Article:
Chi, JD, Su, X, Tang, Y et al. (1 more author) (2020) Is Language an Economic Institution? Evidence from R&D Investment. Journal of Corporate Finance, 62. 101578. ISSN 0929-1199

https://doi.org/10.1016/j.jcorpfin.2020.101578

Is Language an Economic Institution? Evidence from R&D Investment *

Jianxin Daniel Chi† Xunhua Su‡ Yun Tang§ Bin Xu¶

Forthcoming in *Journal of Corporate Finance*

Abstract

Some languages encode future timing more ambiguously than others. We identify two economic channels through which more ambiguous reference to future timing leads to higher levels of R&D investment. Our empirical tests on country- and firm-level R&D investment confirm this prediction, even after controlling for an extensive set of formal and informal economic institutions and addressing endogeneity concern in multiple ways. Tests on patent generation provide further evidence that ambiguous reference to future timing leads to more innovation.

Keywords: language, R&D, innovation, economic institution, future-time reference

JEL Classification: D02, G30, O30, O32, Z13

*This version: Dec. 16, 2019 (first version: Apr. 13, 2016). We are very grateful to an anonymous reviewer whose comments and suggestions significantly improved this paper. For helpful comments and suggestions, we thank Rita Biswas, Murillo Campello, Si Cheng, Henrik Cronqvist, Vincent Hsu, Mel Jameson, Yanxi Li, Laura Liu, Bibo Liu, Qi Liu, Ian McDonough, Yuandong Mu, Nagpurnanand Prabhala, Marc Rieger, Francisco Santos, Xuan Tian, Andrew Vivian, Mei Wang, Lucy White, Yangru Wu, Han Xia, Kuo Zhang, Xinde Zhang, Rui Zhong; seminar participants at CUFE (Beijing), DUT (Dalian), Hainan University, NHH (Bergen), Peking University, Qingdao University, Qingdao University of Technology, Queen's University Belfast, Shandong University, SUFE (Shanghai), UIBE (Beijing), UNLV, and Zhejiang Gongshang University; and conference participants at CICF (2016 Xiamen), RBFC (2016 Amsterdam), FMA (2016 Las Vegas), FRCFM (2017 Takamatsu), and AsianFA (2018 Tokyo). Chi acknowledges support from the Key Program of National Natural Science Foundation of China (71731003). The views expressed in this paper are those of the authors and do not necessarily represent those of the OECD.

†Department of Finance, University of Nevada, Las Vegas (UNLV). daniel.chi@unlv.edu

‡Department of Finance, Norwegian School of Economics (NHH). xunhua.su@nhh.no

§Organisation for Economic Co-operation and Development (OECD). yun.tang@oecd.org

¶Department of Finance, University of Leeds. b.xu@leeds.ac.uk

1 Introduction

Institutions determine resource allocation and economic growth. While North (1991) and Williamson (2000) highlight the importance of both formal and informal institutions, the finance literature has traditionally focused on formal institutions such as law and investor protection. Relatively little attention has been paid to informal institutions such as culture (Karolyi, 2016). One of the early studies on informal institutions and finance is Stulz & Williamson (2003), who ask an intriguing question: since overwhelming evidence shows that investor protection (a formal institution) promotes economic growth, why does investor protection still differ substantially across countries? They find that religion (an informal and persistent institution) is a significant determinant of investor protection. Since then, studies of culture and finance have grown and gained more acceptance, and culture has been generally recognized as an informal economic institution that has wide-ranging impacts on finance.[1]

Language, like culture, affects cognition and decision making. In recent years, linguistics and psychology research has produced increasing evidence that people who speak different languages think differently.[2] Is language also an economic institution that affects resource allocation and economic outcome? To operationalize this research question, we study R&D investment. Since at least Joseph Schumpeter, R&D investment has been recognized as a key resource allocation decision that affects innovation and growth. Yet despite this recognition, there remain substantial and *persistent* differences in the level of R&D investment across countries. How could this be? One piece of this puzzle, as documented by numerous studies, is the difference in formal economic institutions, such as property rights protection (e.g., Allred & Park, 2007). The second piece lies in culture, such as individualism (Shao et al., 2013), which is persistently different across countries. Some language features, when compared with various cultural attributes, are even more persistent. We hypothesize that another piece of the puzzle lies in language.

The particular language feature that underlies our hypothesis is *future-time reference* (FTR),

[1] See reviews on culture and finance by Aggarwal et al. (2016).

[2] This is the central hypothesis of *linguistic relativity*. For example, see reviews by Boroditsky (2011) and Wolff & Holmes (2011).

which refers to how explicitly a language marks future timing. Strong-FTR languages, such as English and French, have explicit tense to mark the future. For example, when referring to future timing in English, the usage of future tense as denoted by *will* or *going to* is grammatically obligatory. In contrast, weak-FTR languages, such as German or Chinese, have no tense and no obligatory marking of future timing. Another way to think about the difference is that compared to strong-FTR languages, weak-FTR languages have more ambiguous reference to future timing. Boroditsky (2001) finds that language is most powerful in shaping thought about abstract domains like time. Chen (2013), in a pioneering study that links language to economic decision making, finds that individuals speaking weak-FTR languages exhibit more long-term-oriented behavior, such as saving more for retirement and making more health-conscious decisions.

Chen's findings notwithstanding, more evidence is needed to establish language as an economic institution. Specifically, does language FTR affect resource allocation at the firm and country levels? And if so, through what channels, and what are the real economic consequences? Also importantly, how distinct is the language effect from the cultural effects that have been documented in the literature, such as the effects of religion, trust, and individualism? R&D investment provides a suitable setting to study these questions.

Why do we expect language FTR features to affect R&D investment? R&D is long-term investment with highly uncertain rewards. Given the long-term and uncertain nature, there are two economic channels through which FTR can affect the level of R&D investment, and both channels suggest that weak FTR (i.e., more ambiguous reference to future timing) leads to higher levels of R&D investment.

The first channel is "time preference". Time preference means that to a weak-FTR language speaker, the future seems closer, and thus the present value of a future reward appears larger, which leads to higher current investment. Panel A in Figure 1 illustrates this channel. The second channel is "reward-timing uncertainty". Speakers of weak-FTR languages, because of their ambiguous marking of future timing, perceive greater uncertainty in future timing. This perceived reward-timing uncertainty, combined with the convexity of the time discount function, leads to a higher perceived present value of future rewards, and therefore higher investment. This effect, illustrated in Panel B in Figure 1, is analogous to Jensen's Inequality. In the hypothesis-development section, we present a more rigorous illustration of these two channels. But it is worth emphasizing that both channels suggest that weak FTR leads to higher R&D investment, which is our main empirical hypothesis.

Our central finding is that weak FTR is associated with significantly higher R&D investment and patent generation. Is the association between FTR and R&D causal? Since language FTR features are highly stable over time (Roberts et al., 2015), reverse causality running from R&D to FTR is unlikely. Therefore, potential endogeneity can only arise from omitted correlated variables. We tackle this identification challenge in four ways. First, we perform a cross-country analysis while controlling for a wide range of country characteristics that are potentially correlated with language and may also affect R&D investment. These characteristics include economic and financial-market development, legal origin, investor rights protection, patent protection, and various cultural attributes. Employing a panel data set of 52 countries during 1996-2013, we find robust evidence that weak FTR is statistically and economically related to higher levels of R&D investment. The p-value is below 0.1% in most specifications, and depending on the regression specification, weak FTR is related to a business-sector R&D/GDP ratio that is 20%-30% above the sample mean.

Since the FTR measure is time invariant, we are not able to control for country fixed effects. Is it likely that a country omitted variable is driving the positive relation between FTR and R&D investment? Note that the strong- vs. weak-FTR classification groups together many countries that have drastically different cultural, legal, political, historical, and economic attributes. For example, weak-FTR countries include Brazil, China, Germany, and Japan. For a country omitted variable to drive the result, this omitted variable needs to be distinct from the variables that are already controlled for and needs to exert a common effect across all these countries. *In addition*, this omitted variable also needs to exert a common effect (in the opposite direction) across strong-FTR countries, which is a diverse group that includes Argentina, Israel, South Korea, and United States. In light of extant literature, it is difficult to conceive what this omitted correlated variable might be.

To further alleviate the concern about omitted-variable bias, we conduct a second set of tests at the firm level. This allows us to control for firm characteristics that affect R&D and are potentially correlated with FTR. The firm-level sample consists of over 31,000 firms from 34 countries and spans from 1985 to 2013. After controlling for important firm characteristics that affect R&D (firm size, leverage, market-to-book, operating profit, cash holdings, asset tangibility, and capital intensity), FTR continues to exhibit a statistically significant and economically large effect on R&D. The p-value is consistently below 0.1%, and depending on the regression specification, weak FTR is related to a R&D/Assets ratio that is 20%-36% above the sample mean.

Our third identification strategy is to use pre-1500 CE crop return as an instrumental variable for *Weak-FTR*. Galor et al. (2018) show that pre-1500 CE crop return is positively associated with the use of periphrastic future tense in a language. Consistent with their finding, we find that pre-1500 CE Crop Return is negatively associated with *Weak-FTR*. The instrumented *Weak-FTR* has a significantly positive effect on both country-level and firm-level R&D investment.

Our fourth identification strategy is to conduct within-country tests. This essentially controls for all observable and unobservable country-level attributes and directly addresses the concern about country-level omitted-variable bias. We identify a group of CEOs of U.S. firms that have non-U.S. nationalities with weak-FTR languages. Employing a propensity score methodology, we match these CEOs with a control group of strong-FTR CEOs on multiple firm characteristics. We find that firms with CEOs who have weak-FTR nationalities invest significantly more in R&D.[3]

Overall, the country-level, firm-level, and within-country tests provide robust empirical evidence, consistent with our theoretical prediction that weak FTR leads to higher R&D investment. These findings survive numerous robustness checks, including alternative measures of language FTR, alternative regression specifications, subperiod tests, and a correlated-random-effects model that explicitly controls for country-level omitted correlated variables. We also examine patent generation as an alternative measure of innovation activity and find consistent evidence that weak FTR is associated with significantly more patent generation. Despite all the

[3]We conduct two more within-country tests and find consistent results. The first test exploits the language difference within Belgium, and the second test exploits the handover of Hong Kong from U.K. to China.

evidence, we acknowledge that without an exogenous shock that produces random assignment of weak- vs. strong-FTR languages, we cannot definitively conclude causality. As we discussed earlier, however, the bar is high for a correlated omitted variable to explain away the findings.

In recent years, the *linguistic relativity* literature has produced "*a solid body of empirical evidence showing how languages shape thinking*" (Boroditsky, 2011, p. 63). Chen (2013) documents that language is associated with individuals' economic behavior. We extend Chen's work and document how linguistic features can affect resource allocation at corporate and country levels. R&D investment is crucial to generate innovation and support economic growth. Therefore, the evidence that linguistic features affect R&D and patent generation suggests that language is an important economic institution, similar to the now widely accepted notion that culture is an important economic institution. The language effect is robust to an extensive set of culture proxies, suggesting that the language effect is distinct from the culture effects that have been documented in the literature.

Our study contributes to the emerging literature on how language affects economic decision making. Guin (2016) and Paule-Paludkiewicz et al. (2016) confirm Chen's linguistic-savings hypothesis. Chen et al. (2017) find similar savings behavior by corporations. Pérez & Tavits (2017) survey a group of bilingual speakers and randomly assign the language used in the survey (weak- or strong-FTR). They find that when being asked in the weak-FTR language, the subjects are more likely to support future-oriented policies. In accounting literature, Fasan et al. (2016) and Kim et al. (2017) find that firms in weak-FTR countries are less likely to engage in earnings management. In a recent survey on linguistic structures and economic outcomes, Mavisakalyan & Weber (2017) review more evidence that language affects the cognitive process and decision making.[4]

The literature on language and institutions is in its infancy. But just as psychology and sociology (culture) have enriched our understanding of finance and economics, language can

[4]Our study also contributes to the growing literature on the determinants of corporate innovation (e.g., Hsu et al., 2014; Luong et al., 2017; Bhattacharya et al., 2017) by suggesting that the language of corporate managers is an important driver of corporate R&D investment. In addition, our study adds to the literature on the positive relation between R&D and firm stock returns (e.g., Cohen et al., 2013; Gu, 2016; Croce et al., 2019). Specifically, our evidence suggests that the language of corporate managers may affect stock returns through decisions on R&D investment.

potentially provide new vantage points. A particular challenge, however, is to identify the economic channels that link linguistic features to economic decisions and outcomes. We provide plausible theoretical explanations of how a language's FTR feature affects R&D investment, but a different theoretical explanation may be required for another language feature or another economic outcome. Viewing this challenge through a more optimistic lens, we heed Oliver Williamson's recommendation that while "*awaiting a unified theory, we should be accepting of pluralism*" (Williamson, 2000, p.595).

The remainder of the paper proceeds as follows. Section 2 develops our main hypothesis and describes the empirical models. Section 3 examines the effect of language FTR on country-level R&D investment. Section 4 examines the effect of language FTR on firm-level R&D investment. Section 5 addresses endogeneity concerns and conducts further robustness checks. Section 6 tests the relation between language FTR and country-level innovation outputs. Section 7 concludes.

2 Hypothesis and Empirical Strategy

2.1 Language FTR

Future-time reference (FTR) describes how explicitly a language encodes the timing of future events, and FTR differs substantially across languages (Dahl, 2000). A strong-FTR language speaker is required to grammatically mark future time, while a weak-FTR language speaker does not need to grammatically distinguish between the present and the future. For example, in the sentence *It will/is going to be cloudy tomorrow*, English marks the future using "will/is going to". It is grammatically wrong to use the present tense of the copula, "is", to talk about the weather tomorrow. However, in Chinese there are no tenses. Chinese speakers can simply replace *jintian* (today) with *mingtian* (tomorrow) without changing any other part of the sentence.[5] Among the main languages, German, Chinese, and Japanese have weak FTR, while English, French, and Spanish have strong FTR. Appendix I summarizes FTR classification for the 52 countries in our sample.

[5]See Chen (2013) for a more detailed discussion on how different languages mark time differently.

English and Chinese speakers talk about the weather *today*:

English speaker:	*It*	*is*	*cloudy*	*today.*
Chinese speaker (pinyin):	*Jintian*	*shi*	*duoyun.*	
Chinese (translation):	*Today*	*is*	*cloudy.*	

English and Chinese speakers talk about the weather *tomorrow*:

English speaker:	*It*	*will/is going to be*	*cloudy*	*tomorrow.*
Chinese speaker (pinyin):	*Mingtian*	*shi*	*duoyun.*	
Chinese (translation):	*Tomorrow*	*is*	*cloudy.*	

The difference in FTR can cause at least two effects. First, relative to a strong-FTR speaker, a weak-FTR speaker likely has a more blurred distinction between present and future, and may perceive future events closer to the present. When discounting a future reward, it is conceivable that the weak-FTR speaker will apply a smaller discount factor and hence perceive a higher present value. This is the "time preference" effect. Second, relative to a strong-FTR speaker, a weak-FTR speaker likely has a less precise, or more ambiguous, perception about future timing. That is, the weak-FTR speaker may perceive a wider distribution in the timing of a future reward. This is the "timing uncertainty" effect.

Do language features actually influence how the speaker think and behave? The *linguistic relativity hypothesis* (LRH), also known as Sapir-Whorf Hypothesis (Whorf, 1956), says yes. The debate of whether LRH holds true went on for the better part of the 20th century. But since the 1990s, the linguistics and psychology literatures have produced increasing evidence supporting LRH. The question of whether language features, particularly how a language encodes time, affect *economic behavior* does not attract much attention until Chen (2013). He finds that weak-FTR language speakers save more and make more health-conscious decisions.

There is limited evidence, however, that language FTR affects resource allocation at firm or country level. We next develop the hypothesis on how FTR can affect R&D investment decisions.

2.2 Hypothesis Development

Based on Chen (2013), we first develop a simple model to show how time preference and timing uncertainty can affect a manager's perception of the timing of future rewards and hence the decision on R&D investment.

Consider a firm with an R&D project. By investing x today, the firm will be rewarded $R(x)$ at time t. We assume that the reward function R is increasing and concave in x, i.e., $R'(x) > 0$ and $R''(x) < 0$, reflecting diminishing return to scale from R&D investment. As one of the key features of R&D investment, the reward time, t, is uncertain.[6] Let t follow a distribution $F(t)$, defined in interval $[0, +\infty]$.

The firm maximizes its net expected profit:

$$\max_{x} \ P(x) = -x + \int_{0}^{+\infty} e^{-\delta t} R(x) \ dF(t) \tag{1}$$

where δ denotes the discount rate. The first-order condition for this problem is

$$R'(x^*) \cdot \int_{0}^{+\infty} e^{-\delta t} \ dF(t) = 1 \tag{2}$$

From Equation (2), we can solve for the optimal R&D investment, x^*. Given the project (and hence the R function), this optimal investment is determined by the discount rate δ and the distribution function $F(t)$.

First, a lower discount rate increases the present value of the reward and hence increases the current optimal investment. Mathematically, taking the first-order derivative of δ for Equation (2), we have

$$\frac{\partial x^*}{\partial \delta} = \frac{1}{R''(x^*)} \cdot \left[\int_{0}^{+\infty} e^{-\delta t} \ dF(t) \right]^{-2} \int_{0}^{+\infty} t e^{-\delta t} \ dF(t) < 0 \tag{3}$$

Notice that $R''(x^*) < 0$, so $\partial x^* / \partial \delta < 0$. That is, x^* increases as δ decreases. If a weak-FTR speaker perceives the future to be closer to the present, she will apply a smaller δ, perceive a higher x^*, and therefore make higher investment in R&D. This is the "time preference" channel,

[6] Another key feature of R&D is that the level of reward is uncertain. For simplicity, we consider only deterministic reward. Randomizing R does not change the conclusion of the model.

which we illustrate in Panel A of Figure 1.

Second, due to the convexity of the exponential function (in terms of t), a "wider" distribution $F(t)$ (i.e., a greater timing uncertainty of the reward) increases the present value of the reward, and hence increases current investment. To see this, consider two distributions, $F_1(t)$ and $F_2(t)$. Assume $F_2(t)$ exhibits more uncertainty than $F_1(t)$ in the sense that $F_2(t)$ is a mean-preserving spread (MPS) of $F_1(t)$ (Rothschild & Stiglitz, 1970). As $-e^{-\delta t}$ is a strictly concave function of t, appealing to the important property of MPS,[7] we have

$$\int_0^{+\infty} -e^{-\delta t}\ dF_1(t) > \int_0^{+\infty} -e^{-\delta t}\ dF_2(t) \tag{4}$$

or equivalently

$$\int_0^{+\infty} e^{-\delta t}\ dF_1(t) < \int_0^{+\infty} e^{-\delta t}\ dF_2(t) \tag{5}$$

Combining (5) and (2), we have $R'(x_1^*) > R'(x_2^*)$ and hence $x_1^* < x_2^*$. That is, due to the convexity of the time discount function, greater reward-timing uncertainty increases the optimal R&D investment.[8] This is the "timing uncertainty" channel, which we illustrate in Panel B of Figure 1.

We use a more specific example to further illustrate the intuition of the two channels. Suppose that $F(t)$ has only two values, $T - \epsilon$ and $T + \epsilon$, with equal probabilities, where ϵ is a positive number, capturing timing uncertainty. A higher ϵ indicates a higher timing uncertainty. Applying Equation (2), we have

$$\frac{1}{2} \cdot e^{-\delta T} \cdot (e^{\epsilon} + e^{-\epsilon}) \cdot R'(x^*) = 1$$

[7]MPS has the following property. Given that $F_2(t)$ is a MPS of $F_1(t)$, for any convex function g, $\int_0^{+\infty} g(t)\ dF_1(t) < \int_0^{+\infty} g(t)\ dF_2(t)$. The proof can be seen in Rothschild & Stiglitz (1970).

[8]This relationship can be illustrated using a simple example. Suppose an R&D investment today (time 0) is expected to generate one unit of cash flow at time 2, and the discount factor is β $(0 < \beta < 1)$. The present value of this future reward is thus β^2. Now suppose the timing of the future reward is perceived to be uncertain: it may occur at time 1 or time 3. The present value thus becomes $(\beta + \beta^3)/2$. It is straightforward to see that $(\beta + \beta^3)/2 > \beta^2$ because $(\beta + \beta^3)/2 - \beta^2 = \beta(1-\beta)^2/2 > 0$. That is, reward-timing uncertainty increases the present value of the payoff, and therefore encourages R&D investment.

It follows that

$$\frac{\partial x^*}{\partial \delta} = \frac{2Te^{\delta T}}{R''(x^*)} \cdot \frac{1}{e^{\epsilon} + e^{-\epsilon}} < 0 \tag{6}$$

$$\frac{\partial x^*}{\partial \epsilon} = -\frac{2e^{\delta T}}{R''(x^*)} \cdot \frac{e^{\epsilon} - e^{-\epsilon}}{(e^{\epsilon} + e^{-\epsilon})^2} > 0 \tag{7}$$

Equations (6) and (7) respectively show that the optimal investment is determined by the discount rate (i.e., time preference) δ and timing uncertainty ϵ.

As Chen (2013) argues, time preference and timing uncertainty are the two channels through which language FTR can affect individuals' savings and health decisions. Both of these two channels, as we illustrate above, suggest that weak FTR leads to higher R&D investment, which is our main testable hypothesis.

Hypothesis: *Weak FTR leads to higher R&D investment.*

2.3 Empirical Design and Main Variables

To test our hypothesis, we conduct cross-country analysis at both the country and firm levels. At the country level, we implement the following empirical model:

$$R\&D_{kt} = \alpha + \beta \cdot \textit{Weak-FTR}_k + \Omega \mathbb{X}_{kt} + \epsilon_{kt} \tag{8}$$

$R\&D_{kt}$ is R&D expenditures of country k in year t. We use two measures of national R&D expenditures: (1) aggregate R&D of the business sector as a percentage of GDP (*Business R&D/GDP*), and (2) aggregate R&D as a percentage of GDP (*Total R&D/GDP*). $\textit{Weak-FTR}_k$ is a binary variable that equals 1 if the main language of a country is classified as a weak-FTR language, and 0 otherwise. Alternatively, we replace this binary variable by two continuous measures of FTR strength, *Verb Ratio* and *Sentence Ratio*, developed by Chen (2013). The ratios are defined respectively as the frequency of verbs and sentences that are grammatically future-marked in weather forecasts of various languages. By construction, they are inverse measures of weak-FTR and range from zero to one. The above variables are defined in Panels A and B of Appendix II.

$\mathbb{X}_{kt}$ is a set of country-specific control variables, including economic, legal, cultural, and religious variables. First, the three economic variables are the logarithm of GDP per capita in U.S. dollars (*GDP per capita*), stock market capitalization as a percentage of GDP (*Stock Market*), and domestic credit to private sectors as a percentage of GDP (*Credit Market*). Low GDP countries are typically less innovative and invest less in R&D. Including GDP per capita in the regression controls for such an effect. Stock and credit markets are important for a country's financial development that affects R&D (e.g., Maskus et al., 2012; Hsu et al., 2014).

Second, the legal variables consist of legal origins (La Porta et al., 2008) and proxies for protection of *Shareholder Rights*, *Creditor Rights*, and *Patent Rights*, defined by Djankov et al. (2007); Djankov et al. (2008); and Park (2008). Countries in our sample belong to four legal origins: UK, French, German, and Scandinavia.

Third, we control for important cultural attributes including *Individualism*, *Uncertainty Avoidance*, *Masculinity*, *Power Distance*, and *Long-term Orientation* (Hofstede, 1984, 2001). In addition, we control for *Trust* because trust affects the development of financial markets (Guiso et al., 2008) and hence may affect investment in R&D. Various studies have shown that culture affects firms' investment policy (e.g., Li et al., 2013; Shao et al., 2013; Ahern et al., 2015).

Fourth, the religious variable (*Catholic*) is a dummy indicating whether the majority of a country's inhabitants are catholic. All the country-level control variables are defined in Panel D of Appendix II. We also control for continent fixed effects, as different languages within a continent may share similar components. For example, the Chinese language has influenced the languages of Japan and Korea.

At the firm level, we run the following OLS regression:

$$R\&D_{ikt} = \alpha + \beta \cdot \textit{Weak-FTR}_k + \Omega \mathbb{X}_{kt} + \Gamma \mathbb{Z}_{ikt} + \epsilon_{ikt} \tag{9}$$

where $R\&D_{ikt}$ is R&D investment of firm i of country k in year t. We construct two measures of firm R&D investment, $R\&D/Assets$ and $R\&D/Sales$, defined as R&D expenditures divided by total assets or sales. We replace missing R&D values by zero. Koh & Reeb (2015) document that over 10% of U.S. firms with missing R&D in their financial statements receive patents, so it might be problematic to recode all missing R&D as zero. Following Koh & Reeb (2015), we include a dummy variable indicating missing R&D (*R&D-missing*). $\mathbb{X}_{kt}$ includes country-specific control variables, as described above. $\mathbb{Z}_{ikt}$ includes firm-specific control variables, such as *Firm Size*, *Tangibility*, *Profitability*, *Leverage*, *Market-to-Book ratio*, *Capital Intensity*, and *Cash Ratio*, which may all affect R&D investment (e.g., Brown et al., 2009). We also control for *CF Volatility*, defined as the industry-average standard deviation of firm profitability in the past five years. *CF Volatility* is a proxy for reward-size uncertainty and is expected to have a negative effect on firm R&D investment. All these control variables are defined in Panel C of Appendix II. We also control for industry, year, and continent fixed effects, and cluster standard errors at the firm level.

3 Country-level R&D Investment

3.1 Data, Sample, and Statistics

We obtain data on national R&D investment between 1996 and 2013 from United Nations Educational, Scientific and Cultural Organization (UNESCO) Institute for Statistics.[9] The full sample has 52 countries on six continents, including almost all large economies of the world. Among them, 15 are classified as countries with weak-FTR languages.

As mentioned earlier, we construct a large set of country-level economic, legal, cultural, and religious variables. Table 1 presents summary statistics. Business R&D and total R&D are on average 1% and 1.5% of GDP. 29% of the country-year observations are of weak FTR. Summary statistics of the rest are comparable to related studies in the literature.

We compute time-series average of *Business R&D/GDP* for each country. We tabulate the averages in Appendix I and plot them in Figure 2. Figure 2 shows that R&D appears to be higher for weak-FTR countries. Seven of the top 10 countries are weak-FTR countries, while among the bottom 20 countries only one country is a weak-FTR country. We also find a similar pattern in Figure 3, which shows each country's average *Total R&D/GDP*. Of course, although the univariate comparisons are suggestive, we need to control for plausible determinants of R&D

[9] The data are downloadable at `http://www.uis.unesco.org/DataCentre/Pages/BrowseScience.aspx`

that are also potentially correlated with language FTR. But first we want to see how distinct FTR is from all these determinants.

Table 2 reports the correlation between FTR and a host of economic, cultural, legal, and religious variables. First, as expected, *Weak-FTR* is negatively correlated with the two inverse measures of language FTR, *Verb Ratio* and *Sentence Ratio*. Second, *Weak-FTR* seems to be sufficiently distinct from the set of control variables, including those proxies for informal economic institutions of culture and religion.

3.2 FTR and Country-level R&D Investment

Table 3 reports the regression result of Equation (8). Column (1) shows that *Weak-FTR* is positively associated with *Business R&D/GDP* when controlling for economic and financial market development. In Columns (2)-(3) we progressively add controls of formal economic institutions (legal and property-rights protection) and informal economic institutions (culture and religion). It is interesting to note that the magnitude and statistical significance of *Weak-FTR* coefficient increases substantially after we control for culture attributes in Column (3). That is, not only is the language effect not subsumed by the culture effects, it actually becomes more distinctive when the culture effects are controlled for. This suggests that language FTR likely captures a new aspect of informal institution than those captured by the culture attributes. In Column (4) we control for continent and year fixed effects. The coefficient of *Weak-FTR* remains highly significant and economically large. The coefficient of 0.002 in Column (4) suggests that weak-FTR countries on average have 0.2 percentage-points higher Business R&D/GDP ratio, which is 20% of the sample mean of 1% (based on the summary statistics in Table 1). In Columns (5)-(8) we use *Total R&D/GDP* as dependent variable and obtain similar finding.

The coefficients on the control variables are mostly consistent across specifications. R&D is positively associated with the level of economic development – in all columns *GDP per capita* has a positive coefficient. Stock market development, *Stock Market*, has a negative coefficient. This may seem counterintuitive because stock market financing should be supportive of R&D financing. We do find that without other control variables, *Stock Market* has a significantly positive coefficient. Other cross-country studies also find that in a multivariate setting, stock market capitalization is negatively related to R&D investment (e.g., Shao et al., 2013). Shareholder and creditor rights are negatively related to R&D, consistent with findings in prior literature (e.g., Li et al., 2013; Acharya et al., 2011). In contrast, *Patent Rights* is positively associate with R&D because stronger patent protection allows innovators to retain more benefits from their R&D investment (Varsakelis, 2001; Czarnitzki & Toole, 2011).

Among the culture variables, *Individualism* is not significant in explaining Business R&D after controlling for fixed effects, but is significantly positively in explaining total R&D, consistent with the finding in Shao et al. (2013). *Uncertainty Avoidance* is insignificant after controlling for fixed effects. *Power Distance* has a negative and significant effect on R&D investment. This could be attributed to the fact that in low power distance societies, social mobility is relatively stronger. To improve social status, people in low power distance countries need to invest more in technology and knowledge (Varsakelis, 2001). *Masculinity* has an ambiguous relation with R&D. As expected, *Long-term Orientation* is positively related to R&D because R&D investment requires long-term commitment. *Trust* is negatively related to R&D, which seems to be consistent with the findings in Helliwell (1996) and Roth (2009) that trust is negatively related to economic growth. Finally, Catholic-dominated countries on average have lower levels of R&D expenditures, perhaps because Catholics have higher levels of risk aversion (Halek & Eisenhauer, 2001).

Notwithstanding various possible explanations on the control variables, the main empirical finding on language FTR is consistent with our theoretical prediction and also suggests that the effect of language FTR is distinct from the effects of other formal and informal institutions. Is the effect of FTR on R&D causal? According to John Stuart Mill (1884), one can infer causality when (1) the cause precedes the effect, (2) the cause is related to the effect, and (3) there is no other explanation for the effect other than the cause. For our case, the first two criteria are satisfied, so the challenge is to satisfy the third one. Theoretically, we provide an explanation of why FTR can affect R&D investment. Empirically, the main concern is that some country-level omitted variable is causing spurious correlation between FTR and R&D. But what might this omitted variable be?

As mentioned earlier, for a country omitted variable to drive the result, this omitted variable needs to be distinct from the variables that are already controlled for. More importantly, it needs to exert a common effect across weak-FTR countries, which have drastically different cultural, legal, political, historical, and economic attributes. In addition, this omitted variable also needs to exert a common effect (in the opposite direction) across strong-FTR countries. So the bar seems high for such an omitted variable to drive the finding.

A country-fixed-effects specification will go a long way to address the omitted variable concern. Unfortunately, FTR is time invariant and will be perfectly collinear with country fixed effects. We address the omitted variable concern in three additional ways. First, in Section 4 we conduct firm-level tests and control for additional firm-level characteristics that are known determinants of R&D and may also be correlated with FTR. Second, in Section 5 we perform 2SLS regressions and within-country tests. The within-country tests essentially control for all country-level characteristics, observable or not.

4 Firm-level R&D Investment

4.1 Data, Sample, and Statistics

We obtain firm accounting and financial data from Thomson Reuters Worldscope database. The Worldscope database offers fundamental data on the world's leading public companies and covers more than 95% of global market capitalization. We restrict our sample to the primary quotes (i.e., non-cross-listings), and exclude financial firms (ICB code 8000) and utility firms (ICB code 7000). We further delete firm-years with missing data for any of the firm-specific and country-specific control variables used in Equation (9). This results in a final sample consisting of 34 countries, 31,650 firms and 215,877 observations for the 1985-2013 period. We start our sample from 1985 because Worldscope coverage is relatively limited prior to 1985. We winsorize all firm-level variables at the 1st and 99th percentiles to reduce the influence of extreme outliers.

Table 4 presents summary statistics for firm-level R&D investment. The sample means for *R&D/Assets* and *R&D/Sales* are 2.5% and 4.2%. 34.9% of the observations are of weak FTR. The two alternative FTR measures, *Verb Ratio* and *Sentence Ratio*, have means of 54% and 60% with sufficiently wide standard deviation, suggesting that our observations are not concentrated on any one type of countries (weak- or strong-FTR).

4.2 FTR and Firm-level R&D Investment

Table 5 reports regression result of Equation (9). In Column (1), we regress $R\&D/Assets$ on *Weak-FTR* and country-level variables as well as continent fixed effects. Column (2) further controls for industry fixed effects and a set of firm-specific variables that are potential determinants of R&D, including firm size, asset tangibility, profitability, financial leverage, market-to-book ratio, capital intensity, cash ratio, and cash-flow volatility. Column (3) further controls for year fixed effects. In Columns (4)-(6), we use $R\&D/Sales$ as dependent variable. The consistent result is that weak FTR is associated with significantly higher R&D. Based on the result in Column (3), firms in weak-FTR countries invest 0.8 percentage points higher in R&D, which is one third of the sample mean of 2.5%.

The coefficient estimates of some firm-specific control variables also warrant some discussion. First, *Firm Size* and *Profitability* both have a negative coefficient, consistent with previous studies (e.g., Barker & Mueller, 2002; Hitt et al., 1991). Second, high leverage is negatively associated with R&D investment because firms are often reluctant to invest in long-term and risky R&D projects if they face high financial distress costs associated with high debt levels. Third, the positive coefficient of *Market-to-Book* suggests that firms with greater growth opportunities invest more in R&D (e.g., Ryan & Wiggins, 2002). The positive coefficient of *Capital Intensity* indicates that highly capital-intensive firms are often in industrial or manufacturing sectors rather than service sectors and therefore tend to be more R&D intensive (Hirshleifer et al., 2012). The positive relation between *Cash Ratio* and R&D investment is consistent with the well-established notion that firms finance R&D largely with internal funds. Finally, *CF Volatility* is negatively associated with R&D investment, consistent with real-options theory that reward-size uncertainty increases the option value of waiting and therefore lowers current investment (Dixit & Pindyck, 1994).

In Table 6, we employ two alternative measures of language FTR, *Verb Ratio* and *Sentence Ratio*. As expected, these two inverse measures of *Weak-FTR* have significantly negative effects on R&D. These firm-level results corroborate the country-level results and further alleviate the omitted variable concern.

4.3 Correlated Random Effects

If we take a look at Equation (9) and think through the potential endogeneity caused by an omitted correlated variable, it is possible that some country-level unobservable heterogeneity, call it a_k, is hidden in ϵ_{ikt} in the following form: $\epsilon_{ikt} = a_k + u_{ikt}$, where a_k is correlated with $Weak\text{-}FTR_k$, and u_{ikt} is the idiosyncratic error term. We cannot explicitly control for a_k through a country-fixed-effects model because the country FEs will be perfectly collinear with $Weak\text{-}FTR_k$. It is plausible that (1) a_k affects both $Weak\text{-}FTR_k$ and the control variables $\mathbb{X}_{kt}$ and $\mathbb{Z}_{ikt}$, and (2) a_k's effects on $\mathbb{X}_{kt}$ and $\mathbb{Z}_{ikt}$ and its effect on $Weak\text{-}FTR_k$ are correlated. If so, controlling for a_k's effects on $\mathbb{X}_{kt}$ and $\mathbb{Z}_{ikt}$ will (at least partly) control for a_k's effects on $Weak\text{-}FTR_k$. To operationalize this reasoning, we implement a correlated-random-effects (CRE) model as follows.

Suppose $a_k = a + \Pi\bar{\mathbb{X}}_k + \Phi\bar{\mathbb{Z}}_{ik} + r_k$, where the overbar denotes the time averages of the control variables $\mathbb{X}$ and $\mathbb{Z}$, and r_k is the remaining country effect that is unrelated to $\mathbb{X}$ or $\mathbb{Z}$. We run a CRE model by explicitly controlling for $\bar{\mathbb{X}}_k$ and $\bar{\mathbb{Z}}_{ik}$ in Equation (9). This way we have fully corrected the bias caused by a_k on the coefficient estimates of $\mathbb{X}_{kt}$ and $\mathbb{Z}_{ikt}$. If the two assumptions in the previous paragraph hold, we will have also (partly) controlled for the effect of the omitted variable on $Weak\text{-}FTR_k$. See Wooldridge (2013) Section 14.3 for a more detailed, yet succinct, explanation on CRE.[10]

Table 7 reports the CRE-model result by running Equation (9) while also controlling for the time averages of all the timing-varying control variables. All the time averages have significant coefficients, suggesting that these time averages are indeed correlated with R&D investment. More relevant to our interest, the coefficients on *Weak-FTR* remain highly significant and actually become larger. For example, in Column (1) the coefficient on *Weak-FTR* is 0.010 vs. 0.008 as in Column (3) of Table 5.

Interpretation of the CRE result depends on one's confidence in the two assumptions earlier.

[10] As explained in Wooldridge (2013), including the time averages of $\mathbb{X}_{kt}$ and $\mathbb{Z}_{ikt}$ is equivalent to controlling for FEs for $\mathbb{X}_{kt}$ and $\mathbb{Z}_{ikt}$. The benefit of a CRE model over a FE model is that we can estimate the effect of a time-invariant variable, which in our case is $Weak\text{-}FTR_k$.

While we think the assumptions are plausible, the exact extent to which the CRE model mitigates the omitted-variable bias is unknowable. So caution is warranted when interpreting the CRE result. Nevertheless, we take comfort in that the coefficient estimates on *Weak-FTR* actually increase in the CRE model, and therefore regard the CRE result as one piece of evidence consistent with our central prediction and with the rest of the empirical evidence.

4.4 Robustness Checks

In Table 8, we conduct the following robustness checks.

Non-US firms. Over 25% of observations in our sample are U.S. firms. Since language FTR is invariant within a country, a large presence of U.S. observations may have overly influenced the estimation. We rerun the test by excluding U.S. firms. The effect of *Weak-FTR* on R&D remains positive and highly significant.

Subperiods. We also run the test for three subperiods of 1985-1999, 2000-2007, and 2008-2013, and find consistent results. All the coefficients on *Weak-FTR* are positive and statistically significant.

Excluding observations with missing R&D. In our main analysis, we replace missing R&D values with zero and control for a dummy for missing R&D. We rerun the tests with only observations with non-missing R&D. We still find a positive and significant relation between *Weak-FTR* and firm R&D.

The adoption of IFRS. Many firms in various countries started following the International Financial Reporting Standards (IFRS) during our sample period. The adoption of IFRS makes accounting information more comparable across countries (e.g., Daske et al., 2008) and may affect corporate R&D policies by changing credit supply. We include a binary variable, *IFRS*, which equals to one if a firm follows IFRS instead of local accounting standards. The coefficient estimate on *Weak-FTR* remains positive and highly significant.

Tobit model. Firm R&D/Assets and R&D/Sales ratios are limited at the lower bound of zero. In such case, the OLS estimator may be biased and a Tobit maximum likelihood specification may be more appropriate. We run Tobit regressions and find consistent results.

The coefficient estimates on *Weak-FTR* remain positive and highly significant, and in most cases become larger.

Controlling for lagged R&D. One way to mitigate the omitted-variable bias is to control for lagged dependent variable. To the extent that the effect from the omitted variable is persistent over time, the lagged dependent variable will subsume a significant portion of the effect of the omitted variable. But this specification is punishing for other explanatory variables because a significant portion of their effects will also be subsumed by the lagged dependent variable. We control for lagged R&D and find that the coefficient estimate on *Weak-FTR* remains positive and highly significant (p-value < 0.001).

4.5 Time Preference or Reward-timing Uncertainty?

Earlier we outlined two economic channels, both predicting that weak FTR leads to higher R&D investment. The first channel is time preference, and the second channel is reward-timing uncertainty. Although both channels lead to the same prediction, it will be informative to examine the relative importance of these two channels. To this end, we conduct two tests.

First, we control for a measure of time preference, defined as the percentage of survey participants in each country who prefer a larger but later payment rather than a smaller but sooner payment (Wang et al., 2016). Intuitively, people with stronger time preference tend to discount future payoff less. We find that the correlation between the time preference measure and weak FTR is 0.6. If weak FTR mainly captures stronger time preference, controlling for another measure of time preference should reduce the significance of weak FTR. The results in Column (1) of Table 9 show that the coefficient of *Weak-FTR* actually becomes larger in magnitude and remains highly significant, suggesting that time preference is probably not the main channel at work.

Second, we compare the effect of FTR on R&D with the effect on another major type of long-term investment, capital expenditures (Capex). Although both types of investment are for the long term, Capex, relative to R&D, has much less reward-timing uncertainty. This is because any serious Capex decision involves a capital budgeting process that explicitly lays out the expected timing of future cash flows. This explicit forecasting process likely breaks down the timing ambiguity related to weak FTR. In contrast, the capital budgeting process for R&D investment is less explicit because payoff timing from R&D is inherently more uncertain. Therefore, R&D decision is more subjective, and the timing ambiguity imbedded in a language's FTR is more likely to be at work. If weak FTR mainly captures time preference, it should have a positive effect on both R&D and Capex. In contrast, if weak FTR mainly captures reward-timing uncertainty, it should have a more pronounced effect on R&D than on Capex.

Columns (2)-(3) of Table 9 reports the finding. *Weak-FTR* has no effect on Capex, but a significantly positive effect on the ratio of R&D over the sum of R&D and Capex. Thus, the effect of FTR on R&D seems to come through the channel of reward-timing uncertainty, not the channel of time preference. We acknowledge that to distinguish these two channels, we ideally want to observe objective measurements of time preference and timing uncertainty, and then estimate the effect of each. Unfortunately, data limitations do not allow us to do this. Thus, the evidence here should be viewed as suggestive rather than definitive. Separating the two channels remains a rewarding topic for future research.

5 Further Endogeneity and Robustness Tests

We have documented that weak FTR is positively associated with R&D investment. As already discussed, reverse causality running from R&D to FTR is unlikely. Therefore we have focused on addressing the omitted-variable concern through controlling for a large set of country- and firm-level characteristics as well as unobservable country effects through a correlated-random-effects model. In this section, we attempt to further address endogeneity issues using the instrumental variable approach and several within-country tests. We also provide a battery of additional robustness tests of both country-level and firm-level evidence.

5.1 2SLS

Galor et al. (2018) show that pre-1500CE crop return has a positive effect on the use of periphrastic future tense in a language. We thus use the historical Crop Return, which measures the average pre-1500CE annual crop yield, as an instrumental variable for language FTR in our

2SLS regressions.[11] We find that Crop Return has a significantly negative effect on *Weak-FTR*, as shown in the first-stage regressions of Table 10. The first-stage F-statistics are well above 10, suggesting strong IV. In the second-stage regressions, the instrumented *Weak-FTR* has a significantly positive effect on both country-level and firm-level R&D investment.[12] These 2SLS results confirm our baseline evidence and suggest a causal relation between language FTR and R&D investment at both the firm and country levels.

5.2 Within-country Tests

Next, we further address the omitted-variable concern by conducting within-country tests, which essentially control for all country-level effects, whether observable or unobservable.

5.2.1 The United States

Our tests so far assume that firm managers within a country speak the same language. In reality, managers in the same country may have different national origins and different home languages. For example, although all CEOs in the U.S. may speak English in their working environment, those from weak-FTR countries might be influenced by their native languages when making decisions. Boroditsky (2001) finds that "*one's native language appears to exert a strong influence over how one thinks about abstract domains like time*" (p. 18). As the U.S. is a highly internationalized country with a large proportion of immigrants, we are able to collect a sufficient sample of CEOs with non-U.S. nationalities, especially those from weak-FTR language countries.

We collect information on CEO nationality from *BoardEx*. We limit our sample period to 1991-2013 because observations from early years contain very few weak-FTR observations. We then exclude observations from industries that have no weak-FTR CEOs. This results in 19,946 observations. Among them, 229 are identified to have CEO nationalities with weak-FTR languages. Since the proportion of weak-FTR observations is low, we might be picking up other firm characteristics if we simply compare the two subsamples of weak- vs. strong-FTR CEOs. For this reason, we employ a propensity-score-matching approach.

[11] See Galor & Özak (2016) for more details on the construction of Crop Return.

[12] We thank the reviewer for suggesting this test.

We first estimate the probability or the propensity score of a firm having a weak-FTR CEO by regressing an indicator variable for whether there is a weak-FTR CEO on firm size, leverage, market-to-book ratio, profitability, tangibility, capital intensity, and industry and year dummies. We then match each observation with a weak-FTR CEO to an observation with a strong-FTR CEO based on the propensity score. We conduct the match without replacement, and the maximum difference in the propensity score allowed for a match is 1%. This results in a sample of 458 firm-years, in which 229 firm-years with weak-FTR CEOs are matched with 229 firm-years with strong-FTR CEOs.

Our matched sample satisfies three important validity criteria (see e.g., Fang et al., 2014). First, in Panel A of Table 11, we run a Logit model using the original sample and the matched sample. The post-match regression in Column (2) shows that all the control variables are statistically insignificant after matching. Second, as shown in Panel B, there is no difference in the propensity score between observations with weak-FTR CEOs and those with strong-FTR CEOs. Third, as shown in the lower part of Panel C, the differences in firm controls are mostly insignificant.

The first two rows in Panel C show that *R&D/Assets* and *R&D/Sales* of firms with weak-FTR CEOs are significantly higher than that of firms with strong-FTR CEOs. In Panel D, we run regressions based on the matched sample, and find consistent evidence that weak-FTR CEOs invest significantly more in R&D. This result confirms our earlier finding that weak FTR has a positive effect on R&D, and more importantly, this positive effect is not driven by an omitted unobservable country characteristic.

5.2.2 Belgium

Belgium has 11 provinces, and each province uses a dominant language of either Dutch or French. We create a dummy variable *Weak-FTR* equal to 1 if a firm is located in a Dutch-speaking province and 0 if a firm is located in a French-speaking province. We run the regression of Equation (9) without the country-level controls. The result, reported in Table 12, shows that *R&D/Assets* and *R&D/Sales* in weak-FTR regions are significantly higher than those in strong-FTR regions.[13]

5.2.3 Hong Kong: pre- and post-1997

An ideal experiment is to randomly assign weak- or strong-FTR languages to firms or countries and see how R&D investment is affected. This type of natural experiment did not occur in contemporary time.[14] The 1997 handover of Hong Kong from the U.K. to China is a quasi natural experiment in the sense that after 1997, Chinese (weak FTR) becomes a more dominant language over English (strong FTR).[15] We run a test to compare Hong Kong firms' R&D investment before and after 1997. We use other "Asian Tigers" as the control group. Consistent with our hypothesis, the test (tabulated in Appendix Table I) shows a significant increase in R&D investment by Hong Kong firms after 1997. A caveat is that besides the language effect, the 1997 event is accompanied by other confounding economic and political changes. Nevertheless, this quasi natural experiment result is consistent with our theoretical prediction and other empirical results.[16]

5.3 Further Robustness

Finally, we conduct a range of further robustness checks in Table 13. First, the classification of the langauge FTR of multilingual countries can be difficult especially when the people of a country speak both weak-FTR and strong-FTR languages. To address this issue, we drop the multilingual countries, including Belgium, Switzerland, and Singapore, from our sample and rerun the analysis. The results of both country-level and firm-level analyses are robust to

[13]Switzerland is another country where both strong- and weak-FTR languages are used. However, most Swiss are multilingual, and firm CEOs' first language is difficult to identify. Canada uses English and French, but both are strong-FTR languages.

[14]It may have occurred in ancient times such as after the Tower of Babel or after colonization by a foreign country, but the economic data are not available for those times.

[15] Chapter 1 Article 9 of the Hong Kong Basic Law, which serves as the constitutional document of Hong Kong after 1997, states that: *"In addition to the Chinese language, English may also be used as an official language by the executive authorities, legislature, and judiciary of the Hong Kong Special Administrative Region."* Chen et al. (2017) also employ the 1997 Hong Kong handover as a quasi-natural experiment.

[16]In Appendix Table II, we perform a Placebo test based on Canada with two strong-FTR official languages. Given that English is weaker than French in terms of the degree of language of FTR, we test the effect a Placebo Weak-FTR that takes the value of one if English is the official language in a province, and zero otherwise. As expected, the coefficient on the Placebo Weak-FTR is insignificant.

this subsample. In addition, to better capture the strength of language FTR of multilingual countries, we use the weighted FTR measure proposed by Chen (2013). In particular, the FTR values of multilingual countries are weighted by the proportions of their population speaking each language. The results of both country-level and firm-level analyses are robust to this alternative FTR measure.[17] Second, we check whether our results are driven by small economies. We exclude countries with relatively small economies – whose GDP are in the bottom quintile of the sample. *Weak-FTR* are still significantly positive. Finally, our findings are also robust to the inclusion of high-dimensional continent-by-year fixed effects that control for unobservable time-varying continent-specific heterogeneity.

The empirical tests present consistent evidence that weak FTR is associated with higher R&D investment. Various endogeneity checks suggest that the finding is unlikely to be driven by omitted correlated variables. A natural extension of the R&D test is to examine whether weak FTR also leads to more innovation output, such as measured by patent generation. We empirically test this in the following section.

6 Innovation Output

We obtain country-level patent-based innovation output measure from the Organization for Economic Cooperation and Development (OECD), namely Pat_{OECD}, defined as the number of triadic patents reported by the OECD. The OECD patent data are particularly useful for cross-country comparison of patents because the data cover triadic patents that are filed at three major patent offices in different countries and thus are much less subject to the home bias that domestic inventors tend to file only in their home country. After merging these patent measures with the control variables used in Table 2, we have data for 34 countries over the period 1996-2013. Table 14 reports the regression result of patent on *Weak-FTR*. In each regression we control for the full set of control variables used in Table 2 and continent fixed effects. The

[17] One may argue that the effect of language FTR might be weaker for countries with higher international exposure simply because people in these countries are exposed to foreign languages. We test this conjecture in Appendix Table III. We use International Migrant (i.e. international migrant stock (% of population)) as a proxy for a country's international exposure. We find a significantly negative coefficient on the interaction between Weak-FTR and International Migrant, suggesting that international exposure weakens the effect of language FTR on R&D investment.

results suggest that weak FTR is related to a higher level of innovation output, corroborating the evidence based on R&D investment.

The patent result is intriguing – not only does weak FTR lead to greater R&D investment, but this higher investment is also socially beneficial. R&D policy is at the center of economic policy because R&D is a key driver of innovation and economic growth, yet private R&D falls substantially below the socially optimal level (Jones & Williams, 1998; Brown et al., 2017). If aware of the dampening effect of strong-FTR on R&D investment, policy makers in those countries may start thinking about countervailing policy initiatives to promote R&D investment and bring innovation closer to the socially optimal level.

7 Discussion and Conclusion

Is language an economic institution that affects important resource-allocation decisions? Our evidence suggests yes. We outline two economic channels through which a language's ambiguous encoding of future timing (weak FTR) leads to higher R&D investment. Country-level, firm-level, and CEO-level tests all provide consistent evidence supporting the theoretical prediction. The empirical evidence survives various robustness and endogeneity checks. Consistent with the R&D evidence, we find that weak FTR is also associated with more innovation output as measured by patent generation.

It is an interesting finding that language FTR is not highly correlated with various culture attributes, and that the effect of FTR on R&D remains economically and statistically significant while controlling for these culture attributes. Extensive evidence has shown that culture is an important economic institution that influences a wide range of economic and finance decisions. Does language, like culture, exert similar influences? This study provides evidence towards an affirmative answer and suggests that language likely presents a fruitful area for future finance research.

A particular challenge for future research is to identify the economic channels that FTR or other language features affect decision making. We identify two economic channels through which FTR can affect R&D investment. Future research may need to provide a different explanation for another language feature or another economic outcome. To overcome this challenge, finance research will benefit from working with linguistics research to better understand how linguistic features can influence decision making. This interdisciplinary approach can potentially yield great rewards. Just as psychology and sociology (culture) have enriched our understanding of finance, language can potentially provide new vantage points.

Finally, we want to note the prospects for policy implications. There is a strong theoretical argument that because the benefit of R&D investment cannot be fully captured by the R&D investor, private R&D investment will fall below the socially optimal level, potentially by as much as 75% (Jones & Williams, 1998). As a result, significant attention has been given to how formal institutions, such as policies, can encourage R&D investment (Brown et al., 2017). We show that language, as an informal institution, significantly affects R&D and patent generation. Does language FTR impose a constraint on the making and effectiveness of R&D policies? And if so, in what ways? Similar policy-related questions have been raised in political science research (Pérez & Tavits, 2017), and the answers may have wide-ranging implications.

References

Acharya, V., Amihud, Y., & Litov, L. (2011). Creditor rights and corporate risk-taking. *Journal of Financial Economics*, *102*, 150–166.

Aggarwal, R., Faccio, M., Guedhami, O., & Kwok, C. C. (2016). Culture and finance: An introduction. *Journal of Corporate Finance*, (41), 466–474.

Ahern, K. R., Daminelli, D., & Fracassi, C. (2015). Lost in translation? the effect of cultural values on mergers around the world. *Journal of Financial Economics*, *117*, 165–189.

Allred, B. B., & Park, W. G. (2007). Patent rights and innovative activity: evidence from national and firm-level data. *Journal of International Business Studies*, *38*(6), 878–900.

Barker, V., & Mueller, G. (2002). CEO characteristics and firm R&D spending. *Management Science*, *48*, 782–801.

Bhattacharya, U., Hsu, P.-H., Tian, X., & Xu, Y. (2017). What affects innovation more: Policy or policy uncertainty? *Journal of Financial and Quantitative Analysis*, *52*(5), 1869–1901.

Boroditsky, L. (2001). Does language shape thought?: Mandarin and english speakers' conceptions of time. *Cognitive psychology*, *43*(1), 1–22.

Boroditsky, L. (2011). How language shapes thought. *Scientific American*, *304*(2), 62–65.

Brown, J., Fazzari, S., & Petersen, B. (2009). Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom. *Journal of Finance*, *64*, 151–185.

Brown, J. R., Martinsson, G., & Petersen, B. C. (2017). What promotes R&D? comparative evidence from around the world. *Research Policy*, *46*(2), 447–462.

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.

Chen, S., Cronqvist, H., Ni, S., & Zhang, F. (2017). Languages and corporate savings behavior. *Journal of Corporate Finance*, *46*, 320–341.

Cohen, L., Diether, K., & Malloy, C. (2013). Misvaluing innovation. *The Review of Financial Studies*, *26*(3), 635–666.

Croce, M. M., Nguyen, T. T., Raymond, S., & Schmid, L. (2019). Government debt and the returns to innovation. *Journal of Financial Economics*, *132*(3), 205–225.

Czarnitzki, D., & Toole, A. A. (2011). Patent protection, market uncertainty, and R&D investment. *The Review of Economics and Statistics*, *93*, 147–159.

Dahl, Ö. (2000). The grammar of future time reference in european languages. In O. Dahl (Ed.) *Tense and Aspect in the Languages of Europe*, (pp. 309–328). Berlin: Mouton de Gruyter.

Daske, H., Hail, L., Leuz, C., & Verdi, R. (2008). Mandatory IFRS reporting around the world: Early evidence on the economic consequences. *Journal of Accounting Research*, *46*, 1085–1142.

Dixit, A. K., & Pindyck, R. S. (1994). *Investment Under Uncertainty*. NJ, USA: Princeton University Press.

Djankov, S., La Porta, R., Lopez-de Silanes, F., & Shleifer, A. (2008). The law and economics of self-dealing. *Journal of Financial Economics*, *88*, 430–465.

Djankov, S., McLiesh, C., & Shleifer, A. (2007). Private credit in 129 countries. *Journal of Financial Economics*, *84*, 299–329.

Fang, V. W., Tian, X., & Tice, S. (2014). Does stock liquidity enhance or impede firm innovation? *Journal of Finance*, *69*, 2085–2125.

Fasan, M., Gotti, G., Kang, T., & Liu, Y. (2016). Language FTR and earnings management: International evidence. *Available at SSRN: https://ssrn.com/abstract=2763922*.

Galor, O., & Özak, Ö. (2016). The agricultural origins of time preference. *American Economic Review*, *106*(10), 3064–3103.

Galor, O., Özak, Ö., & Sarid, A. (2018). Geographical origins of language structures. *Available at SSRN 3097220*.

Gu, L. (2016). Product market competition, R&D investment, and stock returns. *Journal of Financial Economics*, *119*(2), 441–455.

Guin, B. (2016). Culture and household saving. *Available at SSRN: https://ssrn.com/abstract=2698872*.

Guiso, L., Sapienza, P., & Zingales, L. (2008). Trusting the stock market. *Journal of Finance*, *63*(6), 2557–2600.

Halek, M., & Eisenhauer, J. G. (2001). Demography of risk aversion. *Journal of Risk and Insurance*, (pp. 1–24).

Helliwell, J. F. (1996). Economic growth and social capital in Asia. Tech. rep., National Bureau of Economic Research.

Hirshleifer, D., Low, A., & Teoh, S. H. (2012). Are overconfident CEOs better innovators? *Journal of Finance*, *67*, 1457–1498.

Hitt, M., Hoskisson, R., Ireland, D., & Harrison, J. (1991). Effects of acquisitions on R&D inputs and outputs. *Academy of Management Journal*, *34*, 693–706.

Hofstede, G. (1984). *Culture's Consequences: International Differences in Work-Related Values*, vol. 5. CA: Sage.

Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations*. CA: Sage.

Hsu, P.-H., Tian, X., & Xu, Y. (2014). Financial development and innovation: Cross-country evidence. *Journal of Financial Economics*, *112*, 116–135.

Jones, C. I., & Williams, J. C. (1998). Measuring the social return to R&D. *The Quarterly Journal of Economics*, *113*(4), 1119–1135.

Karolyi, G. A. (2016). The gravity of culture for finance. *Journal of Corporate Finance*, *41*, 610–625.

Kim, J., Kim, Y., & Zhou, J. (2017). Languages and earnings management. *Journal of Accounting and Economics*, *63*, 288–306.

Koh, P.-S., & Reeb, D. (2015). Missing R&D. *Journal of Accounting and Economics*, *60*, 73–94.

La Porta, R., Lopez-de Silanes, F., & Shleifer, A. (2008). The economic consequences of legal origins. *Journal of Economic Literature*, *46*, 285–332.

Li, K., Griffin, D., Yue, H., & Zhao, L. (2013). How does culture influence corporate risk-taking? *Journal of Corporate Finance*, *23*, 1–22.

Luong, H., Moshirian, F., Nguyen, L., Tian, X., & Zhang, B. (2017). How do foreign institutional investors enhance firm innovation? *Journal of Financial and Quantitative Analysis*, *52*(4), 1449–1490.

Maskus, K., Neumann, R., & Seidel, T. (2012). How national and international financial development affect industrial R&D. *European Economic Review*, *56*, 72–83.

Mavisakalyan, A., & Weber, C. (2017). Linguistic structures and economic outcomes. *Journal of*

Economic Surveys, *forthcoming*, 1–24.

Mill, J. S. (1884). *A system of logic, ratiocinative and inductive: Being a connected view of the principles of evidence and the methods of scientific investigation*, vol. 1. Longmans, Green, and Company.

North, D. C. (1991). Institutions. *Journal of Economic Perspectives*, *5*(1), 97–112.

Park, W. G. (2008). International patent protection: 1960–2005. *Research Policy*, *37*, 761–766.

Paule-Paludkiewicz, H., Fuchs-Schündeln, N., & Masella, P. (2016). Cultural determinants of household saving behavior. Tech. rep., Verein für Socialpolitik/German Economic Association.

Pérez, E. O., & Tavits, M. (2017). Language shapes people's time perspective and support for future-oriented policies. *American Journal of Political Science*, *61*(3), 715–727.

Roberts, S. G., Winters, J., & Chen, K. (2015). Future tense and economic decisions: Controlling for cultural evolution. *PloS One*, *10*(7), 1–46.

Roth, F. (2009). Does too much trust hamper economic growth? *Kyklos*, *62*(1), 103–128.

Rothschild, M., & Stiglitz, J. (1970). Increasing risk: I. a definition. *Journal of Economic Theory*, *2*, 225–243.

Ryan, H., & Wiggins, R. (2002). The interactions between R&D investment decisions and compensation policy. *Financial Management*, *31*, 5–29.

Shao, L., Kwok, C. C., & Zhang, R. (2013). National culture and corporate investment. *Journal of International Business Studies*, *44*, 745–763.

Stulz, R. M., & Williamson, R. (2003). Culture, openness, and finance. *Journal of Financial Economics*, *70*, 313–349.

Varsakelis, N. (2001). The impact of patent protection, economy openness and national culture on R&D investment: a cross-country empirical investigation. *Research Policy*, *30*, 1059–1068.

Wang, M., Rieger, M. O., & Hens, T. (2016). How time preferences differ: Evidence from 53 countries. *Journal of Economic Psychology*, *52*, 115–135.

Whorf, B. (1956). *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. Edited by John B. Carroll. Cambridge, MA: MIT Press.

Williamson, O. E. (2000). The new institutional economics: Taking stock, looking ahead. *Journal of Economic Literature*, *38*(3), 595–613.

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(3), 253–265.

Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach, Fifth Edition*. South-Western Cengage Learning.

Appendix I: Country Language FTR and Business R&D/GDP

This table presents language FTR (weak vs. strong) and average Business R&D expenditures as a percentage of GDP (R&D/GDP) of 52 countries over the period 1996-2013.

Nation	Language FTR	R&D/GDP	Nation	Language FTR	R&D/GDP
Argentina	strong	0.116	Korea (South)	strong	2.126
Australia	strong	1.068	Luxembourg	weak	1.214
Austria	weak	1.679	Malaysia	weak	0.520
Belgium	weak	1.348	Mexico	strong	0.139
Bermuda	strong	0.015	Morocco	strong	0.180
Bulgaria	strong	0.172	Netherlands	weak	0.968
Canada	strong	1.038	New Zealand	strong	0.443
Chile	strong	0.118	Norway	weak	0.867
China	weak	0.850	Peru	strong	0.015
Colombia	strong	0.045	Philippines	strong	0.078
Croatia	strong	0.367	Poland	strong	0.222
Cyprus	strong	0.068	Portugal	strong	0.400
Czech Republic	strong	0.735	Romania	strong	0.249
Denmark	weak	1.684	Russian Federation	strong	0.730
Finland	weak	2.281	Singapore	strong	1.268
France	strong	1.339	South Africa	strong	0.424
Germany	weak	1.710	Spain	strong	0.571
Greece	strong	0.185	Sri Lanka	strong	0.028
Hong Kong	weak	0.269	Sweden	weak	2.474
Hungary	strong	0.481	Switzerland	weak	1.896
India	strong	0.214	Thailand	strong	0.096
Indonesia	weak	0.018	Turkey	strong	0.238
Ireland	strong	0.885	United Arab.	strong	0.140
Israel	strong	3.071	United Kingdom	strong	1.071
Italy	strong	0.562	United States	strong	1.845
Japan	weak	2.393	Vietnam	strong	0.040

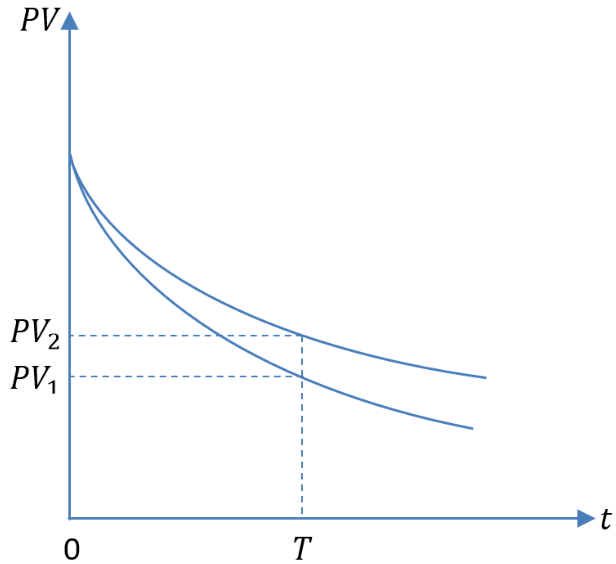
Appendix II: Variable Definitions

Variable	Definition	Source
Panel A: Dependent variables		
Business R&D/GDP	Country-level: domestic expenditures on R&D by business enterprises divided by GDP	UNESCO
Total R&D/GDP	Country-level: total domestic expenditures on R&D divided by GDP	UNESCO
Pat_{OECD}	Country-level: the number of triadic patents filed at the European, Japanese, and U.S. patent offices (scaled by 1000)	OECD
R&D/Assets	Firm-level: R&D expenditures, divided by total assets	Worldscope
R&D/Sales	Firm-level: R&D expenditures, divided by sales	Worldscope
Capex/Assets	Firm-level: Capital expenditures, divided by total assets	Worldscope
Panel B: Language-based measures of perceived reward-timing uncertainty		
Weak-FTR	Dummy=1 if the official language of a nation is classified as a weak future-time reference (FTR) language	Chen (2013)
Verb Ratio	In the country's weather forecast: the number of verbs that are grammatically future-marked, divided by the total number of future-referring verbs	Chen (2013)
Sentence Ratio	In the country's weather forecast: the proportion of sentences regarding the future which contain a grammatical future marker	Chen (2013)
Panel C: Firm-level control variables		
Firm Size	The logarithm of total assets in U.S. dollars	Worldscope
Leverage	Total debt divided by total assets	Worldscope
Market-to-Book	Firm market value (i.e., total assets - common equity + market capitalization), divided by total assets	Worldscope
Profitability	EBITDA divided by total assets	Worldscope
Tangibility	PP&E divided by total assets	Worldscope
Cash Ratio	Cash divided by total assets	Worldscope
Capital Intensity	The logarithm of the ratio of total assets in U.S. dollars to the number of employees	Worldscope
CF Volatility	The industry-average standard deviation of firm profitability in the past five years, using DataStream Level 3 industry classification.	Worldscope
R&D-missing	Dummy=1 if R&D observation is missing	Worldscope
IFRS	Dummy=1 if a firm follows the International Financial Reporting Standards (IFRS), and 0 otherwise	Worldscope
Panel D: Country-level control variables		
GDP per capita	The logarithm of GDP in constant 2005 U.S. dollars, divided by midyear population	World Bank
Stock Market	Market capitalization of listed domestic companies, as a percentage of GDP (scaled by 100)	World Bank
Credit Market	Domestic credit to private sectors, as a percentage of GDP (scaled by 100)	World Bank

(continued from the previous page)

UK Legal Origin	Dummy=1 for UK legal origin	La Porta et al. (2008)
French L. O.	Dummy=1 for French legal origin	La Porta et al. (2008)
German L. O.	Dummy=1 for German legal origin	La Porta et al. (2008)
Scand. L. O.	Dummy=1 for Scandinavian legal origin	La Porta et al. (2008)
Shareholder Rights	The anti-self-dealing index, a measure of shareholder protection against expropriation by firm directors	Djankov et al. (2008)
Creditor Rights	Creditor rights aggregate score between 0 and 4, measuring the legal rights of creditors against defaulting debtors	Djankov et al. (2007)
Patent Rights	An index of patent protection, considering five aspects of patents: coverage; membership in international treaties; duration of protection; enforcement mechanisms; and restrictions	Park (2008)
Catholic	Dummy=1 if the largest proportion of a country's population practices catholic religion	La Porta et al. (2008)
Individualism	National culture index related to the level of individualism in a society (scaled by 100)	Hofstede Website
Uncertainty Avo.	National culture index related to the level of uncertainty avoidance in a society (scaled by 100)	Hofstede Website
Power Distance	National culture index related to power distance between different members of a society (scaled by 100)	Hofstede Website
Masculinity	National culture index related to the level of masculinity in a society (scaled by 100)	Hofstede Website
Long-term Ort.	National culture index related to the long-term orientation of a society (the score is constructed based on the World Value Survey) (scaled by 100)	Hofstede Website
Trust	The percentage of survey respondents who believe that most people can be trusted	World Value Survey
Time Preference	The percentage of survey participants in each country who prefer option B in the following binary choice question: "Which offer would you prefer? A. a payment of $3400 this month; B. a payment of $3800 next month."	Wang et al. (2016)

Panel A



Panel B

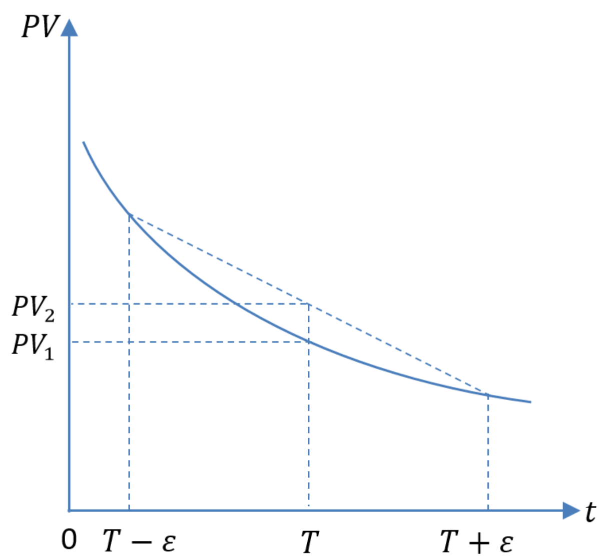


Figure 1: The effects of future timing perception on valuation

The curves in both panels represent the perceived present value (PV) of a future reward over time t. In Panel A, we show how "time preference" affects valuation over the same future reward, assuming that there is no timing uncertainty. A speaker of weak-FTR language perceives the future to be closer to the present and discounts a future reward less, so her valuation (PV_2) is higher. In Panel B, we show how "timing uncertainty" affects the valuation of the same future reward, assuming the same discount rate. A speaker of weak-FTR language perceives greater timing uncertainty, so her valuation (PV_2) is higher.

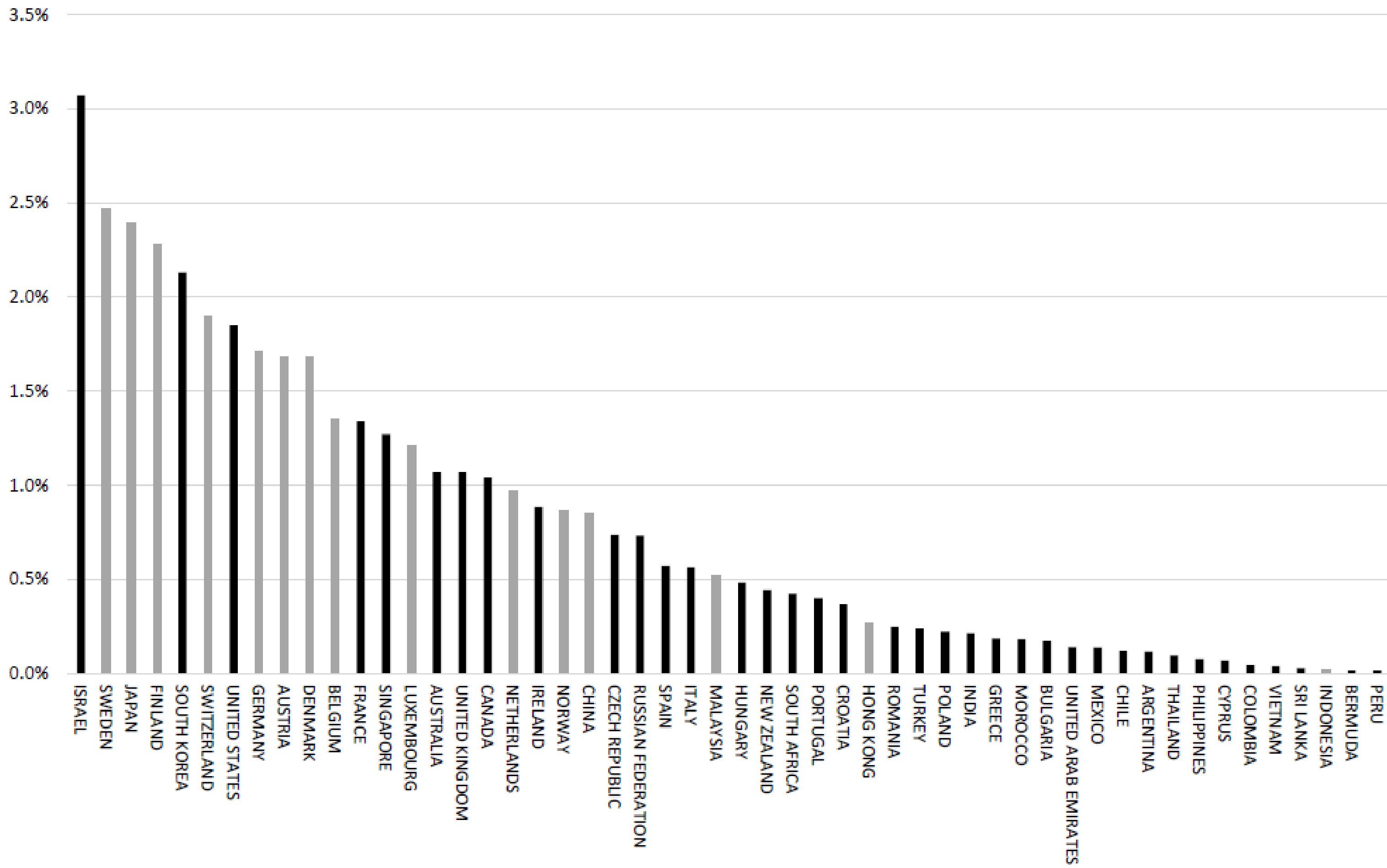


Figure 2: Mean Business R&D expenditures as a percentage of GDP (52 countries, 1996-2013)

The figure shows the mean Business R&D expenditures as a percentage of GDP in different countries. Light-colored bars represent weak-FTR countries, while dark-colored bars represent strong-FTR countries. The classification of weak-FTR and strong-FTR languages is based on Chen (2013). All variables are defined in Appendix II.

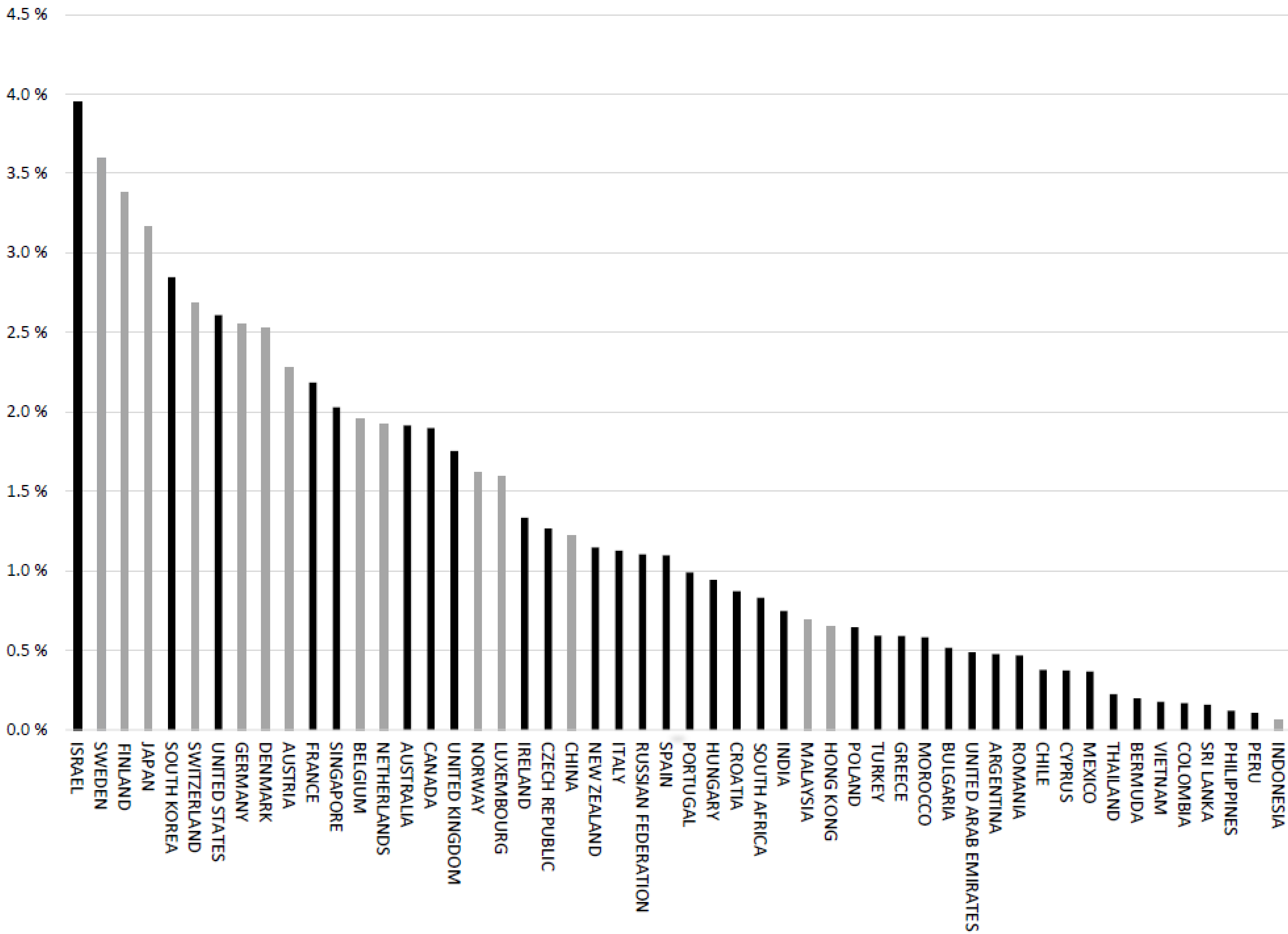


Figure 3: Mean Total R&D expenditures as a percentage of GDP (52 countries, 1996-2013)

The figure shows the mean Total R&D expenditures as a percentage of GDP in different countries. Light-colored bars represent weak-FTR countries, while dark-colored bars represent strong-FTR countries. The classification of weak-FTR and strong-FTR languages is based on Chen (2013). All variables are defined in Appendix II.

Table 1: Country-level Summary Statistics

This table presents summary statistics for the main variables used in our country-level analysis. Our sample comprises over 400 country-year observations from 39 countries in the 1996-2013 period. All variables are defined in Appendix II.

	N	Mean	Median	SD	Min	Max
Business R&D/GDP	435	0.010	0.007	0.009	0.000	0.038
Total R&D/GDP	459	0.015	0.012	0.010	0.001	0.045
Weak-FTR	459	0.292	0.000	0.455	0.000	1.000
Verb Ratio	386	0.519	0.716	0.371	0.000	1.000
Sentence Ratio	386	0.559	0.741	0.390	0.000	1.000
GDP per capita	459	9.455	9.908	1.142	6.399	11.125
Stock Market	459	0.824	0.580	0.942	0.004	6.060
Credit Market	459	0.907	0.910	0.541	0.088	2.036
Individualism	459	0.493	0.480	0.245	0.130	0.910
Uncertainty Avo.	459	0.657	0.750	0.230	0.080	0.950
Power Distance	459	0.564	0.570	0.193	0.130	1.040
Masculinity	459	0.524	0.520	0.182	0.050	0.950
Long-term Ort.	459	0.516	0.499	0.226	0.131	1.000
Trust	459	0.303	0.281	0.142	0.056	0.693
UK Legal Origin	459	0.329	0.000	0.470	0.000	1.000
French L. O.	459	0.342	0.000	0.475	0.000	1.000
German L. O.	459	0.281	0.000	0.450	0.000	1.000
Shareholder Rights	459	0.530	0.461	0.244	0.178	1.000
Creditor Rights	459	1.911	2.000	1.150	0.000	4.000
Patent Rights	459	3.916	4.000	0.706	1.380	4.880
Catholic	459	0.401	0.000	0.491	0.000	1.000

Table 2: Correlations among R&D and Formal and Informal Institutions

This table presents correlation coefficients between country-level R&D and country-specific language, economic, cultural, legal, and religion variables. All variables are defined in Appendix II.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
(1) Business R&D/GDP																				
(2) Total R&D/GDP	0.99																			
(3) Weak-FTR	0.28	0.26																		
(4) Verb Ratio	-0.09	-0.08	-0.88																	
(5) Sentence Ratio	-0.08	-0.07	-0.90	0.99																
(6) GDP per capita	0.60	0.67	0.22	-0.09	-0.08															
(7) Stock Market	0.10	0.11	0.30	-0.23	-0.23	0.34														
(8) Credit Market	0.48	0.51	0.33	-0.17	-0.17	0.61	0.45													
(9) Individualism	0.29	0.41	0.01	0.03	0.03	0.64	-0.05	0.32												
(10) Uncertainty Avo.	-0.07	-0.10	-0.29	0.26	0.25	-0.17	-0.54	-0.36	-0.08											
(11) Power Distance	-0.53	-0.57	-0.05	-0.03	-0.03	-0.58	0.04	-0.27	-0.66	0.01										
(12) Masculinity	0.03	0.01	-0.07	-0.06	-0.06	-0.04	-0.03	0.07	0.08	0.13	0.08									
(13) Long-term Ort.	0.41	0.41	0.36	-0.33	-0.31	0.21	0.10	0.33	-0.11	-0.05	0.08	0.06								
(14) Trust	0.32	0.39	0.46	-0.44	-0.43	0.45	0.24	0.46	0.38	-0.44	-0.46	-0.22	0.26							
(15) UK Legal Origin	0.13	0.10	-0.18	0.31	0.31	0.06	0.44	0.32	0.10	-0.63	-0.16	0.04	-0.25	0.12						
(16) French L. O.	-0.41	-0.39	-0.12	0.16	0.13	-0.16	-0.21	-0.31	-0.14	0.40	0.26	-0.20	-0.30	-0.42	-0.50					
(17) German L. O.	0.19	0.19	0.15	-0.30	-0.26	-0.00	-0.24	0.03	-0.04	0.31	0.03	0.40	0.63	0.08	-0.44	-0.45				
(18) Shareholder Rights	0.03	0.03	-0.06	0.28	0.27	0.02	0.48	0.34	-0.18	-0.62	0.05	-0.01	-0.05	0.12	0.76	-0.46	-0.26			
(19) Creditor Rights	0.22	0.25	0.23	-0.13	-0.12	0.31	0.40	0.43	0.07	-0.45	-0.24	-0.10	0.43	0.35	0.41	-0.46	0.10	0.53		
(20) Patent Rights	0.55	0.62	0.17	-0.11	-0.09	0.77	0.18	0.54	0.60	-0.11	-0.37	0.06	0.31	0.33	-0.12	-0.07	0.15	-0.08	0.16	
(21) Catholic	-0.39	-0.35	-0.18	0.08	0.06	-0.04	-0.20	-0.25	0.12	0.41	0.11	0.03	-0.36	-0.35	-0.45	0.68	-0.17	-0.57	-0.61	0.03

Table 3: Weak Future-Time Reference (FTR) and National R&D investment

This table presents evidence on how language FTR affects national R&D investment. Our sample comprises country-year observations from 39 countries in the 1996-2013 period. We use two alternative measures of national R&D investment: Business R&D/GDP in Columns (1)-(4) and Total R&D/GDP in Columns (5)-(8). Business R&D/GDP is defined as domestic expenditures on R&D by business enterprises as a percentage of GDP. Total R&D/GDP is defined as total domestic expenditures on R&D during a given year as a percentage of GDP. The measure of language future-time reference (FTR) is the dummy variable, *Weak-FTR*, which is equal to one if a country's dominant language is identified as a weak-FTR language based on Chen (2013). All columns use OLS regressions. Year and continent fixed effects are denoted as Year FE and Continent FE. Standard errors are clustered at the country level and corrected for heteroscedasticity. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. An intercept is included but not reported. All variables are defined in Appendix II.

(*continuing on the next page...*)

Dependent Var.	Business R&D/GDP				Total R&D/GDP			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Weak-FTR	0.002***	0.002**	0.003***	0.002***	0.003***	0.004***	0.004***	0.002***
	(3.17)	(2.38)	(4.64)	(3.34)	(3.61)	(3.87)	(6.15)	(3.59)
GDP per capita	0.004***	0.003***	0.002***	0.002***	0.005***	0.004***	0.003***	0.004***
	(12.30)	(8.09)	(4.46)	(6.09)	(13.83)	(8.81)	(7.99)	(10.31)
Stock Market	-0.002***	-0.002***	-0.001***	-0.001***	-0.003***	-0.003***	-0.002***	-0.001***
	(-6.35)	(-5.15)	(-5.61)	(-3.71)	(-7.41)	(-6.64)	(-7.07)	(-4.45)
Credit Market	0.003***	0.002**	0.001**	0.001**	0.004***	0.002*	0.000	0.001
	(4.68)	(2.10)	(2.15)	(2.09)	(4.78)	(1.67)	(0.57)	(0.99)
UK Legal Origin		-0.000	0.007***	0.004		0.002	0.007***	0.003
		(-0.04)	(3.65)	(1.60)		(0.78)	(3.39)	(1.28)
French L. O.		-0.007***	-0.006***	-0.005**		-0.008***	-0.004**	-0.005**
		(-3.45)	(-3.57)	(-2.41)		(-3.63)	(-2.25)	(-2.23)
German L. O.		-0.001	-0.007***	-0.002		-0.002	-0.004**	-0.000
		(-0.60)	(-3.78)	(-1.07)		(-0.85)	(-2.05)	(-0.12)
Shareholder Rights		-0.000	-0.010***	-0.006***		-0.004***	-0.006***	-0.003*
		(-0.23)	(-5.56)	(-3.92)		(-2.83)	(-2.93)	(-1.66)
Creditor Rights		-0.000	-0.003***	-0.004***		-0.000	-0.003***	-0.004***
		(-1.54)	(-13.39)	(-13.28)		(-1.17)	(-11.28)	(-13.68)
Patent Rights		0.003***	0.004***	0.003***		0.004***	0.006***	0.004***
		(6.19)	(7.19)	(5.42)		(7.30)	(8.32)	(6.70)
Individualism			-0.015***	-0.000			-0.012***	0.010***
			(-9.40)	(-0.00)			(-6.38)	(4.10)
Uncertainty Avo.			0.004**	0.002			0.004**	0.001
			(2.54)	(0.82)			(2.27)	(0.55)
Power Distance			-0.025***	-0.022***			-0.023***	-0.020***
			(-11.28)	(-10.09)			(-10.20)	(-9.11)
Masculinity			0.001	-0.006***			-0.001	-0.011***
			(1.01)	(-4.16)			(-0.66)	(-6.91)
Long-term Ort.			0.019***	0.016***			0.018***	0.018***
			(13.16)	(6.43)			(10.30)	(7.39)
Trust			-0.014***	-0.016***			-0.012***	-0.016***
			(-5.11)	(-5.08)			(-5.01)	(-5.80)
Catholic			-0.005***	-0.005***			-0.005***	-0.005***
			(-7.34)	(-6.37)			(-6.90)	(-6.02)
Continent FE	No	No	No	Yes	No	No	No	Yes
Year FE	No	No	No	Yes	No	No	No	Yes
N	435	435	435	435	459	459	459	459
adj. R^2	0.418	0.520	0.798	0.824	0.510	0.617	0.812	0.858

Table 4: Firm-level Summary Statistics

This table presents firm-level summary statistics for the main variables used in our study. All variables are defined in Appendix II.

	N	Mean	Median	SD	Min	Max
R&D/Assets	215,877	0.025	0.000	0.066	0.000	0.433
R&D/Sales	215,877	0.042	0.000	0.134	0.000	0.826
Weak-FTR	215,877	0.349	0.000	0.477	0.000	1.000
Verb Ratio	191,286	0.541	0.769	0.390	0.000	1.000
Sentence Ratio	191,286	0.599	0.875	0.428	0.000	1.000
GDP per capita	215,877	10.024	10.431	1.013	6.519	10.880
Stock Market	215,877	1.178	0.923	1.525	0.143	10.769
Credit Market	215,877	1.347	1.350	0.484	0.218	2.193
Individualism	215,877	0.628	0.710	0.279	0.140	0.910
Uncertainty Avo.	215,877	0.567	0.460	0.226	0.080	0.950
Power Distance	215,877	0.510	0.400	0.171	0.130	1.040
Masculinity	215,877	0.614	0.620	0.180	0.050	0.950
Long-term Ort.	215,877	0.551	0.511	0.259	0.212	1.000
Trust	215,877	0.354	0.372	0.100	0.056	0.693
UK Legal Origin	215,877	0.557	1.000	0.497	0.000	1.000
French L. O.	215,877	0.116	0.000	0.321	0.000	1.000
German L. O.	215,877	0.308	0.000	0.462	0.000	1.000
Shareholder Rights	215,877	0.625	0.651	0.209	0.178	1.000
Creditor Rights	215,877	2.098	2.000	1.164	0.000	4.000
Patent Rights	215,877	4.333	4.540	0.581	1.020	4.880
Catholic	215,877	0.128	0.000	0.334	0.000	1.000
Firm Size	215,877	12.213	12.193	2.096	6.258	16.961
Tangibility	215,877	0.297	0.258	0.220	0.003	0.913
Profitability	215,877	0.036	0.098	0.335	-2.358	0.462
Leverage	215,877	0.236	0.189	0.245	0.000	1.565
Market-to-Book	215,877	1.906	1.267	2.141	0.456	16.120
Capital Intensity	215,877	5.418	5.371	1.173	2.641	8.756
Cash Ratio	215,877	0.109	0.062	0.133	0.000	0.703
CF Volatility	215,877	0.778	0.285	1.604	0.021	9.112

Table 5: Weak Future-Time Reference (FTR) and Firm-level R&D

This table presents evidence on how FTR affects firm R&D investment. The dependent variable is R&D/Assets in Columns (1)-(3) and R&D/Sales in Columns (4)-(6). R&D/Assets is defined as R&D expenditures divided by book assets. R&D/Sales is defined as R&D expenditures divided by annual sales. *Weak-FTR* is equal to one if a country's dominant language is identified as a weak-FTR language, based on Chen (2013). All columns use OLS regressions. We control for a large set of country and firm characteristics, which are all defined in Appendix II. Year, industry, and continent fixed effects are denoted as Year FE, Industry FE, and Continent FE. Standard errors are clustered at the firm level and corrected for heteroscedasticity. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level. An intercept is included but not reported.

(*continuing on the next page...*)

Dependent Var.	R&D/Assets			R&D/Sales		
	(1)	(2)	(3)	(4)	(5)	(6)
Weak-FTR	0.005***	0.009***	0.008***	0.003*	0.017***	0.013***
	(5.04)	(10.49)	(8.81)	(1.66)	(9.61)	(7.36)
Firm Size		-0.000***	-0.000***		-0.001***	-0.001***
		(-3.09)	(-2.79)		(-5.28)	(-4.64)
Tangibility		-0.001	-0.001		-0.005**	-0.006**
		(-1.16)	(-1.23)		(-2.31)	(-2.52)
Profitability		-0.047***	-0.047***		-0.111***	-0.110***
		(-30.17)	(-29.96)		(-36.95)	(-36.72)
Leverage		-0.015***	-0.015***		-0.042***	-0.042***
		(-10.20)	(-10.17)		(-15.00)	(-14.95)
Market-to-Book		0.002***	0.002***		0.003***	0.003***
		(11.52)	(11.38)		(8.61)	(8.35)
Capital Intensity		0.003***	0.003***		0.018***	0.019***
		(10.80)	(10.61)		(29.95)	(29.99)
Cash Ratio		0.047***	0.047***		0.128***	0.130***
		(17.84)	(17.93)		(21.89)	(22.14)
CF Volatility		-0.001***	-0.001***		-0.001***	-0.001***
		(-6.11)	(-6.41)		(-6.65)	(-4.40)
R&D-missing	-0.045***	-0.034***	-0.034***	-0.075***	-0.054***	-0.055***
	(-70.63)	(-65.31)	(-65.54)	(-56.22)	(-52.06)	(-52.32)
GDP per capita	-0.000	0.000	0.001**	-0.005***	-0.003***	-0.002*
	(-0.02)	(0.44)	(2.07)	(-3.91)	(-2.82)	(-1.83)
Stock Market	-0.000	-0.001***	-0.001***	0.000	-0.003***	-0.002***
	(-1.40)	(-10.10)	(-8.34)	(0.48)	(-9.26)	(-7.31)
Credit Market	0.016***	0.003***	0.001*	0.039***	-0.002	-0.004**
	(15.41)	(3.09)	(1.71)	(17.77)	(-1.42)	(-2.18)
UK Legal Origin	0.022***	0.013***	0.013***	0.025***	0.033***	0.035***
	(6.53)	(4.45)	(4.28)	(3.68)	(5.35)	(5.67)
French L. O.	0.026***	0.011***	0.013***	0.038***	0.022***	0.029***
	(7.00)	(3.59)	(4.13)	(4.98)	(3.56)	(4.57)
German L. O.	0.018***	0.002	0.005	0.020**	0.004	0.013*
	(4.22)	(0.58)	(1.51)	(2.27)	(0.50)	(1.79)
Shareholder Rights	-0.006**	-0.001	-0.000	-0.004	0.009*	0.008
	(-2.14)	(-0.45)	(-0.01)	(-0.81)	(1.90)	(1.56)
Creditor Rights	-0.006***	-0.003***	-0.003***	-0.009***	-0.004***	-0.005***
	(-10.81)	(-6.04)	(-6.04)	(-7.63)	(-4.28)	(-4.36)
Patent Rights	0.001	-0.005***	-0.007***	0.003***	-0.010***	-0.013***
	(1.11)	(-8.63)	(-9.43)	(2.64)	(-9.58)	(-9.20)
Individualism	-0.032***	-0.016***	-0.006	-0.065***	-0.023***	-0.008
	(-8.41)	(-4.52)	(-1.61)	(-8.43)	(-3.11)	(-0.98)
Uncertainty Avo.	-0.022***	-0.003	-0.005**	-0.043***	-0.002	-0.006
	(-9.54)	(-1.56)	(-2.35)	(-9.41)	(-0.36)	(-1.41)
Power Distance	-0.043***	-0.029***	-0.025***	-0.085***	-0.047***	-0.037***
	(-9.14)	(-7.72)	(-6.45)	(-8.52)	(-5.63)	(-4.42)
Masculinity	-0.014***	-0.002	-0.003	-0.025***	-0.008	-0.010*
	(-5.18)	(-0.80)	(-1.21)	(-4.25)	(-1.55)	(-1.95)
Long-term Orientation	-0.030***	-0.011***	-0.011***	-0.053***	-0.016**	-0.018***
	(-8.44)	(-3.76)	(-3.47)	(-7.19)	(-2.46)	(-2.72)
Trust	-0.031***	-0.023***	-0.020***	-0.068***	-0.032***	-0.020***
	(-7.92)	(-6.55)	(-5.64)	(-8.53)	(-4.37)	(-2.77)
Catholic	-0.014***	-0.003**	-0.004**	-0.018***	0.000	-0.002
	(-7.14)	(-2.06)	(-2.48)	(-4.52)	(0.10)	(-0.64)
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Yes	No	Yes	Yes
Year FE	No	No	Yes	No	No	Yes
N	215,877	215,877	215,877	215,877	215,877	215,877
adj. R^2	0.204	0.386	0.387	0.150	0.377	0.378

Table 6: Alternative Measures of FTR and Firm-level R&D

This table presents evidence on how alternative measure of FTR affects firm R&D investment. The dependent variable is R&D/Assets in Columns (1)-(2) and R&D/Sales in Columns (3)-(4). R&D/Assets is defined as R&D expenditures divided by book assets. R&D/Sales is defined as R&D expenditures divided by annual sales. The alternative FTR measures are Verb Ratio and Sentence Ratio, proposed by Chen (2013). Verb Ratio is the number of verbs in a country's weather forecast that are grammatically future-marked divided by the total number of future-referring verbs. Sentence Ratio is the proportion of sentences in a country's weather forecast regarding the future which contains a grammatical future marker. All columns use OLS regressions. Country controls are the full set of country-level control variables in Table 5. Firm controls are the full set of firm-level control variables in Table 5. Year, industry, and continent fixed effects are denoted as Year FE, Industry FE, and Continent FE. Standard errors are clustered at the firm level and corrected for heteroscedasticity. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level. All variables are defined in Appendix II.

Dependent Var.	R&D/Assets		R&D/Sales	
	(1)	(2)	(3)	(4)
Verb Ratio	-0.004***		-0.011***	
	(-2.88)		(-3.94)	
Sentence Ratio		-0.004***		-0.011***
		(-3.09)		(-4.16)
Firm Controls	Yes	Yes	Yes	Yes
Country Controls	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	191,286	191,286	191,286	191,286
adj. R^2	0.390	0.390	0.384	0.384

Table 7: FTR and Firm-level R&D: Correlated Random Effects

This table presents correlated-random-effects estimation results. The dependent variable is R&D/Assets in Column (1) and R&D/Sales in Column (2). R&D/Assets is defined as R&D expenditures divided by book assets. R&D/Sales is defined as R&D expenditures divided by annual sales. *Weak-FTR* is equal to one if a country's dominant language is identified as a weak-FTR language, based on Chen (2013). We control for a large set of country and firm characteristics, which are all defined in Appendix II. To estimate the correlated-random-effects model, in addition to the full set of control variables in Table 5, we also include the averages of the country-level and firm-level controls which are reported in this table. Year, industry, and continent fixed effects are denoted as Year FE, Industry FE, and Continent FE. Standard errors are clustered at the firm level and corrected for heteroscedasticity. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level. An intercept is included but not reported.

Dependent Var.	R&D/Assets	R&D/Sales
	(1)	(2)
Weak-FTR	0.010***	0.015***
	(7.62)	(5.19)
Avg_GDP per capita	-0.014***	-0.015***
	(-8.62)	(-4.58)
Avg_Stock Market	-0.002***	-0.003***
	(-4.09)	(-3.13)
Avg_Credit Market	-0.011***	0.003
	(-6.61)	(0.88)
Avg_Firm Size	0.006***	0.002***
	(24.42)	(4.68)
Avg_Tangibility	-0.015***	-0.023***
	(-8.10)	(-5.89)
Avg_Profitability	-0.023***	-0.077***
	(-18.88)	(-29.90)
Avg_Leverage	-0.021***	-0.045***
	(-13.77)	(-14.04)
Avg_Market-to-Book	0.001***	0.001***
	(4.76)	(3.37)
Avg_Capital Intensity	0.010***	0.015***
	(27.58)	(19.09)
Avg_Cash Ratio	0.066***	0.161***
	(22.30)	(25.83)
Avg_CF Volatility	-0.003***	-0.004***
	(-9.45)	(-5.76)
Firm Controls	Yes	Yes
Country Controls	Yes	Yes
Continent FE	Yes	Yes
Industry FE	Yes	Yes
Year FE	Yes	Yes
N	215,877	215,877
adj. R^2	0.398	0.392

Table 8: Firm-level Robustness Tests

This table presents various firm-level robustness tests. Columns (1)-(2) exclude US firms. Columns (3)-(8) are based on three sub-periods. Columns (9)-(10) exclude observations with missing R&D. Columns (11)-(12) control for IFRS. Columns (13)-(14) use the Tobit estimator. Column (15)-(16) control for lagged R&D. The dependent variable is R&D/Assets in odd-numbered columns, and R&D/Sales in even-numbered columns. R&D/Assets is defined as R&D expenditures divided by book assets. R&D/Sales is defined as R&D expenditures divided by annual sales. *Weak-FTR* is equal to one if a country's dominant language is identified as a weak-FTR language, based on Chen (2013). All columns use OLS regressions. We control for a large set of country and firm characteristics, which are all defined in Appendix II. Year, industry, and continent fixed effects are denoted as Year FE, Industry FE, and Continent FE. Standard errors are clustered at the firm level and corrected for heteroscedasticity. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level. An intercept is included but not reported.

	Non-US Firms		Sub-period: 1993-1999		Sub-period: 2000-2007		Sub-period: 2008-2014	
Dependent Var.	R&D/Assets	R&D/Sales	R&D/Assets	R&D/Sales	R&D/Assets	R&D/Sales	R&D/Assets	R&D/Sales
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Weak-FTR	0.007***	0.010***	0.006***	0.005*	0.011***	0.023***	0.005***	0.007**
	(8.17)	(5.64)	(3.87)	(1.69)	(8.83)	(8.23)	(4.05)	(2.50)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	159871	159871	61393	61393	87694	87694	66790	66790
adj. R^2	0.322	0.295	0.421	0.386	0.387	0.383	0.382	0.383

	Non-missing R&D		Control for IFRS		Tobit Regression		Control for lagged R&D	
Dependent Var.	R&D/Assets	R&D/Sales	R&D/Assets	R&D/Sales	R&D/Assets	R&D/Sales	R&D/Assets	R&D/Sales
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Weak-FTR	0.009***	0.018***	0.008***	0.013***	0.016***	0.034***	0.002***	0.003***
	(4.02)	(3.87)	(8.78)	(7.34)	(8.15)	(9.01)	(6.48)	(6.54)
Lagged R&D/Assets							0.826***	
							(151.67)	
Lagged R&D/Sales								0.821***
								(151.64)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	110409	110409	215877	215877	215877	215877	171767	171767
adj. R^2	0.419	0.455	0.387	0.378			0.802	0.821

Table 9: Time Preference or Reward-timing Uncertainty?

This table presents OLS estimation results to examine whether the positive relation between weak-FTR and R&D is due to managers' time preference or reward-timing uncertainty. Column (1) controls for *Time Preference*, an indicator of low subjective discount rate constructed by Wang et al. (2016). Column (2) investigates the effects of language FTR on firm capital expenditures (Capex). Column (3) further examines how language FTR affects the ratio of R&D over R&D plus Capex. Weak-FTR is a dummy variable equal to one if a country's dominant language is a weak-FTR language. Country controls and firm controls are those in Table 5. We also control for year, industry, and continent fixed effects. Standard errors are clustered at the firm level and corrected for heteroscedasticity. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level. All variables are defined in Appendix II.

Dependent Var.	R&D/Assets	Capex/Assets	R&D/(R&D+Capex)
	(1)	(2)	(3)
Weak-FTR	0.013***	0.001	0.048***
	(11.10)	(0.57)	(11.45)
Time Preference	-0.016**		
	(-2.37)		
Firm Controls	Yes	Yes	Yes
Country Controls	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	197,301	215,877	213,859
adj. R^2	0.388	0.296	0.552

Table 10: 2SLS

This table presents 2SLS results. The dependent variable is *Weak-FTR* in all the first stage regressions. The instrumental variable is Crop Return which is the pre-1500 CE crop return (Galor et al., 2018). Columns (1)-(4) present the country-level regressions. The dependent variable is Business R&D/GDP in Column (1) and Total R&D/GDP in Column (3). Columns (5)-(8) present the firm-level regressions. The dependent variable is R&D/Assets in Column (5) and R&D/Sales in Column (7). R&D/Assets is defined as R&D expenditures divided by book assets. R&D/Sales is defined as R&D expenditures divided by annual sales. *Weak-FTR* is equal to one if a country's dominant language is identified as a weak-FTR language, based on Chen (2013). All columns use OLS regressions. We control for a large set of country and firm characteristics, which are all defined in Appendix II. Year, industry, and continent fixed effects are denoted as Year FE, Industry FE, and Continent FE. Standard errors are clustered at the firm level and corrected for heteroscedasticity. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level. An intercept is included but not reported.

	Country Level				Firm Level			
Dependent Var.	Business R&D/GDP (2^{nd} stage)	Weak-FTR (1^{st} stage)	Total R&D/GDP (2^{nd} stage)	Weak-FTR (1^{st} stage)	R&D/Assets (2^{nd} stage)	Weak-FTR (1^{st} stage)	R&D/Sales (2^{nd} stage)	Weak-FTR (1^{st} stage)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Weak-FTR (Instrumented)	0.023***		0.035***		0.013***		0.025***	
	(4.28)		(2.88)		(4.13)		(3.68)	
Crop Return		-0.036***		-0.024***		-0.075***		-0.075***
		(-5.09)		(-2.79)		(-18.51)		(-18.51)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	435	435	459	459	215,877	215,877	215,877	215,877
adj. R^2	0.438		0.105		0.387		0.377	
F-stat		25.91		17.79		342.73		342.73

Table 11: Within-country Evidence: United States

This table provides within-country evidence for U.S. firms over the sample period 1991-2013. For each observation with a weak-FTR CEO, we match an observation with a strong-FTR CEO in the same year, employing a propensity-score-matching approach. In Panel A, we report results from Logit regressions used to calculate the propensity scores for the matching procedure, where the dependent variable is an indicator variable set to one if the firm has a weak-FTR CEO and zero otherwise. Columns (1) and (2) respectively show marginal effects for the full sample before matching and the subsample with matched observations. Panel B displays the distribution of propensity scores from the regression in Column (2) of Panel A. Panel C compares R&D levels between firms with weak-FTR CEOs vs. firms with strong-FTR CEOs. Panel D presents the regression results based on the matched sample. ***, **, and * indicate significance at the 1%, 5%, and 10% level. All variables are defined in Appendix II.

Panel A: Pre-Match and Post-Match Logit Regression

	Pre-Match Regression	Post-Match Regression
	(1)	(2)
Firm Size	0.221***	-0.003
	(5.261)	(-0.051)
Leverage	-1.076***	-0.065
	(-2.814)	(-0.094)
Market-to-Book	-0.018	0.026
	(-0.495)	(0.349)
Profitability	-1.323***	-0.567
	(-6.767)	(-1.225)
Tangibility	0.092	-0.009
	(0.216)	(-0.013)
Capital Intensity	-0.315***	0.058
	(-4.033)	(0.487)
N	19,946	458
pseudo R^2	0.042	0.008

Panel B: Estimated Propensity Score Distributions

Propensity score	Mean	Min	P25	P50	P75	Max	SD
Weak-FTR CEO	0.023	0.002	0.011	0.018	0.029	0.157	0.019
Strong-FTR CEO	0.023	0.002	0.011	0.018	0.028	0.156	0.019
Difference	0.000	0.000	0.000	0.000	-0.001	0.001	0.000

Panel C: Difference in R&D and Other Firm Characteristics

	With weak-FTR CEOs			With strong-FTR CEOs			**Difference**
	Mean	**SD**	**N**	**Mean**	**SD**	**N**	
R&D/Assets	0.073	0.109	229	0.043	0.079	229	0.030***
R&D/Sales	0.264	0.631	229	0.129	0.428	229	0.135***
Firm controls:							
Firm Size	13.498	2.065	229	13.634	1.940	229	-0.136
Leverage	0.178	0.166	229	0.184	0.169	229	-0.006
Market-to-Book	2.159	1.565	229	1.997	1.287	229	0.162
Profitability	0.015	0.289	229	0.058	0.249	229	-0.043*
Tangibility	0.229	0.185	229	0.227	0.183	229	0.002
Capital Intensity	5.434	1.084	229	5.377	0.960	229	0.057

Panel D: Regression Using Matched Sample

	R&D/Assets	R&D/Sales
	(1)	(2)
Weak-FTR CEO	0.017**	0.057**
	(2.42)	(2.17)
Firm Controls	Yes	Yes
Industry FE	Yes	Yes
Year FE	Yes	Yes
N	458	458
adj. R^2	0.252	0.196

Table 12: Within-country Evidence: Belgium

This table presents within-country estimation of FTR on firm R&D investment in Belgium, where some regions speak weak-FTR langues while other regions speak strong-FTR languages. The dependent variable is R&D/Assets in Columns (1) and R&D/Sales in Columns (2). R&D/Assets is defined as R&D expenditures divided by book assets. R&D/Sales is defined as R&D expenditures divided by annual sales. Weak-FTR is an indicator variable equal to one if a firm is headquartered in a region where the main language is of weak FTR, and zero otherwise. Country controls are the full set of country-level control variables in Table 5. Firm controls are the full set of firm-level control variables in Table 5. Year, industry, and continent fixed effects are denoted as Year FE, Industry FE, and Continent FE. Standard errors are clustered at the firm level and corrected for heteroscedasticity. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level. All variables are defined in Appendix II.

Dependent Var.	R&D/Assets	R&D/Sales
	(1)	(2)
Weak-FTR	0.018***	0.073**
	(2.60)	(2.07)
Firm Controls	Yes	Yes
Industry FE	Yes	Yes
Year FE	Yes	Yes
N	1,495	1,495
adj. R^2	0.441	0.324

Table 13: Further Robustness Tests

This table presents further robustness tests. Panel A presents country-level regressions and Panel B presents firm-level regressions. Columns (1)-(2) exclude multilingual countries. Columns (3)-(4) use population-weighted Weak-FTR as an alternative explanatory variable. Columns (5)-(6) exclude countries with small economies (i.e., total GDP in the bottom quintile). Columns (7)-(8) control for continent-by-year fixed effects. are based on three sub-periods. In Panel A, the dependent variable is Business R&D/GDP in odd-numbered columns, and Total R&D/GDP in even-numbered columns. In Panel B, the dependent variable is R&D/Assets in odd-numbered columns, and R&D/Sales in even-numbered columns. *Weak-FTR* is equal to one if a country's dominant language is identified as a weak-FTR language, based on Chen (2013). All columns use OLS regressions. We control for a large set of country and firm characteristics, which are all defined in Appendix II. Year, industry, and continent fixed effects are denoted as Year FE, Industry FE, and Continent FE. Standard errors are clustered at the firm level and corrected for heteroscedasticity. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level. An intercept is included but not reported. (Note: B. R&D/GDP = Business R&D/GDP, T. R&D/GDP = Total R&D/GDP)

Panel A: Country-level Regressions

	Excluding Multilingual Countries		Population-weighted Weak-FTR		Excluding Small Economies		Continent-Year Fixed Effects	
Depdent Var.	B. R&D/GDP	T. R&D/GDP	B. R&D/GDP	T. R&D/GDP	B. R&D/GDP	T. R&D/GDPs	B. R&D/GDP	T. R&D/GDP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Weak-FTR	0.002**	0.002***	0.002***	0.002***	0.003***	0.002***	0.002***	0.003***
	(2.15)	(3.17)	(3.35)	(3.62)	(2.90)	(2.94)	(3.38)	(3.82)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	413	437	435	459	353	367	435	459
adj. R^2	0.825	0.856	0.824	0.857	0.841	0.879	0.819	0.852

Panel B: Firm-level Regressions

	Excluding Multilingual Countries		Population-weighted Weak-FTR		Excluding Small Economies		Continent-Year Fixed Effects	
Dependent Var.	R&D/Assets	R&D/Sales	R&D/Assets	R&D/Sales	R&D/Assets	R&D/Sales	R&D/Assets	R&D/Sales
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Weak-FTR	0.009***	0.016***	0.008***	0.015***	0.0041**	0.013***	0.009***	0.017***
	(9.94)	(8.53)	(8.72)	(7.74)	(2.21)	(3.18)	(10.08)	(8.81)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	210,210	210,210	215,877	215,877	195,809	195,809	215,877	215,877
adj. R^2	0.387	0.379	0.387	0.378	0.391	0.385	0.388	0.379

Table 14: Innovation Output

This table presents evidence on how FTR affects country-level innovation output. The dependent variable in Column (1) is Pat_{OECD}, defined as the number of triadic patents (registered in multiple countries) reported in the Organization for Economic Co-operation and Development (OECD) triadic patent families database. The dependent variable in Column (2) is Pat_{OECD} measured on a per capita basis. Weak-FTR is a dummy variable equal to one if a country's dominant language is identified as a weak-FTR language. We control for the full set of country-level control variables used in Table 3 and continent and year fixed effects. All columns use OLS regressions. Standard errors are clustered at the country level and corrected for heteroscedasticity. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level. An intercept is included but not reported. All variables are defined in Appendix II.

Dependent Var.	Pat_{OECD}	Pat_{OECD} per capita
	(1)	(2)
Weak-FTR	1.86***	0.03***
	(7.38)	(11.10)
GDP per capita	1.08***	0.01***
	(5.06)	(6.42)
Stock Market	-0.40***	-0.00**
	(-3.18)	(-2.50)
Credit Market	2.96***	0.02***
	(9.38)	(6.41)
Shareholder Rights	0.65	-0.02***
	(0.71)	(-2.79)
Creditor Rights	-1.66***	-0.02***
	(-10.81)	(-12.85)
Patent Rights	1.76***	0.01***
	(6.65)	(4.26)
Individualism	6.46***	0.00
	(6.37)	(0.13)
Uncertainty Avo.	5.95***	0.02**
	(5.86)	(2.27)
Power Distance	4.64***	-0.07***
	(6.06)	(-7.13)
Masculinity	3.64***	-0.03***
	(5.05)	(-3.11)
Long-term Ort.	-3.37***	0.03***
	(-3.67)	(3.40)
Trust	6.27***	0.03*
	(4.68)	(1.67)
Catholic	-7.25***	-0.03***
	(-18.14)	(-14.73)
Continent FE	Yes	Yes
Year FE	Yes	Yes
N	424	424
adj. R^2	0.855	0.817

Appendix Table I: Hong Kong – before and after 1997

This table compares R&D investment in Hong Kong before and after 1997. In Columns (1) and (2), we use firms located in the other two Asian Tigers, namely Singapore and South Korea, as the control group. The sample comprises 4,083 firm-year observations from 1994 to 2001. In Columns (3) and (4), we use firms located in countries affected by the 1997 Asian Financial Crisis as the control group. The sample comprises 6,868 firm-year observations from 1994-2001. The dependent variable is R&D/Assets or R&D/Sales. Hong Kong is a dummy variable that is equal to one for firms based in Hong Kong, and zero otherwise. Post-1997 is a dummy variable that is set to one after the 1997 Hong Kong transfer of sovereignty from the U.K. to China, and zero otherwise. Country controls are the full set of country-level control variables in Table 5. Firm controls are the full set of firm-level control variables in Table 5. Year and industry fixed effects are denoted as Year FE and Industry FE. Standard errors are clustered at the firm level and corrected for heteroscedasticity. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level. All variables are defined in Appendix II.

Sample	Asian Tigers		Countries Affected by the 1997 Asian Financial Crisis	
Dependent Var.	R&D/Assets	R&D/Sales	R&D/Assets	R&D/Sales
	(1)	(2)	(3)	(4)
Hong Kong * Post-1997	**0.004*** **	**0.005** **	**0.002** **	**0.004** **
	(2.99)	**(2.35)**	**(2.55)**	**(2.42)**
Post-1997	-0.004***	-0.004***	-0.001**	-0.002**
	(-3.41)	(-2.88)	(-2.54)	(-2.35)
Firm Controls	Yes	Yes	Yes	Yes
Country Controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
N	4,083	4,083	6,868	6,868
adj. R^2	0.171	0.123	0.149	0.108

Appendix Table II: Placebo Tests - Canada

This table presents within-country Placebo tests of FTR on firm R&D investment in Canada, where there are two strong-FTR official languages (English and French). The dependent variable is R&D/Assets in Columns (1) and R&D/Sales in Columns (2). Placebo Weak-FTR is an indicator variable equal to one if a firm is headquartered in a province where the main language is English (English is relatively weaker than French in terms of language FTR), and zero otherwise. Firm controls are the full set of firm-level control variables in Table 5. Year and industry fixed effects are denoted as Year FE and Industry FE, respectively. Standard errors are clustered at the firm level and corrected for heteroscedasticity. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level. All variables are defined in Appendix II.

Dependent Var.	R&D/Assets	R&D/Sales
	(1)	(2)
Placebo Weak-FTR	0.001	-0.020
	(0.255)	(-0.812)
Firm Controls	Yes	Yes
Industry FE	Yes	Yes
Year FE	Yes	Yes
N	4,897	4,897
adj. R^2	0.560	0.464

Appendix Table III: The Moderating Effect of International Exposure

This table presents the moderating effects of international exposure on the relation between FTR and R&D investment at both the country and firm levels. Columns (1)-(2) present country-level regressions and Columns (3)-(4) present firm-level regressions. In Columns (1)-(4), the dependent variable is Business R&D/GDP, Total R&D/GDP, R&D/Assets, and R&D/Sales, respectively. *Weak-FTR* is equal to one if a country's dominant language is identified as a weak-FTR language, based on Chen (2013). Our proxy for international exposure is *International Migrant* defined as international migrant stock as a percentage of population (data source: World Bank). All columns use OLS regressions. We control for a large set of country and firm characteristics, which are all defined in Appendix II. Year, industry, and continent fixed effects are denoted as Year FE, Industry FE, and Continent FE. Standard errors are clustered at the firm level and corrected for heteroscedasticity. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level.

	Country Level		Firm Level	
Dependent Var.	Business R&D/GDP	Total R&D/GDP	R&D/Assets	R&D/Sales
	(1)	(2)	(3)	(4)
Weak-FTR*International Migrant	**-0.001*** **	**-0.001*** **	**-0.001*** **	**-0.001*** **
	(-9.77)	**(-11.39)**	**(-7.80)**	**(-7.41)**
Weak-FTR	0.006***	0.005***	0.010***	0.017***
	(7.84)	(6.90)	(10.02)	(8.38)
International Migrant	0.000***	0.000***	0.001***	0.001***
	(5.31)	(7.01)	(5.94)	(6.27)
Firm Controls	Yes	Yes	Yes	Yes
Country Controls	Yes	Yes	Yes	Yes
Continent FE	Yes	Yes	Yes	Yes
Industry FE	No	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	435	459	215,877	215,877
adj. R^2	0.864	0.895	0.387	0.378