table 1.1.4 at https://www.bea.gov/iTable/index_nipa.cfm. Results are quantitatively and qualitatively similar using nominal or real holdings values.
[6] The percentage of the population speaking each language within a country is from the Central Intelligence Agency (CIA) Factbook at https://www.cia.gov/library/publications/the-world-factbook/fields/2098.html.

TIC data on foreign holdings of U.S. stocks are at a monthly frequency, and thus we require a monthly aggregate measure of U.S. financial reporting readability. We begin with the Bog scores for firm-level Form 10-K filings created by Bonsall et al. (2017).[7] We create a proxy for aggregate U.S. financial reporting readability by calculating an equal weighted average of the Bog scores for all firms filing a 10-K over the past 12 months from month j-12 to month j-1.[8] Finally, we subtract the sample mean of the time series from each month's value.[9] The resulting aggregate measure of financial reporting readability for month j-1 (BOG_{j-1}) varies from -1.01 to 2.24 with higher values denoting less readability.

Sample

The TIC data described above provides monthly aggregate holdings of U.S. stocks by country for 125 countries. Requiring control variable data (discussed in the next section) reduces our sample to 75 countries. Our sample of 75 countries is quite large compared to other studies in the international literature. For example, Lundholm et al. (2014) examine the disclosure behavior of firms from 45 countries who list on a U.S. stock exchange. Moreover, the countries in our sample exhibit significant cross-sectional variation in size and economic importance.[10]

Table 1 presents linguistic distances and average U.S. stock holdings over the sample period for the 75 countries in our sample (presented in descending order by linguistic distance). Table 1 also presents each country's average holdings of U.S. Treasuries (*US*) and the country's average household consumption expenditures (*CONS*) for the substitution tests performed later in the study. The largest linguistic distances are areas where logographic languages are dominant

---

[7] Firm filing-level Bog data are available at https://kelley.iu.edu/bpm/activities/bogindex.html.
[8] Results are quantitatively and qualitatively similar using an aggregated *BOG* measure weighted by firm size.
[9] "Centering" (i.e. de-meaning) helps facilitate interpretation of the regression main effects in the presence of an interaction term. See Section 5 for further discussion.
[10] Our sample includes all possible G-20 countries. The G-20 is a group of the largest and most important economies which promotes international financial stability. Our sample excludes the U.S. (by construction) and the European Union (TIC data tracks holdings for the individual countries) but includes the other 18 members.

(e.g. Japan's $LD = 3$, China's $LD = 2.58$). The countries with the next largest linguistic distances (e.g. Kuwait and Saudi Arabia with $LD = 2.5$) share Arabic as the dominant language. Arabic is an alphabetic language, but it uses its own alphabet (rather than the Latin alphabet used in English).[11] The countries with the lowest linguistic distances (United Kingdom, Ireland, Australia, New Zealand, Trinidad and Tobago, and Jamaica where $LD = 0$) are or were British settlements. The largest holders of U.S. stocks over the sample period are the United Kingdom ($668 billion) and Canada ($649 billion). These countries have two of the lowest linguistic distances, consistent with the alternative hypothesis for Hypothesis 1. China, a country with one of the highest linguistic distances, has the largest average holdings of U.S. Treasuries (*US*) over the sample period ($1.15 Trillion), consistent with the substitution effect discussed in Section 6.

Figure 1 plots the time-series of aggregate foreign holdings of U.S. stocks for the 75 countries in our sample. The solid (dotted) line shows that foreign holdings of U.S. stocks (U.S. Treasuries) are increasing over the sample period, but the trend is not monotonic. The dashed line plots the time series of *BOG*. The increasing trend suggests that U.S. financial reports have become less readable over time, consistent with Bonsall et al. (2017).

4. Methodology

We empirically test Hypothesis 1, Hypothesis 2, and Hypothesis 3 simultaneously using the following linear regression where the unit of observation is country-month:

$$CS_{i,j} = \beta_0 + \beta_1 LD_i + \beta_2 BOG_{j-1} + \beta_3 LD_i * BOG_{j-1} + \beta_4 CS_{i,j-12} + \beta_5 EQUITY_{j-1} + \beta_6 TREAS_{j-1} +$$

$$\beta_7 OIL_{j-1} + \beta_8 DISCLOSE_{t-1} + \beta_9 POP_{i,t-1} + \beta_{10} BILAT_{i,t-1} + \beta_{11} GDPPC_{i,t-1} + \beta_{12} \Delta GDPPC_{i,t-1}$$

[11] Other systems exist including pictographic systems (e.g. ancient Aztec) and syllabaries (e.g. Cherokee). While Korean has syllabic elements (Taylor 1980), we do not have any "pure" pictographic or syllabic languages in our sample. However, some such languages have historical significance. For example, U.S. armed forces used "Code Talkers" to transmit coded messages based on Native American languages (e.g. Navajo and Choctaw) during World War II. Native American languages were useful because non-native speakers found them difficult to understand, few people outside of the tribes knew the languages, and virtually no books had been published in those languages. The Japanese never cracked these codes (Arbuckle 2017), a testament to the power of linguistic distance.

$$+ \beta_{13}INF_{i,t-1} + \beta_{14}G20_i + \beta_{15}TAX_i + \beta_{16}OFFSHORE_i + \beta_{17}FINSTD_i + \beta_{18}COMMON_i +$$

$$\beta_{19}PROTECT_i + \beta_{20}DIST_i + \beta_{21}POWERDISTANCE_i + \beta_{22}INDIVIDUALISM_i +$$

$$\beta_{23}MASCULINITY_i + \beta_{24}UNCERTAINTYAVOIDANCE_i + \beta_{25}LONGTERM_i +$$

$$\beta_{26}INDULGENCE_i + \varepsilon_{i,j} \tag{1}$$

Variables not discussed in Section 3 are defined below. All variables are also defined in the Appendix. The Δ operator denotes the percentage change from year t-2 to year t-1. The β_1, β_2, and β_3 coefficients test the null for Hypothesis 1, Hypothesis 2, and Hypothesis 3, respectively.

With respect to the control variables, we include $CS_{i,j-12}$ (aggregate U.S. stocks held by investors from country i from 12 months ago). We expect the β_4 coefficient to be positive to the extent that foreign investment levels are "sticky". For example, investors in a country may not all jointly adjust their positions in response to changes in macroeconomic conditions due to transaction costs and the potential for large price impacts by moving simultaneously.

The next group of controls are time-varying but not country specific. *EQUITY*, *TREAS*, and *OIL* are proxies for global and U.S. economic conditions. *EQUITY$_{j-1}$* and *TREAS$_{j-1}$* are monthly returns on the Center for Research in Securities Prices (CRSP) value-weighted index and 30-year U.S. Treasury bonds from month j-12 to month j-1 from the CRSP database, respectively. We expect the β_5 coefficient to be positive because foreign investors' current U.S. equity holdings are a function of their past holdings and recent U.S. equity market returns, and foreign investors "chase" returns (Lin and Swanson 2003). U.S. stock and U.S. Treasury returns are typically negatively correlated, but our sample includes periods of quantitative easing by the U.S. Federal Reserve which resulted in simultaneous rallies in both markets. As such, we make no predictions for the β_6 coefficient. Monthly spot prices of Brent crude oil at the end of month j-

1 (OIL_{j-1}) are from the Federal Reserve Economic Database.[12] We expect the β_7 coefficient on

OIL to be positive because oil prices tend to be pro-cyclical, and foreign investors should have

more money to invest when global macroeconomic conditions are strong. *DISCLOSE* is a proxy

for the level of voluntary disclosure by U.S. firms calculated as the percentage of firms on

Compustat that have held at least one conference call within the past 12 months according to

Seeking Alpha. We include *DISCLOSE* because prior research suggests that U.S. firms may try

to mitigate financial reporting readability difficulties by increasing voluntary disclosure (Guay et

al. 2016). We expect foreign investors to invest more in U.S. stocks when U.S. firms engage in

more voluntary disclosure, and thus we expect the β_8 coefficient on *DISCLOSE* to be positive.

The next group of controls (*POP*, *BILAT, GDPPC, ΔGDPPC, INF*, and *G20*) capture

demographic and macroeconomic conditions in the investor's home country. Annual population

in millions ($POP_{i,t-1}$), gross domestic product per capita in real 2010 U.S. dollars ($GDPPC_{i,t-1}$),

and inflation ($INF_{i,t-1}$) for country *i* at the end of year *t*-1 (i.e. the most recent calendar year

ending before month *j*) are available from the World Bank.[13] We make no predictions on the

signs of the β_9, β_{11}, or β_{12} coefficients because investors from larger, more efficient, and faster

growing economies may have more funds to invest, but they may invest less in the U.S. because

there are more investment opportunities in their home country. $BILAT_{i,t-1}$ denotes bilateral trade

(i.e. the sum of imports and exports) between country *i* and the U.S. in year *t*-1 from the United

Nations Commodity Trade Statistics Database.[14] $G20_i$ is a dummy variable that equals 1 if

country *i* is a member of the G-20, and $G20_i$ equals 0 otherwise. We expect the β_{10} and β_{14}

coefficients to be positive because greater engagement with the U.S. (via bilateral trade or via

---

[12] See https://fred.stlouisfed.org/series/MCOILBRENTEU.
[13] See https://data.worldbank.org/indicator.
[14] See https://comtrade.un.org/. Results are quantitatively similar bifurcating *BILAT* into exports and imports.

participation in G-20 treaties) could lead to greater U.S. equity investment. We expect the β_{13} coefficient on *INF* to be positive because prior research demonstrates that investing in other countries is one way to hedge against home country inflation (Sayek 2009).

The next group of control variables (*TAX*, *OFFSHORE*, and *FINSTD)* are time-invariant measures of the investor's home country tax and financial environments. Estimates of the average corporate tax rate (including corporate income taxes, payroll taxes, retirement contributions, and other taxes) (*TAX$_i$*) for country *i* are available from the World Economic Forum.[15] We expect the β_{15} coefficient to be negative, consistent with research which shows that high corporate tax rates are associated with lower foreign investment (Clark 2008). *OFFSHORE$_i$* is a dummy variable that equals 1 if the U.S. Treasury has identified country *i* as an "offshore banking center", and *OFFSHORE$_i$* equals 0 otherwise.[16] We expect the β_{16} coefficient to be positive if banks in offshore financial centers classify some foreign account holders as domestic investors.[17] Finally, *FINSTD$_i$* is a measure of the strength of the financial reporting and auditing standards in country *i* from the World Economic Forum. *FINSTD$_i$* values are integers ranging from 1 to 7 where higher (lower) values denote stronger (weaker) standards. We expect the β_{17} coefficient to be negative, consistent with research showing that investors in countries with stronger reporting and auditing standards have less incentive to invest abroad (Bradshaw et al. 2004; Aggarwal et al. 2005; Khurana and Michas 2011; Baik et al. 2013).

---

[15] See http://reports.weforum.org/global-competitiveness-index-2017-2018/competitiveness-rankings/.
[16] Recent data leaks (e.g. "Wikileaks", the "Panama Papers", and the "Paradise Papers") reveal that wealthy individuals often use financial institutions in certain countries as intermediaries to evade taxes and maintain privacy (Haldevang 2017; Enrich 2018; Pacaud 2017). The U.S. Treasury monitors countries to identify improper behavior. The only countries identified by the U.S. Treasury as offshore financial centers in our sample are Luxembourg, Switzerland, Belgium, and the United Kingdom. All other countries identified by the U.S. Treasury as offshore financial centers do not have the necessary data to be included in our sample (e.g. Cyprus, Cayman Islands, etc.).
[17] For example, a Russian investor may have a numbered account at a Swiss bank (i.e. an account that does not include any identifying information about the account holder). If the Russian investor uses funds from the numbered account to invest in U.S. stocks, then the Swiss bank may classify the account holder as Swiss (rather than Russian) on TIC Form SLT to mask the account holder's nationality.

The next set of variables (*COMMON* and *PROTECT*) are time-invariant measures of the investor's home country legal environment. $COMMON_i$ is a dummy variable that equals one if country *i* is a former British colony or is listed as a common law country in Table 2 of LaPorta et al. (1998), and $COMMON_i$ equals zero otherwise. We expect the β_{18} coefficient to be positive, consistent with individuals from common law countries preferring to invest in another common law country (Khurana and Michas 2011; Bae et al. 2008a, b). $PROTECT_i$ is a measure of the strength of investor protections in country *i* from the World Economic Forum. $PROTECT_i$ values are integers from 1 to 7 where higher (lower) values denote stronger (weaker) protections. We expect the β_{19} coefficient to be negative, consistent with research showing that investors in countries with stronger investor protections have less incentive to invest abroad (Aggarwal et al. 2005; Khurana and Michas 2011; Baik et al. 2013).

Next, $DIST_i$ is the distance between the capital of country *i* and New York City in kilometers.[18] Geographic distance can be an element of the "psychic distance" between two cultures (Lopez-Duarte and Vidal-Suarez 2010; Lundholm et al. 2014). Prior research shows that greater geographic distance is negatively correlated with the trade of physical goods (Anderson 2011) and financial investment (Portes et al. 2001; Portes and Rey 2005; Ferreira and Matos 2008; Baik et al. 2013). As such, we expect the β_{20} coefficient to be negative.

Finally, we include a set of six time-invariant proxies for the home country cultural environment developed by Hofstede (2001). Gorodnichenko and Roland (2011) review existing cross-country measures of culture and conclude that the Hofstede (2001) measures have been empirically validated and are the most correlated with economic behavior. $UNCERTAINTYAVOIDANCE_i$ is a proxy for risk aversion which prior research has shown to be

---

[18] Pairwise distance searches are from https://www.timeanddate.com/worldclock/distance.html.

correlated with investment behavior (Gandelman and Hernandez-Murillo 2014). Higher values denote a culture more threatened by ambiguity. The remaining culture measures include *POWERDISTANCE$_i$* (higher values denote a greater deference to people in power), *INDIVIDUALISM$_i$* (higher [lower] values denote that the individual [group] is the most important social unit), *MASCULINITY$_i$* (higher values denote emphasis on ambition, wealth acquisition, and differentiated gender roles whereas lower values denote emphasis on nurturing behaviors, sexual equality, environmental awareness, and more fluid gender roles), *LONGTERM$_i$* (higher values denote a culture more prepared to delay short-term gratification), and *INDULGENCE$_i$* (higher values denote a culture more prepared to allow relatively free gratification of natural human desires).[19] We make no predictions on the sign of the β_{21}-β_{26} coefficients because it is not clear how the cultural categories will influence foreign investment.

5. Results

Descriptive Statistics

Table 2 presents descriptive statistics for the variables in model (1). Requiring non-missing control variables results in a sample of 3,582 country-month observations from December 2012 to December 2016. The mean *CS* value of $63,224 indicates that U.S. stock holdings averaged $63 billion.[20] The mean (median) linguistic distance is 1.65 (1.75) which approximates the linguistic distance of Switzerland (Germany). The mean (median) *BOG* value of 0.16 (-0.15) approximates 0 after centering the variable.

---

[19] Culture scores are available from https://www.hofstede-insights.com/product/compare-countries/.
[20] Holding levels are right skewed, and skewness is addressed in multiple ways. First, the inclusion of the control variables limits the possibility for heteroskedasticity. For example, China's holdings are significantly larger than the median holdings, but so is China's lagged holdings, population, etc. Second, we use robust standard errors (clustered by country). Finally, we estimate changes models.

With respect to the control variables, Table 2 shows that equity (30-year U.S. Treasury) returns averaged 12.45% (5.19%) per year. The average oil price was $79.84 per barrel. The average country's population was 72.58 million, and G-20 countries constitute 25% of our sample. The average country's bilateral trade with the U.S. was $47 billion, and the average country's GDP per capita was $23,149. The average country's inflation was 3.10% per year, and the average country's corporate tax rate was 40.87%. Offshore banking center (common law) countries constitute 6% (23%) of our sample. The average country capital is separated from New York City by 8,019 kilometers. The mean *FINSTD* and *PROTECT* values of 5.07 and 6.07, respectively, are more difficult to interpret because the variables are simply integers from 1 to 7. Similarly, the mean values for the culture variables have no real economic connotation because of the arbitrary scoring methodology. Overall, our descriptive statistics appear reasonable.

Table 3 presents univariate correlations. Correlations significant at the 5% level are in bold, and we discuss only the Pearson correlations between the primary variables of interest for brevity. The significantly negative correlation between *CS* and *LD* of -0.28 is consistent with investors from countries with larger linguistic distances investing less in U.S. stocks. The significantly positive correlation between *CS* and *BOG* of 0.04 is inconsistent with foreign investors holding less U.S. stocks when U.S. financial reports are more difficult to read. However, this positive correlation is likely the result of both variables trending upward (see Figure 1). We control for time trends in the regressions by including year fixed effects.

Primary Results

Table 4 Panel A presents the results from regression model (1). We estimate regression model (1) with year fixed effects and cluster the standard errors by country. With respect to the variables of interest, the β_1 coefficient on *LD* is negative and significant at the 5% level. This

result suggests rejection of the null hypothesis for Hypothesis 1 in favor of the alternative that investors from countries with greater linguistic distances hold less U.S. stocks. Interpreting the magnitude of the *LD* coefficient is complicated because linguistic distance is fixed for a country over time. Nevertheless, our results suggest that a one-unit change in *LD* (which would be equivalent to Italians suddenly speaking Arabic) is associated with a reduction in U.S. stock holdings by an economically significant $894 million. Next, the β_2 coefficient on *BOG* of -9,022.46 is negative and significant at the 1% level. This result suggests rejection of the null hypothesis for Hypothesis 2 in favor of the alternative that foreigners invest less in U.S. stocks when U.S. financial reports are more difficult to read. Interpreting the magnitude of the main effect coefficients requires holding all other variables constant (including the interaction). Thus, the magnitude of the β_2 coefficient in the presence of the interaction term reflects the impact of a one-unit change in *BOG* when *LD*=0. Therefore, the results suggest that a one-unit change in *BOG* (approximately 1.5 years of change at the average growth rate) reduces holdings of U.S. stocks for investors in English-speaking countries where *LD*=0 (i.e. the U.K., Ireland, Australia, New Zealand, Trinidad and Tobago, and Jamaica) by an economically significant $9.0 billion.

Next, the β_3 coefficient on the interaction of *LD* and *BOG* of 5,218.70 is significantly positive. Notably, the β_3 coefficient on the interaction term is less in absolute magnitude than the coefficient on *BOG* which suggests the interaction attenuates but does not fully eliminate the impact of financial reporting readability. More formally, the negative β_3 coefficient suggests rejection of the null hypothesis for Hypothesis 3 in favor of the alternative that foreign investors from high linguistic distance countries facing unreadable financial reports reduce their investments less than foreign investors from low language distance countries facing the same unreadable financial reports. This suggests that foreign holdings of U.S. stocks are concave with

respect to total frictions and that foreign investors are reluctant to reduce their investments in U.S. stocks below a certain floor (perhaps because their next best option involves either investing in their home country which reduces global diversification or investing in another country where they will still likely face a linguistic distance and financial reports that are difficult to read). Table 4 Panel A also shows that all control variables with a directional prediction are significant at least the 5% level except *PROTECT*. Finally, the adjusted R^2 of 0.981 suggests our model explains nearly all variation in foreign holdings of U.S. stocks.

One econometric concern with the levels specification presented in Table 4 Panel A is that regression model (1) is an autoregressive model. The coefficient on the lagged holdings variable is statistically significant, which suggests that including the lag term is necessary to prevent autocorrelation in the residuals. However, the coefficient on the lagged holdings variable is indistinguishable from one which raises concerns about a unit root (i.e. non-stationarity) and potentially inflates the R^2. Another concern with model (1) is the potential for unobserved time-invariant heterogeneity. One solution for both concerns is the use of changes specifications.[21]

We estimate two changes analyses. First, we subtract the lagged holdings variable from both sides of model, and the dependent variable becomes ($CS_{i,j} - CS_{i,j-12}$). The result is a "pseudo changes" model in that the dependent variable is an unscaled changes variable, but the remaining independent variables are unaltered. In a second changes specification, we scale the change variable by GDP to reduce heteroskedasticity and transform the independent variables to changes variables where possible. Specifically, ΔBOG, ΔOIL, $\Delta DISCLOSE$, ΔPOP, and $\Delta BILAT$

---

[21] An alternative approach for the potential presence of unobserved time-invariant heterogeneity is a fixed effects model. Unfortunately, estimating a model with country fixed effects is problematic in our setting because one of the primary variables of interest (*LD*) and multiple control variables are all time-invariant for a given country. Additionally, introducing fixed effects into an autoregressive model results in biased parameters (Nickell 1981).

represent percentage changes from twelve months prior. *EQUITY*, *TREAS*, and $\Delta GDPPC$ are already changes variables, and the time-invariant independent variables remain as levels.

Table 4 Panel B presents the results of the pseudo changes model. The coefficients on *LD* and *BOG* remain negative and significant while the interaction coefficient remains positive and significant. The magnitude of these coefficients closely approximates those in Table 4 Panel A, and the behavior of all control variables also matches those in the primary analysis. Finally, the adjusted R^2 drops to 0.284, consistent with moving from a levels model to a changes model.

Table 4 Panel C presents the scaled changes specification. The negative *LD* coefficient suggests that a one-unit change in linguistic distance reduces a country's holdings of U.S. stocks by an economically meaningful 1.43% of the country's GDP (when ΔBOG equals zero). Similarly, the negative ΔBOG coefficient suggests that a one-unit increase in *BOG* for countries where *LD*=0 reduces a country's holdings of U.S. stocks by 1.69% of GDP, although ΔBOG is only marginally significant at the 10% level. The coefficient on the interaction term of 1.09% is significant but smaller in magnitude than the coefficient on *BOG*, suggesting that the interaction attenuates but does not fully eliminate the impact of a change in financial reporting readability on linguistically distant countries. While neither Panel B nor Panel C is a "pure" changes analysis (because *LD* is time-invariant), the inferences are consistent with the levels analysis.

Overall, the results in Table 4 suggest that linguistic distance and financial reporting readability are negatively associated with foreign holdings of U.S. stocks. We acknowledge the limitations of our aggregate data as a caveat to our results and inferences. For example, we cannot identify investor-level characteristics other than nationality (e.g. whether the investor is an individual, institutional investor, firm, or government or whether the investor is sophisticated

or unsophisticated). Similarly, we are unable to identify holdings characteristics (e.g. whether foreign investors tend to own stocks with easy or difficult to read financial reports).

Robustness Checks

The results in Table 4 are consistent with linguistic distance and financial reporting readability reducing foreign investors' willingness to hold U.S. stocks. Table 5 presents the results from a variety of robustness tests. Panel A of Table 5 replicates regression model (1) utilizing an alternative linguistic distance from Isphording (2014). As discussed in Section 2, the Isphording (2014) measure has the advantage of being symmetric (although the phonetic approach relies on an algorithm to measure pronunciation similarity). The coefficients on *LD* and *BOG* remain negative and significant while the interaction term remains positive and significant. Additionally, the control variables all behave similarly to the main results in Table 4. Overall, inferences are similar regardless of which linguistic distance measure is used.

Next, Demirbag et al. (2007) observe that most studies examining the consequences of language use a simple dummy variable for a common language which implicitly treats all foreign languages the same (e.g. Lundholm et al. 2014; Baik et al. 2013). In order to more closely replicate this approach from the prior literature and create a sample where a common language dummy variable would equal zero, we exclude the United Kingdom, Ireland, Australia, New Zealand, Trinidad and Tobago, and Jamaica where *LD* = 0. We also exclude countries where English is a legal official language (e.g. Hong Kong and Singapore) because English may be widely used. Panel B of Table 5 shows that the coefficients on *LD* and *BOG* remain negative and significant while the interaction term remains positive and significant. These results suggest that other studies could benefit from using continuous measures of linguistic distance rather than dummy variables in order to allow the economic consequences to differ across languages.

Our next robustness check uses only annual data. As described previously, TIC holdings data are monthly while some control variables (e.g. World Bank data) are annual. In order to eliminate this inconsistency, we use only TIC holdings data at the end of January of each year. Using January data minimizes the time between the TIC holdings data and the associated control variables and yields a maximum of four observations per country (one observation per country for 2013-2016). However, regression model (1) contains an intercept, five variables that are the same for every country for a given month (*BOG*, *EQUITY*, *TREAS*, *OIL*, and *DISCLOSE*), and is estimated using year fixed effects. Thus, when using annual data, we must suppress the intercept, omit *DISCLOSE*, and eliminate the year fixed effects (and instead cluster by country and year) or else there would be insufficient degrees of freedom to estimate the model. Panel C of Table 5 shows that the coefficients on *LD* and *BOG* remain negative and significant, although *BOG* is only marginally significant at the 10% level. The interaction term remains positive and significant. Overall, our full sample results are robust to the use of annual data.

Finally, we estimate regression model (1) augmented with several additional controls. Specifically, we add (1) the percentage of the population of country *i* speaking a language identified as having a weak future time reference in Chen (2013)[22], (2) an estimate of the total home bias of investors in country *i* from Cooper et al. (2018), (3) a measure of the quality of primary education in country *i* from the World Economic Forum, (4) the variance in exchange rates between country *i*'s currency and U.S. dollars, (5) a variable equal to |1| {2} (3) if the International Monetary Fund classifies country *i* as having |"Open"| {"Gate"} ("Wall") capital controls which limit foreign investment by their citizens, (6) the scaled number of immigrants

---

[22] Languages differ with respect to the delineation between the present and future tense. For example, "It rains tomorrow" is weak compared to the stronger "It will rain tomorrow". Chen (2013) show that individuals from weak future time reference countries save more, smoke less, practice safer sex, and are less obese. Kim et al. (2017) find that managers from weak future time reference countries engage in less earnings management.

from country *i* living in the U.S. as of 2010 from the Migration Policy Institute, (7) the percentage of adults in country *i* that are financially literate according to the S&P Global FinLit Survey, (8) the most recent monthly percentage change in the local stock index of country *i* using data from Thomson Reuters, and (9) the percentage of country *i* identifying as religious from the Pew Center. We do not include these variables in our primary regression specification in Table 4 because inclusion of these variables reduces our sample size by 53% (including the exclusion of economically important countries such as China) and introduces multicollinearity (e.g. education quality is highly correlated with GDP per capita). Nevertheless, the coefficients on *LD* and *BOG* remain negative and significant while the interaction coefficient remains positive and significant.

In untabulated analyses, we also re-estimate the four robustness tests in Table 5 in both pseudo changes and full changes for a total of eight additional specifications. For seven of those eight specifications, the results are quantitatively and qualitatively similar to the changes results reported in Table 4 (statistically significant negative coefficients on *LD* and Δ*BOG* and a significantly positive interaction coefficient). When using annual data in full changes, the sample size decreases by 83% but we continue to find a significantly negative *LD* coefficient and a significantly positive interaction coefficient (like all prior results). The coefficient on Δ*BOG* is not statistically significant, potentially due to a lack of power in the smaller sample.

We also perform additional untabulated robustness checks. First, we re-estimate model (1) excluding all offshore banking center countries. The coefficients on *LD* and *BOG* remain negative and significant while the interaction coefficient remains positive and significant. Thus, the full-sample results do not appear to be driven by the presence of offshore financial centers. Second, we estimate model (1) separately for G-20 countries and non-G-20 countries. The coefficients on *LD* and *BOG* are negative and significant in both sub-samples, and the interaction

coefficient is positive and significant in both sub-samples. This suggests that our results hold for the largest and most developed countries as well as the smaller and less developed countries. Finally, we re-estimate regression model (1) excluding one country at a time. The coefficients on *LD* and *BOG* are negative and significant and the interaction coefficient is positive and significant in all 75 specifications, suggesting no one country in our sample drives our results.

6. Additional Analyses

The results in Tables 4 and 5 are consistent with foreign investors reducing their holdings of U.S. stocks in response to linguistic distance and difficulty reading U.S. financial reports. A natural follow-up question is "Where do foreign investors then direct those funds?" Foreign investors' options in response to linguistic distance and readability frictions include, but are not limited to, (1) re-allocating their U.S. equity investments (e.g. divesting from U.S. firms with less readable financial reports and investing in U.S. firms with more readable financial reports), (2) investing in their home country or other foreign countries, (3) investing in alternative U.S. securities (e.g., U.S. Treasuries), and (4) consuming rather than investing. Data limitations prevent us from examining all these alternatives. For example, TIC data does not include firm-level information, and thus we cannot examine the first alternative (e.g. the *CS* variable would be unaffected if foreign investors sell some of one U.S. stock and buy an equal dollar amount of another). Additionally, we are not aware of data that details foreign investors' equity holdings outside of the U.S. (e.g. Japanese ownership of French stocks), and thus we cannot examine the second alternative. However, we can explore two alternatives for foreign investors. In our final analyses, we examine whether foreign investors direct a portion of their investment funds towards U.S. Treasuries and/or whether they increase their household consumption expenditures in response to linguistic distance and difficulty reading U.S. financial reports.

First, we examine foreign investors' holdings of U.S. Treasuries. See Table 1 for descriptive statistics on country i's holdings of U.S. Treasuries at the end of month j in millions of real 2010 U.S. dollars from TIC Table 1D ($US_{i,j}$). See Figure 1 for a time-series plot of aggregate foreign holdings of U.S. Treasuries. Table 6 Panel A presents regression model (1) estimated with US as the dependent variable (we present only the levels specification for brevity). The coefficients on LD and BOG are positive and significant, although BOG is only marginally significant at the 10% level. Additionally, the coefficient on the interaction term is negative and significant. These results are the mirror image of the U.S. stock results in Tables 4 and 5 and suggest a substitution effect. In other words, some foreign investors who want to invest in U.S. stocks, but are sufficiently deterred by linguistic distance and financial reporting readability frictions, appear to seek the relative ease of investing in U.S. Treasuries.

Next, we investigate annual total household consumption expenditures by country. See Table 1 for descriptive statistics on household consumption for country i in year t in millions of real 2010 U.S. dollars ($CONS_{i,t}$) from the World Bank.[23] Table 6 Panel B presents regression model (1) estimated with $CONS$ as the dependent variable. The coefficients on LD and BOG are positive and significant, and the coefficient on the interaction term is negative and significant. Like the U.S. Treasury results, these consumption results are the mirror image of the U.S. stock results in Tables 4 and 5 and suggest a substitution effect. That is, some foreigners who want to invest in U.S. stocks, but are sufficiently deterred by linguistic distance and financial reporting readability frictions, appear to simply invest less and consume more.

Overall, the results in Table 6 suggest that linguistic distance and financial reporting readability increase the demand for U.S. Treasuries. These results are the opposite of a home bias

---

[23] Consumption data from the World Bank is annual (as opposed to TIC data which is monthly), and consumption data is not available for three countries in our sample (China, Malta, and Trinidad and Tobago).

and are unique because investment frictions generally reduce demand. Moreover, these results suggest that financial reporting readability frictions can induce spillover effects in markets for assets other than stocks, including U.S. Treasuries that do not have accompanying financial statements. Furthermore, our results collectively suggest that linguistic distance and financial reporting readability can alter the relative attractiveness of U.S. stocks, U.S. Treasuries, and household consumption. In turn, changes in U.S. Treasury rates may alter the risk-free rate for all firms. Similarly, consumption-based asset pricing models suggest that changes in the marginal utilities of investment and consumption can alter firms' betas (Cochrane 2009). Thus, our results suggest that linguistic distance, financial reporting readability, and the interaction between the two have the potential to influence firms' cost of capital and overall macroeconomic growth.

7. Conclusion

We show that foreign investors invest less in U.S. stocks when they are from countries with greater linguistic distances and when U.S. financial reports are more difficult to read. Additionally, we find that linguistic distance and financial reporting readability are positively associated with foreign holdings of U.S. Treasuries and with foreign consumption. These results are consistent with a substitution effect. In other words, foreigners who want to hold U.S. stocks, but are sufficiently deterred by linguistic distance and financial reporting readability frictions, appear to (1) invest more in U.S. Treasuries which are less linguistically complex because they do not have financial statements and/or (2) simply invest less and consume more.

We make multiple contributions to the literature. First, our results suggest that linguistic distance and financial reporting readability persist as information processing frictions, even in the U.S. equities market where foreign investors should have the greatest ability, resources, and incentives to overcome these issues. Second, we introduce a continuous measure of linguistic

distance and allow the economic impact of linguistic distance to vary depending on the readability of U.S. financial reports. Third, we use U.S. Treasury data which mitigates deficiencies with data used in prior studies by including foreign holdings of both U.S. equities and U.S. Treasuries, covering all foreign investors, and accurately specifying foreign investors' nationalities. Fourth, we show that linguistic distance and financial reporting readability increase the demand for U.S. Treasuries, and this result is unique (investment frictions generally reduce demand) and the opposite of a home bias. Fifth, our results suggest that financial reporting readability frictions can induce spillover effects in markets for assets other than stocks, including U.S. Treasuries that do not have accompanying financial statements.

Finally, our study has implications for future research. A growing literature examines the relation between firm-level financial reporting outputs and macroeconomic outcomes (see Konchitchki and Patatoukas 2014; Crawley 2015; Gallo et al. 2016; Nallareddy and Ogneva 2017; Shivakumar and Urcan 2017). Future research may quantify the macroeconomic consequences of the substitution effect that we document. Specifically, researchers may examine whether linguistic distance, financial reporting readability, and the interaction between the two affect firms' cost of capital within a consumption-based asset pricing framework. For example, our results suggest that an exogenous decrease in the readability of U.S. financial reports may increase demand for U.S. Treasuries. The associated downward pressure on U.S. Treasury yields may reduce the risk-free rate of return on stocks. The shock may also alter the relative attractiveness of investment and consumption, thereby altering firms' betas. Estimating the net impact on firms' cost of capital and overall macroeconomic outcomes (e.g. gross domestic product and inflation) are interesting empirical questions yet to be resolved.

References

Aggarwal, R., L. Klapper, and P.D. Wysocki. 2005. Portfolio Preferences of Foreign Institutional Investors. *Journal of Banking and Finance* 29: 2919-2946.

Ahearne, A.G., W.L. Griever, and F.E. Warnock. 2004. Information Costs and Home Bias: An Analysis of US Holdings of Foreign Equities. *Journal of International Economics* 62: 313-336.

Anderson, J.E. 2011. The Gravity Model. *Annual Review of Economics* 3: 133-160.

Arbuckle, A.Q. 2017. The Navajo Code Talkers that Helped the U.S. Win WWII. *Mashable*, November 11, https://mashable.com/2017/11/28/navajo-code-talkers/#euauPWarxPqh, retrieved 1/15/2018.

Ardalan, K. 2019. Equity Home Bias: A Review Essay. *Journal of Economic Surveys* 33: 949-967.

Bae, K.-H., R.M. Stulz, and H. Tan. 2008a. Do Local Analysts Know More? A Cross-Country Study of the Performance of Local Analysts and Foreign Analysts. *Journal of Financial Economics* 88: 581-606.

Bae, K.-H., H. Tan, and M. Walker. 2008b. International GAAP Differences: The Impact on Foreign Analysts. *The Accounting Review* 83: 593-628.

Baik, B., J.-K. Kang, J.-M. Kim, and J. Lee. 2013. The Liability of Foreignness in International Equity Investments: Evidence from the US Stock Market. *Journal of International Business Studies* 44: 391-411.

Beugelsdijk, S., and B. Frijns. 2010. A Cultural Explanation of the Foreign Bias in International Asset Allocation. *Journal of Banking & Finance* 34: 2121-2131.

Bonsall IV, S.B., A.J. Leone, B.P. Miller, and K. Rennekamp. 2017. A Plain English Measure of Financial Reporting Readability. *Journal of Accounting and Economics* 63: 329-357.

Bonsall IV, S.B., and B.P. Miller. 2017. The Impact of Narrative Disclosure Readability on Bond Ratings and the Cost of Debt Capital. *Review of Accounting Studies* 22: 608-643.

Bozanic, Z., and M. Thevenot. 2015. Qualitative Disclosure and Changes in Sell-Side Financial Analysts Information Environment. *Contemporary Accounting Research* 32: 1595-1616.

Bradshaw, M.T., B.J. Bushee, and G.S. Miller. 2004. Accounting Choice, Home Bias, and U.S. Investment in Non-U.S. Firms. *Journal of Accounting Research* 42: 795-841.

Brandner, E., F. Cai, and R. Judson. 2012. Improving the Measurement of Cross-Border Securities Holdings: The Treasury International Capital SLT. *Federal Reserve Bulletin* 98: 1-28.

Brennan, M.J., and H.H. Cao. 1997. International Portfolio Investment Flows. *The Journal of Finance* 52: 1851 – 1880.

Bump, P. 2017. The Asterisk that Accompanies the Gains in the Stock Market: A Third of the Shares are Owned by Foreigners. *The Washington Post,* October 24, https://www.washingtonpost.com/news/politics/wp/2017/10/24/the-asterisk-that-accompanies-the-gains-in-the-stock-market-a-third-of-the-shares-are-owned-by-foreigners/?noredirect=on&utm_term=.b3f4f069f86d, retrieved 1/15/18.

Chan, K., V. Covrig, and L. Ng. 2005. What Determines the Domestic Bias and Foreign Bias? Evidence from Mutual Fund Equity Allocations Worldwide. *The Journal of Finance* 60: 1495-1534.

Chen, M.K. 2013. The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets. *American Economic Review* 103: 690-731.

Chiswick, B.R., and W Miller. 2005. Linguistic Distance: A Quantitative Measure of the Distance Between English and Other Languages. *Journal of Multilingual and Multicultural Development* 26: 1-11.

Choi, N., M. Fedenia, H. Skiba, and T. Sokolyk. 2017. Portfolio Concentration and Performance of Institutional Investors Worldwide. *Journal of Financial Economics* 123: 189-208.

Clark, W.S. 2008. Tax Effects on Foreign Direct Investment. *Organization for Economic Co-Operation and Development*.

Cochrane, R. 2009. *Asset Pricing.* Princeton: Princeton University Press.

Cooper, I.A., P. Secru, and R. Vanpee. 2018. A Measure of Pure Home Bias. *Review of Finance* 22: 1469-1524.

Coval, J.D., and T.J. Moskowitz. 1999. Home Bias at Home: Local Equity Preference in Domestic Portfolios. *The Journal of Finance* 54: 2045 – 2073.

Crawley, M.J. 2015. Macroeconomic Consequences of Accounting: The Effect of Accounting Conservatism on Macroeconomic Indicators and the Money Supply. *The Accounting Review* 90: 987-1011.

Demirbag, M. K.W. Glaister, and E. Tatoglu. 2007. Institutional and Transaction Cost Influences on MNEs Ownership Strategies of Their Affiliates: Evidence from and Emerging Market. *Journal of World Business* 42: 418-434.

Elliot, W.B., K.M. Rennekamp, and B.J. White. 2015. Does Concrete Language in Disclosures Increase Willingness to Invest? *Review of Accounting Studies* 20: 839-865.

Enrich, D. 2018. A Swiss Banker Helped Americans Dodge Taxes. Was It a Crime? *The New York Times* January 6, https://www.nytimes.com/2018/01/06/business/stefan-buck-tax-evasion.html, retrieved 1/15/2018.

Ferreira, M.A. and P. Matos. 2008. The Colors of Investors Money: The Role of Institutional Investors Around the World. *Journal of Financial Economics* 88: 499-533.

Firtel, K.B. 1999. Plain English: A Reappraisal of the Intended Audience of Disclosure Under the Securities Act of 1933. *Southern California Law Review* 851: 851-898.

Fredriksson, R., W. Barner-Rasumssen, and R. Piekkari. 2006. The Multinational Corporation as a Multilingual Organization: The Notion of a Common Corporate Language. *Corporate Communications: An International Journal* 11: 406-423.

French, K.R. and J.M. Poterba. 1991. Investor Diversification and International Equity Markets. *The American Economic Review* 81: 222-226.

Gallo, L.A., R.N. Hann, and C. Li. 2016. Aggregate Earnings Surprises, Monetary Policy, and Stock Returns. *Journal of Accounting and Economics* 62: 103-120.

Gandelman, N., and R. Hernandez-Murillo. 2014. Risk Aversion at the Country Level. Working Paper, Federal Reserve Bank of St. Louis, https://ssrn.com/abstract=2646134.

Ginsberg, V., and S. Weber. 2016. *The Palgrave Handbook of Economics and Language.* London, UK: Palgrave MacMillan.

Gorodnichenko, Y. and G. Roland. 2011. Culture, Institutions, and Development: Which Dimensions of Culture Matter for Long-Term Growth? *American Economic Review* 101: 492-498.

Grinblatt, M., and M. Keloharju. 2001. How Distance, Language, and Culture Influence Stockholdings and Trades. *The Journal of Finance* 56: 1053 – 1073.

Guay, W., D. Samuels, and D. Taylor. 2016. Guiding Through the Fog: Financial Statement Complexity and Voluntary Disclosure. *Journal of Accounting and Economics* 62: 234-269.

Guiso L., P. Sapienza, and L. Zingales. 2009. Cultural Biases in Economic Exchange? *The Quarterly Journal of Economics* 124: 1095-1131.

Gunning, R. 1952. *The Technique of Clear Writing.* New York: McGraw-Hill International Book Company.

Haldevang, M. 2017. Why We Can't Trust Basic Economic Figures. *Quartz,* November 28, https://qz.com/1133984/the-global-offshore-system-means-we-cant-trust-foreign-direct-investment-figures-and-other-basic-data/, retrieved 1/15/2018.

Hart-Gonzalez, L., and S. Lindemann. 1993. *Expected Achievement in Speaking Proficiency.* School of Language Studies, Foreign Services Institute, Department of State, mimeo (1993).

Henry, P.B. 2000a. Do Stock Market Liberalizations Cause Investment Booms? *Journal of Financial Economics* 58: 301-334.

Henry, P.B. 2000b. Stock Market Liberalization, Economic Reform, and Emerging Market Equity Prices. *The Journal of Finance* 55: 529-564.

Hofstede, G. 2001. Cultures consequences: International differences in work-related values. 2nd ed. Beverly Hills: Sage.

Isphording, I.E. 2014. Disadvantages of Linguistic Origin: Evidence from Immigrant Literacy Scores. *Economics Letters* 123: 236-239.

Isphording, I.E., and S. Otten. 2013. The Costs of Babylon – Linguistic Distance in Applied Economics. *Review of International Economics* 21: 354-369.

Kang, J.-K., and R. Stulz. 1997. Why is there a Home Bias? An Analysis of Foreign Portfolio Equity Ownership in Japan. *Journal of Financial Economics* 46: 3-28.

Karolyi, G.A., and R. Stulz. 2003. Are financial assets priced locally or globally? *Handbook of the Economics of Finance* 1: 975-1020.

Khurana, I.K., and P.N. Michas. 2011. Mandatory IFRS Adoption and the U.S. Home Bias. *Accounting Horizons* 25: 729-753.

Kim, J. Y. Kim, and J. Zhou. 2017. Languages and Earnings Management. *Journal of Accounting and Economics* 63: 288-306.

Konchitchki, Y., and P.N. Patatoukas. 2014. Accounting Earnings and Gross Domestic Product. *Journal of Accounting and Economics* 57: 76-88.

Kothari, S.P. 2001. Capital Markets Research in Accounting. *Journal of Accounting and Economics* 31: 105-231.

La Porta, R., F. Lopez-De-Silanes, A. Shleifer, and R. Vishny. 1998. Law and Finance. *Journal of Political Economy* 106: 1113-1155.

Lawrence, A. 2013. Individual Investors and Financial Disclosure. *Journal of Accounting and Economics* 56: 130-147.

Lehavy, R., F. Li, and K. Merkley. 2011. The Effect of Annual Report Readability on Analyst Following and the Properties of Their Earnings Forecasts. *The Accounting Review* 86: 1087-1115.

Li, F. 2008. Annual Report Readability, Current Earnings, and Earnings Persistence. *Journal of Accounting and Economics* 45: 221-247.

Lin, A.Y., and P.E. Swanson. 2003. The Behavior and Performance of Foreign Investors in Emerging Equity Markets: Evidence from Taiwan. *International Review of Finance* 4: 189-210.

Lizardo, R.A., and A.V. Mollick. 2009. Do Foreign Purchases of U.S. Stocks Help the U.S. Stock Market? *Journal of International Financial Markets, Institutions, and Money* 19: 969-986.

Lopez-Duarte, C., and M.M. Vidal-Suarez. 2010. External Uncertainty and Entry Mode Choice: Cultural Distance, Political Risk and Language Diversity. *International Business Review* 19: 575-588.

Loughran, T., and B. McDonald. 2014. Measuring Readability in Financial Disclosure. *The Journal of Finance* 69: 1643 – 1671.

Lundholm, R., N. Rahman, and R. Rogo. 2018. The Foreign Investor Bias and Its Linguistic Origins. *Management Science* 64: 4433-4450.

Lundholm, R.J., R. Rogo, and J.L. Zhang. 2014. Restoring the Tower of Babel: How Foreign Firms Communicate with U.S. investors. *The Accounting Review* 89: 1453-1485.

Miller, B.P. 2010. The Effects of Reporting Complexity on Small and Large Investor Trading. *The Accounting Review* 85: 2107-2143.

Nallareddy, S., and M. Ogneva. 2017. Predicting Restatements in Macroeconomics Indicators using Accounting Information. *The Accounting Review* 92: 151-182.

Neely, T. 2012. Global Business Speaks English. *Harvard Business Review*.

Nickell, S. 1981. Biases in Dynamic Models with Fixed Effects. *Econometrica* 49: 1417-1426.

Pacaud, J. 2017. One Man's Fight Against the Swiss Offshore Banking System. *The Economist* December 23, https://www.economist.com/christmas-specials/2017/12/23/one-mans-fight-against-the-swiss-offshore-banking-system, retrieved 1/15/2018.

Portes, R., H. Rey, and Y. Oh. 2001. Information and Capital Flows: The Determinants of Transactions in Financial Assets. *European Economic Review* 45: 783-796.

Portes, R., and H. Rey. 2005. The Determinants of Cross-Border Equity Flows. *Journal of International Economics* 65: 269-296.

Rennekamp, K. 2012. Processing Fluency and Investors Reactions to Disclosure Readability. *Journal of Accounting Research* 50: 1319-1354.

Sayek, S. 2009. Foreign Direct Investment and Inflation. *Southern Economic Journal* 76: 419-443.

Securities and Exchange Commission. 1998. *A Plain English Handbook: How to Create Clear SEC Disclosure Documents.* Washington, D.C.

Shivakumar, L., and O. Urcan. 2017. Why Does Aggregate Earnings Growth Reflect Information about Future Inflation? *The Accounting Review* 92: 247-276.

Stulz, R. 1999. Globalization, Corporate Finance, and the Cost of Capital. *Journal of Applied Corporate Finance* 12: 8-25.

Stulz, R., and W. Wasserfallen. 1995. Foreign Equity Investment Restrictions, Capital Flight, and Shareholder Wealth Maximization: Theory and Evidence. *The Review of Financial Studies* 8: 1019-1057.

Taylor, I. 1980. *The Korean writing system: An alphabet? A syllabary? A logography? Processing of Visible Language.* NATO Conference Series. Boston, MA: Springer.

Tietze, S. 2004. Spending the Management Gospel – in English, Language and Intercultural Communication. *Language and Intercultural Communications* 4: 175-189.

United States Treasury, Federal Reserve Bank of New York. 2016. *Foreign Portfolio Holdings of U.S. Securities*. New York, NY.

You, H., and X.-J. Zhang. 2009. Financial Reporting Complexity and Investor Underreaction to 10-k Information. *Review of Accounting Studies* 14: 559-585.

Appendix

Variable Definitions

|	
$CS_{i,j}$	Total holdings of U.S. stocks by investors in country i at the end of month j from U.S. Treasury International Capital Table 1D in real 2010 U.S. dollars.
$US_{i,j}$	Total holdings of U.S. Treasuries by investors in country i at the end of month j from U.S. Treasury International Capital Table 1D in real 2010 U.S. dollars.
$CONS_{i,t}$	Total household consumption expenditures in country i in year t from the World Bank in real 2010 U.S. dollars.
LD_i	Linguistic distance of country i based on Chiswick and Miller (2005). Higher values denote greater linguistic distance. See Section 3 for further discussion.
BOG_{j-1}	Aggregate U.S. financial reporting readability at the end of month j-1 based on Bonsall et al. (2017). Higher values denote less readability. See Section 3 for further discussion.
$EQUITY_{j-1}$	Monthly returns on the Center for Research in Securities Prices (CRSP) value-weighted index from month j-12 to month j-1.
$TREAS_{j-1}$	Monthly returns on 30-year U.S. Treasury bonds from month j-12 to month j-1 from the CRSP database.
OIL_{j-1}	Monthly spot price of Brent crude oil at the end of month j-1 from the Federal Reserve Economic Database.
$DISCLOSE_{j-1}$	The number of unique firms holding at least one conference call between month j-12 and j-1 (per Seeking Alpha) / the number of unique firms on Compustat in the calendar quarter containing month j-1.
$POP_{i,t-1}$	Population of country i (in millions) at the end of year t-1 (i.e. the most recent calendar year ending before month j) from the World Bank.
$BILAT_{i,t-1}$	Bilateral trade (i.e. the sum of imports and exports) between country i and the U.S. in year t-1 from the United Nations Commodity Trade Statistics Database in real 2010 U.S. dollars.
$GDPPC_{i,t-1}$	Gross domestic product per capita for country i in year t-1 from the World Bank in real 2010 U.S. dollars.
$\Delta GDPPC_{i,t-1}$	Percentage change in gross domestic product per capita for country i from year t-2 to t-1.
$INF_{i,t-1}$	Inflation for country i in year t-1 from the World Bank.
$G20_i$	Indicator variable equal to 1 if country i is a member of the G-20 and 0 otherwise.
TAX_i	Average corporate tax rate (including corporate income taxes, payroll taxes, retirement contributions, and other taxes) for country i from the World Economic Forum.
$OFFSHORE_i$	Indicator variable equal to 1 if the U.S. Treasury has identified country i as an "offshore banking center" and 0 otherwise.
$FINSTD_i$	The strength of the financial reporting and auditing standards in country i from the World Economic Forum. Values are integers from 1 to 7, and higher (lower) values denote stronger (weaker) standards.
$COMMON_i$	Indicator variable that equals 1 if country i is a former British colony or is listed as a common law country in Table 2 of LaPorta et al. (1998) and 0 otherwise.
$PROTECT_i$	The strength of investor protections in country i from the World Economic Forum. Values are integers from 1 to 7, and higher (lower) values denote stronger (weaker) investor protections.

Variable Definitions

$DIST_i$	Distance between the capital of country i and New York City in kilometers.
$POWERDISTANCE_i$	One of six measures of culture for country i based on Hofstede (2001). Higher values denote a greater deference to people in power.
$INDIVIDUALISM_i$	One of six measures of culture for country i based on Hofstede (2001). Higher (lower) values denote that the individual (group) is the most important social unit.
$MASCULINITY_i$	One of six measures of culture for country i based on Hofstede (2001). Higher values denote emphasis on ambition, wealth acquisition, and differentiated gender roles whereas lower values denote emphasis on nurturing behaviors, sexual equality, environmental awareness, and more fluid gender roles.
$UNCERTAINTYAVOIDANCE_i$	One of six measures of culture for country i based on Hofstede (2001). Higher values denote a culture that feels more threatened by ambiguity.
$LONGTERM_i$	One of six measures of culture for country i based on Hofstede (2001). Higher values denote a culture more prepared to delay short-term gratification.
$INDULGENCE_i$	One of six measures of culture for country i based on Hofstede (2001). Higher values denote a culture more prepared to allow relatively free gratification of natural human desires.
$WEAKFTR_i$	The percentage of the population of country i speaking a language identified as having a weak future time reference in Chen (2013) Appendix B Table 1.
$HOMEBIAS_i$	An estimate of the total home bias of investors in country i from Cooper et al. (2017) Table 3.
$EDUC_i$	A measure of the quality of primary education in country i ranging from 1 (low) to 7 (high) from the World Economic Forum.
$ERVAR_{i,t-1}$	The variance in the weekly exchange rate between the local currency in country i and U.S. dollars over year t-1. Weekly exchange rates are from OANDA.
$CAPCONTROL_i$	A numeric variable equal to 1 (2) [3] if the International Monetary Fund classifies country i as having "Open" ("Gate") ["Wall"] capital controls in place which limit foreign investment by their citizens.
$IMMIGRANT_{i,t-1}$	The number of immigrants from country i living in the U.S. as of 2010 from the Migration Policy Institute, scaled by the population of country i in year t-1.
$FINLIT_i$	The percentage of adults in country i that are financially literate according to the S&P Global FinLit Survey.
$HOMERET_{i,j-1}$	The percentage change in the local stock index of country i from month j-2 to j-1. Stock index data are from Thomson Reuters.
$RELIGIOUS_i$	The percentage of country i identifying as religious from the Pew Center.

Figure 1: Foreign Holdings of U.S. Securities and Financial Reporting Readability
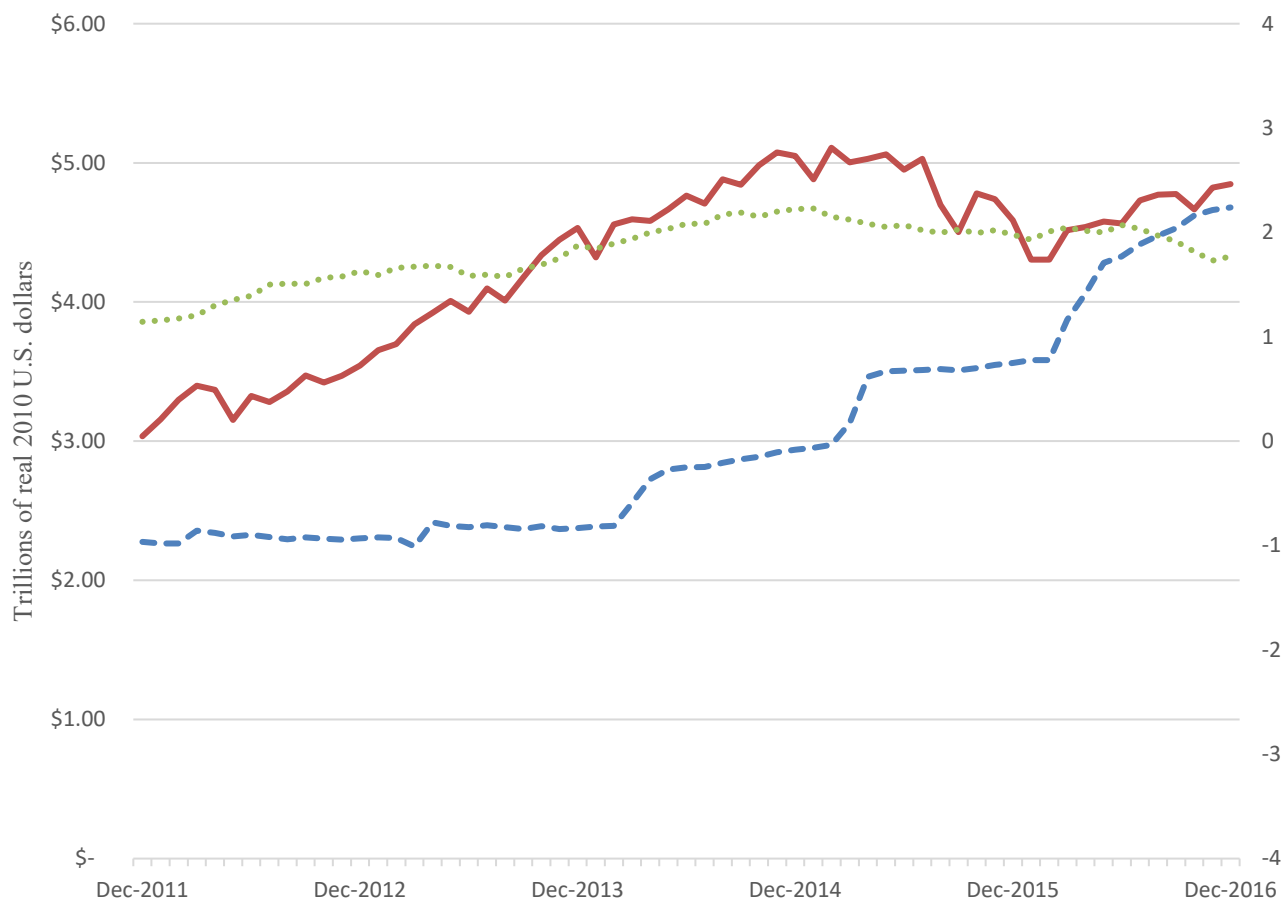
This figure plots the time series of three monthly data series. The sample period begins in December 2012 and ends in December 2016. The solid (dotted) line denotes total foreign holdings of U.S. stocks (U.S. Treasuries) from U.S. Treasury International Capital Table 1D in millions of real 2010 U.S. dollars. The dashed line denotes an aggregate measure of financial reporting readability based on Bonsall et al.'s (2017) plain English "Bog Index" for firm-level 10-K filings (higher values denote less readability). See Section 3 for further discussion.

Table 1

Linguistic Distance, Holdings of U.S. Securities, and Consumption by Country

Country	LD	CS	US	CONS	Country	LD	CS	US	CONS
Japan	3.00	348,413	1,025,088	3,400,545	Uruguay	1.75	1,901	6,402	32,553
Korea, South	3.00	52,804	51,591	605,519	Costa Rica	1.75	725	1,472	30,010
Hong Kong	2.75	65,555	106,062	174,385	Ecuador	1.75	460	101	49,896
China, Mainland	2.58	249,541	1,151,075	N/A	Russia	1.75	428	100,052	896,625
Kuwait	2.50	119,362	25,070	41,557	Dominican Republic	1.75	288	863	46,284
Saudi Arabia	2.50	53,798	84,206	216,556	Guatemala	1.75	265	2,414	42,124
Malta	2.50	576	294	N/A	El Salvador	1.75	200	536	17,733
Lebanon	2.50	520	1	37,883	Latvia	1.75	165	465	17,390
Egypt	2.50	350	2,544	200,516	Honduras	1.75	71	754	14,521
Morocco	2.50	79	1,275	63,715	Ukraine	1.75	54	3,946	101,759
Vietnam	2.50	31	11,254	96,432	Lithuania	1.75	42	6	27,140
Pakistan	2.31	133	918	162,510	Switzerland	1.66	300,365	164,640	339,503
Greece	2.25	3,595	1,045	168,400	Luxembourg	1.63	437,787	121,678	18,955
Sri Lanka	2.25	32	621	49,406	France	1.50	131,213	47,744	1,500,333
Israel	2.05	33,353	17,871	156,699	Italy	1.50	22,709	30,021	1,233,117
Finland	2.00	17,042	5,831	138,050	Brazil	1.50	5,701	218,930	1,476,452
Thailand	2.00	2,629	15,126	199,927	Portugal	1.50	2,634	3,404	145,788
Poland	2.00	2,371	28,443	320,414	Belgium	1.35	29,767	180,608	258,740
Czech Republic	2.00	1,819	5,599	101,913	Netherlands	1.25	181,628	31,819	372,502
Philippines	2.00	1,657	35,683	184,886	Singapore	1.25	127,982	91,467	99,287
Slovenia	2.00	902	190	25,475	United Arab Emirates	1.25	82,318	43,609	125,900
Hungary	2.00	643	746	71,156	Malaysia	1.25	9,720	10,643	166,774
Indonesia	2.00	333	18,212	532,971	Iceland	1.25	637	767	7,339
Turkey	2.00	332	19,303	641,396	Tanzania	1.25	13	1,405	27,538
Croatia	2.00	201	679	33,714	South Africa	1.13	6,819	10,026	248,733
Bulgaria	2.00	65	15	33,555	India	1.13	1,822	76,591	1,244,632
Slovakia	2.00	50	295	51,223	Norway	1.00	165,887	66,552	201,802
Serbia	2.00	15	1,243	30,150	Sweden	1.00	99,409	30,211	245,445
Germany	1.75	127,822	61,970	2,021,083	Romania	1.00	53	3,951	118,338
Denmark	1.75	53,777	13,838	157,936	Kenya	0.63	48	720	40,049
Mexico	1.75	35,117	49,744	773,949	Canada	0.33	649,310	48,348	1,016,420
Chile	1.75	23,230	19,612	158,083	United Kingdom	0.00	667,998	143,017	1,699,629
Colombia	1.75	10,707	30,185	216,243	Ireland	0.00	189,036	143,690	106,663
Panama	1.75	9,866	499	21,479	Australia	0.00	153,122	21,109	725,044
Spain	1.75	9,855	25,673	773,741	New Zealand	0.00	10,960	3,630	98,351
Peru	1.75	8,425	11,837	116,921	Trinidad and Tobago	0.00	2,045	3,042	N/A
Austria	1.75	7,990	3,876	215,174	Jamaica	0.00	147	73	10,933
Argentina	1.75	4,398	974	306,107					

This table presents a measure of linguistic distance by country (*LD*), the country's average holdings of U.S. stocks (*CS*), the country's average holdings of U.S. Treasuries (*US*), and the country's average household consumption expenditures (*CONS*). Linguistic distance captures how different one language is from another. We use a measure of linguistic distance between English and 43 other languages created by Chiswick and Miller (2005). Higher scores denote greater linguistic distance. We map from languages to countries by weighting the linguistic distance of the languages spoken within a country by the percentage of the population speaking each language from the Central Intelligence Agency Factbook. U.S. stock and U.S. Treasury data are from U.S. Treasury International Capital Table 1D. Consumption data are from the World Bank. All dollar amounts are in millions of real 2010 U.S. dollars. The sample is from December 2012 to December 2016. Countries are presented in descending order of linguistic distance.

Table 2

Descriptive Statistics

Variable	N	Mean	Standard Deviation	Q1	Median	Q3
CS	3,582	63,224.07	134,075.34	305.07	4,138.70	58,308.56
US	3,582	61,968.95	181,419.51	1,110.58	11,873.85	44,631.33
CONS	3,482	360,612.21	573,677.57	43,437.13	152,950.97	324,460.25
LD	3,582	1.65	0.70	1.25	1.75	2.00
BOG	3,582	0.16	1.01	-0.82	-0.15	0.73
LD*BOG	3,582	0.26	1.81	-1.22	-0.11	1.31
EQUITY	3,582	12.45	9.68	4.86	14.18	19.82
TREAS	3,582	5.19	12.76	-5.33	4.94	14.42
OIL	3,582	79.84	29.99	47.76	97.09	108.90
DISCLOSE	3,582	39.43	5.17	38.94	41.77	42.75
POP	3,582	72.58	217.30	5.60	16.14	52.23
G20	3,582	0.25	0.43	0.00	0.00	1.00
BILAT	3,582	47,274.18	108,061.18	3,601.39	13,695.13	39,358.02
GDPPC	3,582	23,149.09	22,281.34	5,670.71	13,681.00	39,226.35
INF	3,582	3.10	5.47	0.85	1.97	4.36
TAX	3,582	40.87	14.90	30.90	39.70	49.00
OFFSHORE	3,582	0.06	0.23	0.00	0.00	0.00
FINSTD	3,582	5.07	0.78	4.50	5.00	5.70
COMMON	3,582	0.23	0.42	0.00	0.00	0.00
PROTECT	3,582	6.07	1.16	5.50	6.00	7.00
DIST	3,582	8,018.72	3,685.13	5,850.00	6,924.00	10,873.00
POWERDISTANCE	3,582	61.14	21.96	44.00	65.00	77.00
INDIVIDUALISM	3,582	42.02	22.79	25.00	36.00	63.00
MASCULINITY	3,582	48.02	19.07	40.00	48.00	63.00
UNCERTAINTYAVOIDANCE	3,582	66.54	23.33	49.00	70.00	86.00
LONGTERM	3,582	40.58	25.50	21.00	38.00	60.00
INDULGENCE	3,582	40.13	25.87	20.00	42.00	62.00

This table presents descriptive statistics for the levels variables used in the regressions (changes variables are omitted for brevity). The sample period begins in December 2012 and ends in December 2016. All variables are defined in the Appendix.

Table 3

Univariate Correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
(1) CS		0.41	0.46	-0.28	0.04	0.01	-0.03	0.03	-0.04	0.05	0.08	0.33	0.53	0.60	-0.19	-0.21	0.54	0.47	0.24	0.18	-0.13	-0.31	0.43	0.18	-0.24	0.24	0.27
(2) US	0.75		0.72	0.21	0.02	0.03	-0.02	0.01	-0.02	0.02	0.56	0.36	0.55	0.14	-0.09	0.19	0.12	0.10	-0.06	-0.06	0.14	0.03	0.01	0.32	-0.08	0.39	0.02
(3) CONS	0.55	0.72		-0.02	0.03	0.03	-0.03	0.02	-0.04	0.02	0.34	0.72	0.48	0.21	-0.05	0.27	0.09	0.16	0.01	0.20	0.02	-0.12	0.30	0.35	0.07	0.41	0.19
(4) LD	-0.22	-0.15	-0.12		-0.01	0.06	0.01	-0.01	0.02	-0.01	0.09	-0.03	-0.01	-0.25	0.01	0.08	-0.17	-0.29	-0.49	-0.26	0.16	0.35	-0.39	-0.03	0.37	0.25	-0.31
(5) BOG	0.06	0.08	0.07	-0.02		0.92	-0.72	0.36	-0.88	0.61	0.01	0.02	0.01	0.03	-0.11	0.01	0.01	0.01	0.01	-0.02	0.01	0.00	0.00	0.01	0.00	-0.03	0.02
(6) LD*BOG	0.08	0.09	0.07	-0.10	0.91		-0.66	0.34	-0.81	0.56	0.02	0.02	0.02	0.00	-0.12	0.01	0.00	-0.01	-0.03	-0.04	0.02	0.02	-0.03	0.01	0.02	-0.02	0.00
(7) EQUITY	-0.05	-0.08	-0.06	0.02	-0.75	-0.75		-0.34	0.85	-0.31	-0.01	-0.02	-0.01	-0.02	0.06	-0.01	-0.01	-0.01	-0.01	0.02	-0.01	0.00	0.00	-0.01	0.00	0.03	-0.02
(8) TREAS	0.03	0.04	0.03	-0.01	0.41	0.40	-0.41		-0.40	0.13	0.01	0.01	0.01	0.01	-0.01	0.00	0.01	0.01	0.00	-0.01	0.00	0.00	0.00	0.01	0.00	-0.02	0.01
(9) OIL	-0.05	-0.08	-0.06	0.02	-0.90	-0.83	0.79	-0.32		-0.62	-0.01	-0.03	-0.02	-0.03	0.09	-0.01	-0.01	-0.01	-0.01	0.02	-0.01	0.00	0.00	-0.01	0.00	0.04	-0.02
(10) DISCLOSE	0.06	0.07	0.06	-0.02	0.85	0.78	-0.63	0.42	-0.78		0.01	0.02	0.01	0.02	-0.12	0.01	0.01	0.01	0.01	-0.01	0.01	0.00	0.00	0.01	0.00	-0.02	0.01
(11) POP	0.02	0.41	0.72	0.13	0.04	0.03	-0.04	0.02	-0.04	0.04		0.41	0.45	-0.20	0.05	0.31	-0.06	-0.17	0.07	-0.03	0.23	0.18	-0.10	0.15	-0.21	0.24	-0.09
(12) G20	0.34	0.48	0.71	-0.04	0.02	0.03	-0.02	0.01	-0.02	0.02	0.63		0.52	0.02	0.12	0.30	0.00	-0.02	-0.01	0.17	0.15	0.02	0.20	0.29	0.01	0.33	0.20
(13) BILAT	0.68	0.78	0.78	-0.09	0.03	0.05	-0.04	0.02	-0.03	0.03	0.55	0.62		0.08	-0.08	0.10	0.01	0.13	0.07	0.08	-0.11	0.00	0.08	0.25	-0.13	0.22	0.21
(14) GDPPC	0.78	0.48	0.33	-0.25	0.03	0.05	-0.02	0.01	-0.02	0.02	-0.36	0.09	0.29		-0.30	-0.25	0.47	0.70	0.10	0.31	-0.10	-0.59	0.62	-0.05	-0.18	0.31	0.38
(15) INF	-0.42	-0.21	-0.14	0.02	-0.17	-0.16	0.10	-0.05	0.15	-0.15	0.19	0.09	-0.17	-0.49		0.44	-0.09	-0.30	-0.07	-0.12	0.04	0.06	-0.11	-0.05	0.07	-0.09	-0.10
(16) TAX	-0.17	0.00	0.28	0.04	0.01	0.00	-0.01	0.00	-0.01	0.01	0.42	0.23	0.10	-0.22	0.19		-0.10	-0.27	-0.27	-0.16	-0.09	0.01	0.07	0.11	0.21	0.11	0.11
(17) OFFSHORE	0.33	0.35	0.11	-0.22	0.01	0.03	-0.01	0.00	-0.01	0.01	-0.08	0.00	0.15	0.32	-0.16	-0.13		0.31	0.01	-0.06	-0.14	-0.20	0.33	0.15	-0.02	0.26	0.21
(18) FINSTD	0.65	0.39	0.19	-0.31	0.01	0.04	-0.01	0.00	-0.01	0.01	-0.27	-0.03	0.35	0.65	-0.28	-0.24	0.30		0.22	0.32	-0.03	-0.47	0.49	0.03	-0.31	0.11	0.32
(19) COMMON	0.15	0.07	-0.06	-0.35	0.01	0.05	-0.01	0.00	-0.01	0.01	-0.16	-0.01	0.07	0.09	0.00	-0.29	0.01	0.21		0.35	0.25	-0.17	0.12	0.19	-0.44	-0.16	0.02
(20) PROTECT	0.38	0.31	0.38	-0.23	-0.02	0.00	0.02	-0.01	0.02	-0.02	0.01	0.18	0.29	0.44	-0.20	-0.19	-0.08	0.32	0.33		0.20	-0.27	0.32	-0.05	-0.24	0.28	0.27
(21) DIST	-0.05	0.13	0.13	0.28	0.00	-0.02	0.00	0.00	0.00	0.00	0.30	0.16	0.04	-0.13	0.10	-0.08	-0.17	-0.11	0.21	0.17		0.20	-0.10	0.05	-0.29	0.24	-0.11
(22) POWERDISTANCE	-0.41	-0.10	-0.17	0.30	0.00	-0.03	0.00	0.00	0.00	0.00	0.29	0.02	-0.05	-0.60	0.15	0.05	-0.21	-0.44	-0.15	-0.29	0.23		-0.67	0.05	0.22	-0.13	-0.39
(23) INDIVIDUALISM	0.38	0.23	0.34	-0.33	0.00	0.03	0.00	0.00	0.00	0.00	-0.10	0.20	0.08	0.62	-0.21	0.12	0.30	0.40	0.09	0.29	-0.08	-0.61		0.08	-0.19	0.34	0.39
(24) MASCULINITY	0.23	0.20	0.29	-0.04	0.02	0.03	-0.02	0.01	-0.02	0.01	0.22	0.30	0.31	0.05	-0.10	0.10	0.20	0.08	0.22	-0.03	0.02	-0.03	0.18		-0.03	0.04	0.08
(25) UNCERTAINTYAVOIDANCE	-0.16	-0.14	-0.03	0.27	0.00	-0.03	0.00	0.00	0.00	0.00	-0.01	-0.02	-0.20	-0.04	0.03	0.23	-0.02	-0.28	-0.40	-0.20	-0.20	0.17	-0.18	-0.18		-0.01	-0.09
(26) LONGTERM	0.21	0.40	0.37	0.17	-0.03	-0.06	0.03	-0.02	0.03	-0.03	0.20	0.29	0.25	0.36	-0.20	0.14	0.27	0.10	-0.15	0.25	0.27	-0.13	0.36	0.01	-0.01		0.15
(27) INDULGENCE	0.46	0.32	0.36	-0.36	0.02	0.06	-0.02	0.01	-0.02	0.02	-0.02	0.19	0.26	0.47	-0.24	0.03	0.23	0.36	0.05	0.29	-0.16	-0.46	0.39	0.12	-0.11	0.09	

This table reports univariate correlations for the levels variables used in the regressions (changes variables are omitted for brevity). Pearson (Spearman) correlations are presented above (below) the diagonal. The sample period begins in December 2012 and ends in December 2016. Correlations significant at the 5% level are in bold. All variables are defined in the Appendix.

Table 4

Linguistic Distance, Financial Reporting Readability, and Foreign Holdings of U.S. Stocks

	Predicted Sign	Panel A Dependent Variable = $CS_{i,j}$		Panel B Dependent Variable = $(CS_{i,j} - CS_{i,j-12})$		Panel C Dependent Variable = $(CS_{i,j} - CS_{i,j-12})/GDP_{i,t-1}$	
		Coef.	*t-stat.*	*Coef.*	*t-stat.*	*Coef.*	*t-stat.*
Intercept	?	-22,434.72	(-2.40) **	-8,204.68	(-1.40)	11.30	(4.63) ***
LD_i	?	-894.46	(-2.09) **	-895.48	(-2.07) **	-1.43	(-4.33) ***
BOG_{j-1}	?	-9,022.46	(-5.23) ***	-8,980.36	(-6.39) ***		
ΔBOG_{j-1}	?					-1.69	(-1.90) *
LD_i*BOG_{j-1}	?	5,218.70	(7.32) ***	5,207.09	(6.99) ***		
$LD_i*\Delta BOG_{j-1}$	?					1.09	(3.58) ***
$CS_{i,j-12}$	+	1.00	(76.84) ***				
$EQUITY_{j-1}$	+	284.11	(3.90) ***	337.83	(4.71) ***	0.12	(3.05) ***
$TREAS_{j-1}$	?	179.67	(5.43) ***	144.02	(5.05) ***	0.03	(1.14)
OIL_{j-1}	+	209.59	(3.43) ***	132.68	(3.71) ***		
ΔOIL_{j-1}	+					0.02	(1.55)
$DISCLOSE_{j-1}$	+	465.09	(3.63) ***	245.08	(3.14) ***		
$\Delta DISCLOSE_{j-1}$	+					0.07	(0.96)
$POP_{i,t-1}$	?	-9.64	(-2.91) ***	-9.57	(-2.85) ***		
$\Delta POP_{i,t-1}$	?					1.50	(5.48) ***
$BILAT_{i,t-1}$	+	0.02	(2.23) **	0.02	(2.30) **		
$\Delta BILAT_{i,t-1}$	+					0.03	(2.01) **
$GDPPC_{i,t-1}$	?	0.20	(3.35) ***	0.20	(4.83) ***		
$\Delta GDPPC_{i,t-1}$	?	209.05	(2.31) **	194.62	(2.38) **	-0.07	(-0.79)
$INF_{i,t-1}$	+	320.87	(5.97) ***	329.27	(6.25) ***	0.16	(5.85) ***
$G20_i$	+	5,463.93	(4.56) ***	5,474.04	(3.93) ***	-0.94	(-2.61) ***
TAX_i	-	-158.29	(-6.63) ***	-160.43	(-6.26) ***	-0.13	(-6.32) ***
$OFFSHORE_i$	+	18,423.23	(5.87) ***	18,529.70	(4.55) ***	17.00	(5.11) ***
$FINSTD_i$	-	-1,307.82	(-3.41) ***	-1,291.64	(-3.40) ***	0.42	(2.65) ***
$COMMON_i$	+	4,500.46	(5.76) ***	4,513.56	(5.10) ***	-1.11	(-2.30) **
$PROTECT_i$	-	-11.81	(-0.03)	-13.74	(-0.04)	-0.90	(-3.90) ***
$DIST_i$	-	-0.40	(-4.05) ***	-0.40	(-3.62) ***	0.00	(-4.25) ***
$POWERDISTANCE_i$	?	-25.37	(-1.67) *	-25.06	(-1.71) *	-0.07	(-5.27) ***
$INDIVIDUALISM_i$	?	22.30	(1.10)	22.61	(1.02)	-0.04	(-3.64) ***
$MASCULINITY_i$	?	8.92	(0.64)	9.06	(0.61)	-0.01	(-2.25) **
$UNCERTAINTYAVOIDANCE_i$	?	-42.22	(-3.88) ***	-42.43	(-3.42) ***	0.01	(2.14) **
$LONGTERM_i$	?	-20.06	(-1.91) *	-20.22	(-1.78) *	0.06	(4.71) ***
$INDULGENCE_i$	?	-12.32	(-1.35)	-12.30	(-1.33)	0.02	(4.90) ***
N		3,582		3,582		3,582	
Adjusted R^2		0.981		0.284		0.159	
Year Fixed Effects		Yes		No		No	

This table reports the results from cross-sectional regressions of $CS_{i,j}$ (total holdings of U.S. stocks by investors in country i at the end of month j in millions of real 2010 U.S. dollars) on a measure of country i's linguistic distance, an aggregate measure of U.S. financial reporting readability at month j-1, and a vector of controls. All variables are defined in the Appendix. The sample period begins in December 2012 and ends in December 2016. t-statistics in parentheses are based on robust standard errors clustered by country. ***, **, and * represent two-tailed significance at the 1%, 5%, and 10% levels, respectively.

Table 5

Robustness Checks

	Panel A		Panel B		Panel C		Panel D	
	Alternate Linguistic Distance Measure		Exclude English Speaking Countries		Annual Data		Additional Control Variables	
	Coef.	*t-stat.*	*Coef.*	*t-stat.*	*Coef.*	*t-stat.*	*Coef.*	*t-stat.*
Intercept	-20,702.48	(-2.09) **	-9,941.12	(-0.95)			135,113.16	(5.97) ***
LD_i	-49.11	(-3.40) ***	-2,139.29	(-2.19) **	-2,304.79	(-2.00) **	-11,365.42	(-6.52) ***
BOG_{j-1}	-13,855.92	(-6.71) ***	-11,167.83	(-4.74) ***	-5,794.10	(-1.97) *	-19,222.78	(-7.08) ***
LD_i*BOG_{j-1}	168.26	(8.84) ***	7,222.91	(5.35) ***	4,934.40	(2.37) **	12,198.37	(11.36) ***
$CS_{i\,j-12}$	1.00	(79.79) ***	0.98	(69.86) ***	1.01	(14.29) ***	0.94	(51.02) ***
$EQUITY_{j-1}$	301.46	(3.96) ***	239.36	(3.13) ***	856.29	(2.18) **	384.45	(3.57) ***
$TREAS_{j-1}$	189.93	(5.50) ***	162.05	(4.88) ***	214.45	(3.87) ***	247.10	(4.80) ***
OIL_{j-1}	222.28	(3.50) ***	191.54	(3.16) ***	80.61	(0.56)	254.14	(2.65) ***
$DISCLOSE_{j-1}$	489.77	(3.66) ***	394.01	(2.93) ***			648.10	(3.37) ***
$POP_{i,t-1}$	-10.69	(-3.33) ***	-29.49	(-3.74) ***	-8.73	(-1.57)	6.89	(2.85) ***
$BILAT_{i,t-1}$	0.02	(2.09) **	0.05	(7.28) ***	0.01	(0.17)	0.05	(2.57) **
$GDPPC_{i,t-1}$	0.20	(3.22) ***	0.40	(6.23) ***	0.18	(1.37)	0.03	(0.38)
$\Delta GDPPC_{i,t-1}$	321.85	(3.17) ***	392.91	(3.40) ***	49.64	(0.31)	536.86	(2.19) **
$INF_{i,t-1}$	273.13	(5.17) ***	502.72	(11.28) ***	249.11	(2.64) **	-743.83	(-4.76) ***
$G20_i$	5,668.98	(4.39) ***	291.17	(0.30)	6,605.19	(6.75) ***	15,297.58	(6.73) ***
TAX_i	-139.08	(-5.68) ***	-90.94	(-4.01) ***	-140.17	(-1.79)	-141.95	(-3.07) ***
$OFFSHORE_i$	17,460.86	(5.66) ***	4,351.39	(2.36) **	16,919.78	(1.41)	36,455.92	(10.39) ***
$FINSTD_i$	-1,586.77	(-3.53) ***	-1,544.47	(-3.20) ***	-897.42	(-0.77)	-5,368.35	(-4.32) ***
$COMMON_i$	4,602.79	(5.00) ***	7,193.32	(6.82) ***	4,515.48	(3.23) ***	-12,893.28	(-4.86) ***
$PROTECT_i$	-279.63	(-0.65)	-1,928.49	(-5.42) ***	-198.65	(-0.31)	1,553.15	(1.75) *
$DIST_i$	-0.36	(-3.22) ***	-0.14	(-1.65) *	-0.35	(-2.28) **	-3.09	(-7.19) ***
$POWERDISTANCE_i$	-20.58	(-1.27)	-20.64	(-1.09)	-32.83	(-0.91)	-237.63	(-7.15) ***
$INDIVIDUALISM_i$	18.21	(0.80)	-126.89	(-6.44) ***	11.63	(0.30)	-786.15	(-9.79) ***
$MASCULINITY_i$	10.95	(0.81)	43.83	(3.96) ***	18.53	(0.57)	381.45	(8.47) ***
$UNCERTAINTYAVOIDANCE_i$	-39.35	(-3.52) ***	-59.73	(-5.93) ***	-48.59	(-1.71)	-294.41	(-7.77) ***
$LONGTERM_i$	-17.41	(-1.40)	79.91	(7.38) ***	-9.81	(-0.51)	-784.95	(-9.07) ***
$INDULGENCE_i$	-16.02	(-1.66) *	-31.70	(-3.56) ***	-2.32	(-0.09)	-121.86	(-3.59) ***
$WEAKFTR_i$							-7,770.27	(-4.51) ***
$HOMEBIAS_i$							27,999.75	(4.97) ***
$EDUC_i$							8,391.65	(7.17) ***
$ERVAR_{i,t-1}$							-0.01	(-3.70) ***
$CAPCONTROL_i$							3,379.40	(3.29) ***
$IMMIGRANT_{i,t-1}$							-581,694.62	(-5.47) ***
$FINLIT_i$							727.93	(8.94) ***
$HOMERET_{i,j-1}$							146.89	(2.12) **
$RELIGIOUS_i$							-1,060.28	(-9.60) ***
N	3,382		2,856		286		1,700	
Adjusted R^2	0.981		0.980		0.984		0.986	
Year Fixed Effects	Yes		Yes		No		Yes	

This table reports the results from cross-sectional regressions of $CS_{i,j}$ (total holdings of U.S. stocks by investors in country i at the end of month j in millions of real 2010 U.S. dollars) on a measure of country i's linguistic distance, an aggregate measure of U.S. financial reporting readability at month j -1, and a vector of controls. All variables are defined in the Appendix. The sample period begins in December 2012 and ends in December 2016. t -statistics in parentheses are based on robust standard errors clustered by country. ***, **, and * represent two-tailed significance at the 1%, 5%, and 10% levels, respectively.

Table 6

Linguistic Distance, Financial Reporting Readability, Foreign Holdings of U.S. Treasuries, and Consumption

	Panel A			Panel B	
	Dependent Variable = $US_{i,j}$			Dependent Variable = $CONS_{i,t}$	
	Coef.	*t-stat.*		*Coef.*	*t-stat.*
Intercept	-3,262.04	(-0.30)		-9,581.39	(-1.13)
LD_i	1,703.53	(3.23) ***		2,663.41	(2.16) **
BOG_{j-1}	1,645.41	(1.87) *		3,605.70	(1.75) *
LD_i*BOG_{j-1}	-1,977.95	(-3.86) ***		-1,897.62	(-2.44) **
$US_{i,j-12}$	0.98	(152.20) ***			
$CONS_{i,t-1}$				1.01	(212.45) ***
$EQUITY_{j-1}$	-38.11	(-0.49)		158.29	(0.95)
$TREAS_{j-1}$	112.79	(3.16) ***		-18.38	(-0.34)
OIL_{j-1}	67.86	(1.10)		-46.30	(-0.47)
$DISCLOSE_{j-1}$	-25.87	(-0.20)			
$POP_{i,t-1}$	11.00	(3.32) ***		48.25	(10.66) ***
$BILAT_{i,t-1}$	0.00	(-0.49)		0.01	(1.06)
$GDPPC_{i,t-1}$	0.17	(5.13) ***		-0.08	(-2.02) **
$\Delta GDPPC_{i,t-1}$	582.24	(6.95) ***		1,095.52	(4.14) ***
$INF_{i,t-1}$	110.12	(2.11) **		65.48	(0.33)
$G20_i$	1,417.86	(0.95)		5,047.00	(1.24)
TAX_i	-85.25	(-2.22) **		-101.67	(-1.87) *
$OFFSHORE_i$	3,405.46	(0.65)		7,853.84	(2.76) ***
$FINSTD_i$	-1,374.89	(-2.80) ***		1,863.17	(1.77) *
$COMMON_i$	1,718.38	(2.53) **		4,042.16	(2.00) **
$PROTECT_i$	448.44	(1.81) *		284.63	(0.55)
$DIST_i$	-0.26	(-3.44) ***		-0.07	(-0.28)
$POWERDISTANCE_i$	-36.37	(-1.73) *		-19.48	(-0.40)
$INDIVIDUALISM_i$	-32.95	(-1.56)		31.83	(1.21)
$MASCULINITY_i$	102.88	(6.21) ***		-25.96	(-0.83)
$UNCERTAINTYAVOIDANCE_i$	-32.12	(-1.86) *		-4.80	(-0.22)
$LONGTERM_i$	-19.25	(-1.74) *		-48.92	(-1.00)
$INDULGENCE_i$	6.11	(0.94)		66.17	(2.71) ***
N	3,582			336	
Adjusted R^2	0.987			0.980	
Year Fixed Effects	Yes			No	

This table reports the results from cross-sectional regression of $US_{i,j}$ (total holdings of U.S. Treasuries by investors in country i at the end of month j in millions of real 2010 U.S. dollars) and $CONS_{i,t}$ (total household consumption expenditures in country i in year t in millions of real 2010 U.S. dollars) on a measure of country i's linguistic distance, an aggregate measure of U.S. financial reporting readability at month j-1, and a vector of controls. All variables are defined in the Appendix. The sample period begins in December 2012 and ends in December 2016. t-statistics in parentheses are based on robust standard errors clustered by country. ***, **, and * represent two-tailed significance at the 1%, 5%, and 10% levels, respectively.