# The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets

M. Keith Chen*
Yale University, School of Management and Cowles Foundation

August 2011

**Abstract**

Languages differ dramatically in how much they require their speakers to mark the timing of events when speaking. In this paper I test the hypothesis that being required to speak differently about future events (what linguists call strongly grammaticalized future-time reference) leads speakers to treat the future as more distant, and to take fewer future-oriented actions. Consistent with this hypothesis I find that in every major region of the world, speakers of strong-FTR languages save less per year, hold less retirement wealth, smoke more, are more likely to be obese, and suffer from worse long-run health. This holds true even after extensive controls that compare only demographically similar individuals born and living in the same country. While not dispositive, the evidence does not seem to support the most obvious forms of common causation. Implications of these findings for theories of intertemporal choice are discussed.

*Comments are welcome at 135 Prospect St, New Haven CT, 06511, or at keith.chen@yale.edu. I am indebted to Judy Chevalier, Östen Dahl, Shane Frederick, Emir Kamenica, Emily Oster, Sharon Oster, Ben Polak, and seminar participants at Yale and Berkeley for invaluable feedback. The most recent version of this working paper is available at http://faculty.som.yale.edu/keithchen/. Keywords: language, time preferences, savings behavior, health, national savings rates. JEL Codes: D03, D14, D91, E21, I10.

# 1 Introduction

Languages differ dramatically in how much they require their speakers to indicate the timing of events when speaking about them. For example, a German speaker predicting precipitation can naturally do so in the present tense, saying: "*Es regnet morgen*" which translates to: "It rain tomorrow". In contrast, English would require the use of the future tense, "It *will* rain tomorrow". Could this characteristic of language influence speakers' intertemporal choices?

In this paper I test the hypothesis that being required to speak in a grammatically distinct way about future events leads speakers to treat the future as more distant, and to take fewer future-oriented actions. Put another way, I ask whether a habit of speech which treats the present and future differently, can lead to a habit of mind that treats future rewards as more distant.

To do so, I draw on the Linguistics literature on future-time reference (FTR), which documents large amounts of variation in the degree to which languages require distinct grammatical treatment of present and future events. These differences are surprisingly large, even within small geographical regions. For example Western Europeans speak languages that range from having no future tense (like Finnish), to languages in which verbs have distinct and obligatory future forms (like Spanish).

I examine how these differences in languages' FTR correlate with their speaker's future-oriented behaviors such as saving, exercising, and abstaining from smoking. I also look at the cumulative effects of these behaviors such as retirement savings and long-run health. To avoid conflating differences in languages with other differences in the economic or social environment, my analysis includes extensive controls for individual and family characteristics, including country of birth and residence. Effectively, I only compare individuals who have the same demographics, family structure, and country of birth and residence, but who speak different languages.

Consistent with my hypothesis, I find that speakers of languages with little to no grammatical distinction between the present and future (weak-FTR speakers) engage in much more future-oriented behavior. Weak-FTR speakers are 30% more likely to have saved in any given year, and have accumulated an additional 170 thousand Euros by retirement. Extending my analysis to look at non-monetary investments in health, I find that by retirement, weak-FTR speakers are in better health by numerous measures; they are 24% less likely to have smoked heavily, are 29% more likely to be physically active, and are 13% less likely to be medically obese.

I then attempt to determine if differences in language are directly causing these differences in behavior, or if these correlations derive from cultural values or traits that are coincident with language differences. For example, most (but not all) Germanic languages have a weakly-grammaticalized future tense: could there also be a "Germanic" cultural value towards savings that is widely held by Germanic-language speakers but not directly caused by language? While not dispositive, the evidence does not seem to support the most obvious forms of common causation.

Most notably, several waves of the World Values Survey asked respondents about both their savings behavior, the language which they speak at home, and the degree to which "savings and thrift is an important value to teach children". I find that both a language's FTR and the degree to which a person thinks savings is an important value predict savings behavior. Interestingly though, these effects are completely independent: neither effect attenuates nor boosts the other. Indeed, in the World Values Survey a language's FTR is almost entirely uncorrelated with its speakers' stated values towards savings ($corr = -0.07$). This suggests that the language effects I identify operate through a channel which is independent of conscious attitudes towards savings.

Finally, I examine the effect that this differential propensity to save has on national savings rates of OECD countries. Several interesting patterns emerge. First, the FTR of a country's language has a significant effect on that countries aggregate savings rate. Countries with weak FTR save, on average, 6 percent more of their GDP per year than their strong-FTR counterparts. This effect is unchanged by the addition of life-cycle savings control variables, and holds in every major region of the world.

Second, this finding reverses the long-standing pattern of northern-European countries saving more that their southern counterparts. In specific, language effects induce an *aggregation reversal* in European savings rates. That is, while it is true that northern-European countries tend to save more, northern-Europeans also tend to speak weak-FTR languages. Once the effect of language is accounted for the effect of Latitude flips; within language classes, northern-European countries actually save less than their southern counterparts. This suggests that what has been commonly thought of as a north-versus-south divide in savings rates may actually be more fully explained by language.

The paper proceeds as follows. Section 2 reviews the linguistics literature on future-time reference (FTR), details the ways it differs across languages, and lays out my hypothesis. Section 3 details my empirical methods and the data I use for estimation. Section 4 presents the conditional correlations between a language's FTR and its speakers future-oriented behaviors. More detailed regressions investigate the degree to which these correlations can be taken as evidence of causation. A final set of regressions investigates the relationship between language and national savings rates within the OECD. Section 5 discusses issues surrounding the interpretation of these results before concluding.

## 2 Languages and Future-Time Reference

The ways languages require their speakers to speak about the future differ in two fundamental ways. Languages can differ in both *how* and *when* they require speakers to signal that they are talking about the future. For example, English (like all European languages), marks the future by modifying a sentence's verb. For example, I *walked* to work today, and *will walk* tomorrow if the sun is out. In contrast, many languages require speakers to distinguish future events by modifying a sentence's subject. For example, a Hausa speaker would use the future marker *zā*, more literally saying that "future me" (*zā nì*), walks to work tomorrow, unless "future it" (*zā à*) is raining.[1]

More subtly, languages also differ in *when* they require speakers to specify the timing of events, or when that timing can be left implied. The linguist Roman Jakobson explained this difference as: "Languages differ essentially in what they *must* convey and not in what they *may* convey." For example, if I wanted to explain to an English-speaking colleague why I wasn't at lunch, I would be obliged to tell him that I *went* to a seminar, speaking in the past tense. If I were speaking Mandarin (which has no tenses), it would be quite natural for me to say I *go* (*qù*) to a seminar, omitting all markers of time since the context leaves little room for misunderstanding. In this way, English forces its speakers to habitually attend to the timing of events in a way that Mandarin does not. Of course, this does not mean that Mandarin speakers are unable to understand the concept of time, only that they are not required to attend to it every time they speak.

These differences in the use of the future tense are surprisingly widespread, and even occur within native languages of the same country. For example Thieroff (2000) documents what Dahl (2000) calls a "futureless area" in Northern and Central Europe, including the Finno-Ugrian and all Germanic languages except English. European languages range from a tendency to never distinguish present and future time (like Finnish) to languages like French, which have separate "future" forms

[1] Hausa is a member of the West-Chadic genus and one of the most common language in Nigeria. See Dryer (2011) for a general introduction to Hausa scholarship, and see Newman (2000) for a comprehensive treatment of the future tense in Hausa.

of verbs.[2] A Finnish speaker, for example, would say both *Tänään on kylmää* (today is cold) and *Huomenna on kylmää* (tomorrow is cold) using the unmarked verb *on*, while French speakers would switch from *Il fait froid aujourd'hui* (it is cold today), to *Il fera froid demain* (it will-be cold tomorrow). English is a notable outlier in Europe; in all other Germanic languages the use of the future tense is optional when making *predictions* that have no *intentional* component. That is, while a German speaker predicting precipitation or forecasting a freeze could say: *Es regnet morgen*, or *Morgen ist es kalt* (both in the present tense), an English speaker would have to use the future tense (it *will rain* tomorrow, and tomorrow *will be* cold).

### 2.1 Future-Time Reference and a Linguistic-Savings Hypothesis

In this paper, I investigate the hypothesis that people whose languages require them to habitually mark future events as distinct will treat the future as more distant. Put another way, I ask whether a habit of speech to treat the present and future as distinct, can lead to a habit of mind that treats future rewards as more distant. This would lead speakers to take up fewer future-oriented actions; in general the attractiveness of current pain for future reward is declining in how distant the payoff feels. If this hypothesis is right, holding all else constant people who speak languages in which the future and present are grammatically indistinguishable should save, exercise, and plan more, and spend, smoke, and over-consume less.

## 3 Data and Methods

### 3.1 Coding Languages

In all of the regressions to follow the independent variable of main interest is "strong future-time reference". This is meant to summarize whether a language generally *requires* the use of the future tense when speaking about future events.

Most analyses in this paper (Tables 4 through 9), study speakers of European languages. In those regressions,"StrongFTR" corresponds perfectly with what Dahl (2000) calls "futureless" languages and Thieroff (2000) calls "weakly-grammaticalized future" languages. Dahl defines "futureless" languages as those which do not require "the obligatory use [of the future tense] in (main clause) prediction-based contexts". That is, English is a "strong FTR" language because the future tense is obligatory, even if the speaker has no control over the outcome being predicted (e.g., tomorrow it *will be* sunny). Thieroff notes that at least in Europe, this distinction maps more generally onto whether future events can be left unmarked (i.e. discussed in the present tense). That is, the use of the future tense in prediction-based contexts maps onto the broader question of whether the use of the future tense is obligatory.

Some regressions (Tables 1, 2, and 3) analyze the World-Values Survey, whose participants speak many non-European languages not analyzed in either Dahl or Thieroff. To extend their characterization to this broader set, I rely on several other cross-linguistic analyses that have studied the future tense (most notably Bybee et al. 1994, Dahl & Kós-Dienes 1984, Nurse 2008, and Cyffer et al. 2009), and on individual grammars for languages that are extensively spoken in the WVS but not covered by these broader analyses. A table of all languages included in this study

[2] Languages where verbs or pronouns have distinct future forms are said to have an "inflectional" future. In Europe, this includes most romance languges (except Romanian and Portuguese), and many Slavic and Semitic languages. See Dahl (1985) for source data on inflectional futures in Europe, and Dahl & Velupillai (2011) for a broad survey of inflectional futures around the world.

and their coding is in the appendix, and a complete description of my coding of languages can be found on my website.[3]

### 3.2 Savings Regressions in the WVS

My first set of regressions examines the World-Values Survey (2009), which was intended to be a global survey of world cultures and values. Although five waves of the WVS are available, I study only the last three, which ran from 1994 to 2007. In these (but not earlier) waves, participants were asked what language they normally speak at home, which I use a proxy for the language most likely to structure their thought. This allows me to study individuals across a set of 79 countries for which language data are available.

In these data, I estimate fixed-effect (or conditional) Logit models of an individual's propensity to save (versus not save) in the current year, regressed on the FTR strength of that individual's language and a rich set of fixed-effects for country and individual characteristics.[4] These fixed-effects control for a person's: country of residence, income decile within that country, marital status (with 6 different classifications), sex, education (with 8 different classifications), age (in ten-year bins), number of children, survey wave, and religion (from a set of 74) all interacted (for a total of 1.4 billion categories). Effectively, this analysis matches an individual with others who are identical on every dimension listed above, but who speak a different language. It then asks within these groups of otherwise identical individuals, do those who speak high-FTR languages behave differently than those who speak low-FTR languages? In addition, immigrants are excluded from this analysis so as to avoid conflating differences in a household's primary language with differences between natives and immigrants.

In addition, the WVS allows me to examine the interaction between the effect of language on savings behavior, and several beliefs and values questions asked of participants. This allows me to examine to what degree the measured effect of language on savings behavior is attenuated by such things as how much a person reports trusting other people, or how much they report that saving is an important cultural value. To a limited extent, this allows me to investigate whether language acts as a marker of deep cultural values that drive savings, or whether language itself has a direct effect on savings behavior.

### 3.3 Retirement Assets and Health Behaviors in the SHARE

The second dataset I analyze is the SHARE, the Survey of Health, Ageing and Retirement in Europe (Börsch-Supan & Jürges 2005). The SHARE is a panel survey that measures the socioeconomic status and health of retired households in 13 European countries. This allows me to complement my earlier analysis of saving from the WVS with analyses of both accumulated household wealth, and other future-oriented behavior measures such as smoking, exercise, and long-run health. Like my regressions in the WVS, my analysis of the SHARE looks only at within-country language variation among natives. Unfortunately, the SHARE does not record what language households

[3] Most importantly, several African countries are well represented in the WVS, have several national languages, but are not comprehensively studied by any large cross-language tense study. For these languages I rely on individual grammars which discuss the structure of that languages future tense. Most important were Adu-Amankwah (2003) for Akan, Nurse (2008) for the Bantoid languages, Olawsky (1999) and Lehr, Redden & Balima (1966) for Dagbani and Moore, Newman (2000) for Hausa, Carrell (1970), Emenanjo (1978), Ndimele (2009), and Uwalaka (1997) for Igbo, and Awobuluyi (1978), and Gaye & Beecroft (1964) for Yoruba.

[4] I use Chamberlain's (1980) fixed-effect (or conditional) logit model to estimate these regressions, since I have very few observations within each group defined by my fixed-effects. The Chamberlain model solves the resulting incidental-parameters problem.

speak at home. Instead, I exploit the fact that the survey instrument is offered in multiple languages; households can choose to take the survey in any of the national languages of their country. I use this choice as a proxy for their primary language.

Towards an analysis of the language and accumulated savings, I estimate several OLS models of total net household retirement assets regressed on a household's language and increasingly rich sets of fixed effects. The SHARE survey attempts a comprehensive measure all assets a household has, including income, private and public benefit payments, and all forms of assets (stocks, bonds, housing, etc.) For my other analyses I study the effect of language on several health measures. The SHARE contains several questions on health behaviors (such as smoking and exercise) as well as several physical-health measurements: body-mass-index, walking speed (as measured by a walking test), grip strength (as measured by a dynamometer), and respiratory health (peak expiratory air flow).

All of these regressions include fixed effects similar to those in the WVS so as to aid in comparing results across datasets. The richest of these regressions includes fixed effects for a household's: country of residence (13), income decile within that country, marital status (with 6 different classifications), sex, education (with 8 different classifications), age (in ten-year bins), number of children, and survey wave (2004 and 2006), all interacted for a total of 2.7 million categories. Again, immigrant families are excluded to avoid conflating differences driven by language with differences in immigrant families.

### 3.4 National Savings in the OECD

Finally, I study the relationship between language and the national accounts of the OECD from 1970 to present. These data are collected and harmonized by the OECD for all 34 member countries as well as for the Russian Federation.[5] Details on the exact construction of each OECD measure can be found in the Data Appendix. Importantly, all annual GDP measures are computed using the expenditure method, with constant PPPs using the OECD base year (2000).

These regressions attempt to determine whether the FTR structure of a country's language appears to affect national savings. The form of the national savings equation is a simple linear relation that follows closely from life-cycle savings theory (see Modigliani 1986 for a review). Essentially, I regress national-savings rates on the level and growth rate of GDP as well as a number of other country demographics. To this regression I add a weighted measure of the FTR strength of that country's languages. This is simply the FTR strength of each of that country's major languages, weighted by the percent of the country's population reports speaking those languages.[6] This language measure does not vary by year: these regressions test if the unexplained components of national savings vary cross-sectionally with a country's language, and do not try to identify off of demographic shifts within a country across time.

## 4 Results

If speaking differently about the future lead individuals to discount the future more, then the propensity to save should be negatively correlated with strong future-time reference. I examine this correlation in a regression framework which allows for a rich set of controls.

[5] I include the Russian Federation in this analysis because as of the writing of this paper they are in the process of joining the OECD, and were included in the harmonized OECD data.

[6] These relative language shares were obtained for each country from their national census taken closest to the year 2000.

## 4.1 Language, Beliefs and Savings

My first set of regressions examines the savings behavior of individuals in the World Values Survey. These regressions are carried out using fixed-effect (or conditional) logistic analysis, where the dependant variable $save_{it}$ is an individual reporting having saved in net this year.[7] I estimate the equation:

$$\Pr(save_{it}) = \frac{\exp(z_{it})}{1+\exp(z_{it})}, \tag{1}$$

where

$$z_{it} = \beta_1 StrongFTR + \beta_2 X_{it} + F_{it}^{ex} \times F_{it}^{en} \times F_t^c.$$

In equation 1, the main variable of interest $StrongFTR$ is a binary-coded characteristic of the language that the individual speaks at home. $X_{it}$ are characteristics of individual $i$ at time $t$, such as their self-reported beliefs about trust and savings. The $F$ variables are sets of fixed effects that are jointly interacted to form groups for the basis of analysis: the conditional-likelihood function is calculated relative to these groups. That is, individuals are compared only with others who are identical on every $F$ variable. $F_{it}^{ex}$ is a set of fixed effects that can be taken as exogenous, these are non-choice variables such as age and sex. $F_{it}^{en}$ is a set of fixed effects that are likely endogenous to an individual's discount rate, such as income, education and family structure. $F_t^c$ is a set of country-wave fixed effects. Empirical estimates of equation 1 are presented in Table 1; all coefficients are reported as odds ratios.

**Table 1: An Individual Saved This Year (WVS)**

| | (1) Saved | (2) Saved | (3) Saved | (4) Saved | (5) Saved | (6) Saved |
|---|---|---|---|---|---|---|
| Strong FTR | 0.462 | 0.717 | 0.720 | 0.706 | 0.695 | 0.697 |
| | [0.070]** | [0.113]* | [0.115]* | [0.102]* | [0.091]** | [0.092]** |
| Unemployed | | | 0.677 | 0.693 | 0.687 | 0.688 |
| | | | [0.031]** | [0.044]** | [0.044]** | [0.044]** |
| Trust | | | | | 1.082 | 1.083 |
| | | | | | [0.045] | [0.045] |
| Saving is Important | | | | | | 1.111 |
| (to teach children) | | | | | | [0.043]** |
| Fixed Effects: | | | | | | |
| Age × Sex | Yes | Yes | Yes | Yes | Yes | Yes |
| Country × Wave | No | Yes | Yes | Yes | Yes | Yes |
| Income × Edu | No | Yes | Yes | Yes | Yes | Yes |
| Married × Num Chil | No | No | No | Yes | Yes | Yes |
| All FEs Interacted | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 152,056 | 64,017 | 64,017 | 24,933 | 23,658 | 23,658 |

Regressions are fixed-effect (or conditional) logistic regressions with coefficients reported as odds ratios. Immigrants are excluded from all regressions. Robust standard errors are reported in brackets; all regressions are clustered at the country level.
* significant at 5%; ** significant at 1%.

[7] See Chamberlain (1980) for details on conditional-logistic analysis, and the data appendix for the exact wording of this and other questions in the WVS.

Regression 1 controls only for $F_{it}^{ex}$, (non-choice variables age and sex), so as to summarize the average difference in the propensity to save between strong and weak-FTR individuals. The coefficient of 0.462 can be interpreted as strong FTR families saving only 46% as often (at the yearly level) as weak FTR families. Regressions 2 and 3 add fully-interacted fixed effects for country, time, income, and education. On top of these, regressions 4 through 6 include controls for family structure. Regression 4 can be interpreted as demonstrating that even when comparing only individuals that are identical on every dimension discussed above, individuals who speak a language with strong FTR are roughly 30% less likely to report having saved this year. This effect is nearly as large as being unemployed (31%).

Regression 5 adds "Trust", (the most studied variable in the large literature on social capital) as an additional control. "Trust" measures whether an individual thinks "most people can be trusted". This measure has a large and marginally significant effect on the propensity of an individual to save; individuals who think others are generally trustworthy are on average 8% more likely to have saved this year. Interestingly, this effect appears to be largely independent of the effect of language. Indeed, by comparing regressions 4 and 5 we see that the inclusion of "Trust" if anything, *increases* the measured effect of language.

Regression 6 adds a variable intended to measure saving as an important cultural value. Specifically, this question asks whether "thrift and saving money" is a value which is important to teach children.[8] Unsurprisingly, individuals who report that saving money is important are more likely to save. Interestingly though, this effect is both smaller than the effect of language (11% versus 30%), and does not attenuate the effect of language on savings behavior. This can be seen by comparing regressions 5 and 6. Indeed, across individuals the belief that saving is an important value is almost completely uncorrelated with the FTR of their language ($corr = 0.07$).

Parameter estimates from this first set of regressions suggest that a language's FTR is an important predictor of savings behavior. This effect is both large (larger than that of other widely-studied variables) and survives an aggressive set of controls. Interestingly, it appears to be statistically independent of what was designed to be a good marker of saving and thrift as a cultural value. This suggests that the channel through which language affects the propensity to save is largely independent of the saving as a self-reported value. Later, I will discuss what this non-attenuation result suggests about the causal link between language and savings behavior.

Next, I look at which countries in the WVS have numerous native speakers of both weak and strong FTR languages. Figure 1 plots a histogram of countries in the WVS, organized by what percent of the country's survey respondents report speaking a high-FTR language at home. As Figure 1 shows, the vast majority of countries (72 of 79) have basically no intra-country variation in future-time reference. This is largely because in most countries one language dominates, and in many countries with multiple languages those languages share a common FTR structure. For example, even though Canada has both large English and French speaking populations, both French and English are high-FTR languages.

In 7 of 79 WVS countries however, at least 5% of the population speak languages that has a different FTR structure than the majority language. These are the countries which provide the majority of identification for the full fixed-effect regressions. Table 2 enumerates these countries, and reports the coefficient on strong FTR when my regression with the most aggressive controls (regression 6 in Table 1) is estimated in only that country. Also listed are the percents of the sample that speak either strong or weak-FTR languages in that country, the languages they speak, and the N of the country-specific regression.

[8] See the data appendix for the full wording of these questions in the WVS.

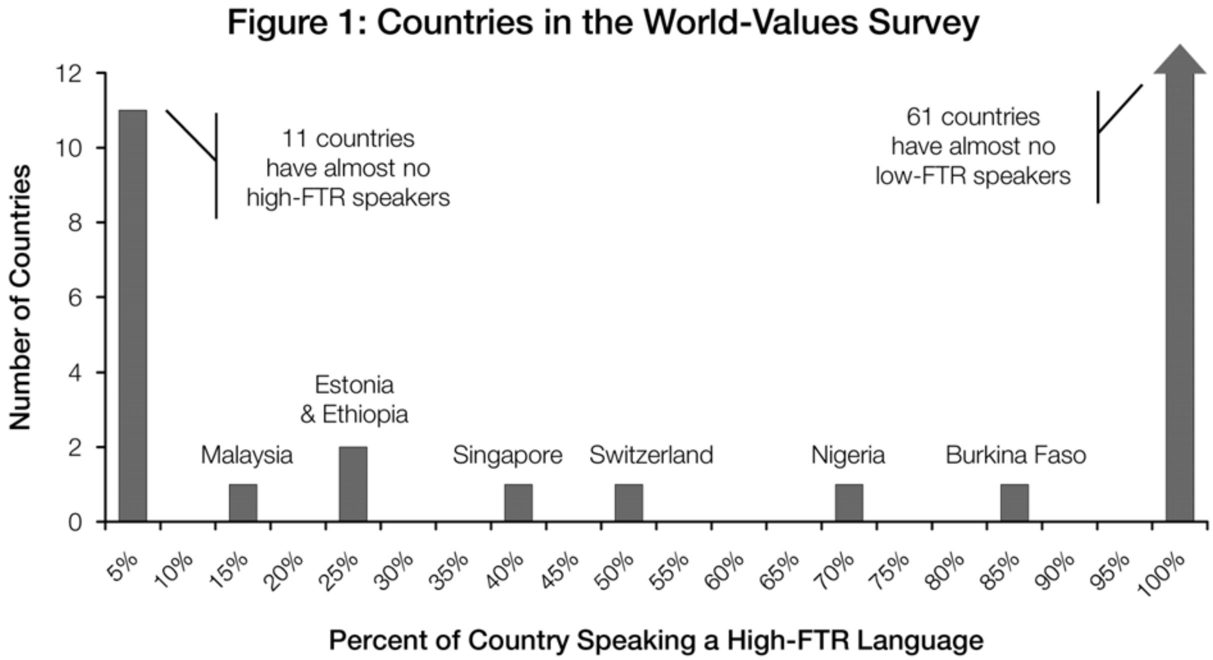


**Table 2: Countries with Large Within-Country FTR Differences in the WVS**

| Country | Weak-FTR Languages | % | Strong-FTR Languages | % | Coef. and SE on Strong FTR | N |
|---|---|---|---|---|---|---|
| Burkina Faso | Dyula | 16 | Fula, French, Moore | 84 | 0.700, [0.391] | 137 |
| Estonia | Estonian | 78 | Russian | 22 | 0.000, [0.000] | 31 |
| Ethiopia | Amharic, Oromo, Sidamo | 78 | Chaha, Gamo, Tigrinya | 22 | 0.825, [0.359] | 208 |
| Malaysia | Malay, Mandarin | 87 | English, Tamil | 13 | 0.742, [0.230] | 449 |
| Nigeria | Yoruba | 30 | English, Hausa, Igbo | 70 | 0.764, [0.355] | 121 |
| Singapore | Malay, Mandarin | 63 | English, Tamil | 37 | 0.821, [0.151] | 664 |
| Switzerland | German | 52 | French, Italian | 48 | 0.362, [0.132] | 172 |

Coefficients (reported as odds ratios) are from fixed-effect (or conditional) logistic regressions with the same specification as regression 6 in Table 1. Immigrants are excluded from all regressions.

Notably, all 7 regressions display coefficients less than 1, consistent with our overall effect. The coefficient in Estonia is 0 because in that regression, no Estonian speaker who was able to be matched with a Russian speaker reported not saving. Other than this outlier (which is largely driven by the small sample size in Estonia) the estimated effect is remarkably stable across this set of countries, which span multiple continents, regions, and sets of languages.

To confirm this and to explore the robustness of my initial results to additional controls, I estimate an additional set of regressions summarized in Table 3. First, I estimate the full regression (regression 6 in Table 1) separately in the 72 countries with little, and the 7 countries with sizable within-country FTR variation. I also examine whether these results are being driving by minority languages, by including as additional regressors for each household both the share of a country's speakers who speak their language, and the share that speak a language with the same FTR structure. Finally, I add as an additional control fixed-effects for self-reported religious denomination (74 in total), interacted with all of our previous fixed effects.

**Table 3: Additional Control Regressions in the WVS**

| | (1) Saved | (2) Saved | (3) Saved | (4) Saved | (5) Saved |
|---|---|---|---|---|---|
| Strong FTR | 0.934 | 0.678 | 0.679 | 0.528 | 0.529 |
| | [0.261] | [0.100]** | [0.101]** | [0.115]** | [0.115]** |
| Unemployed | 0.692 | 0.637 | 0.688 | 0.749 | 0.748 |
| | [0.046]** | [0.155] | [0.044]** | [0.068]** | [0.067]** |
| Trust | 1.071 | 1.273 | 1.083 | 1.068 | 1.068 |
| | [0.046] | [0.136]* | [0.044] | [0.051] | [0.051] |
| Saving is Important | 1.124 | 0.979 | 1.110 | 1.057 | |
| (to teach children) | [0.047]** | [0.082] | [0.043]** | [0.060] | |
| Language Share | | | 0.759 | 0.700 | 0.699 |
| | | | [0.119] | [0.129] | [0.130] |
| FTR Share | | | 1.071 | 0.467 | 0.461 |
| | | | [0.190] | [0.193] | [0.192] |
| Full set of FEs from Table 1 | Yes | Yes | Yes | Yes | Yes |
| Religion FEs | No | No | No | Yes | Yes |
| All FEs Interacted | Yes | Yes | Yes | Yes | Yes |
| Country's FTR Variation | < 5% (72) | > 5% (7) | All | All | All |
| Observations | 21,876 | 1,782 | 23,658 | 13,263 | 13,263 |

Regressions are fixed-effect (or conditional) logistic regressions with coefficients reported as odds ratios. Immigrants are excluded from all regressions. Robust standard errors are reported in brackets; all regressions are clustered at the country level.
* significant at 5%; ** significant at 1%.

Regressions 1 and 2 confirm our intuition that only the seven countries enumerated in Table 2 have enough within country variation to identify our full regression with country fixed effects. The coefficient of 0.678 is statistically indistinguishable from the coefficient of 0.697 I measure when the regression is run on the whole sample.

Returning to the whole sample: as an additional control, regression 3 demonstrates that the effect of language is not driven either by minority languages nor by minority FTR structures. Regressions 4 and 5 include additional fixed-effects for religious denomination (74 in total), interacted with all of our previous fixed effects. This inclusion does not attenuate the effect of language; comparing regression 3 to 4, the measured effect actually increases by 15%. Comparing regression 4 to 5 replicates our earlier non-attenuation finding: the addition saving as a self-reported value does not attenuate the main language effect. After the inclusion of religious controls both "trust" and saving as a cultural value attenuate enough to become statistically insignificant, while the effect of language, if anything, strengthens.

### 4.2 Language and Retirement Assets in Europe

If individuals who speak more strong FTR languages save less in any given year, then we would expect them to accumulate less savings over time. My next set of regressions examines the cumulative retirement assets of individuals in the retired households in the Survey of Health, Aging and Retirement in Europe. Table 4 summarizes regressions which estimate the equation:

$$ra_{it} = \alpha + \beta_1 StrongFTR + \beta_2 (F_{it}^{ex} \times F_{it}^{en} \times F_t^c) + \varepsilon_{it} \tag{2}$$

In equation 2 the dependant variable $ra_{it}$ is the estimated value of a retired household's net worth, including all real assets (homes, businesses and cars), and financial assets (money, stocks, bonds, and life insurance), minus any debt. Unfortunately, unlike the WVS, the SHARE does not ask households what language they speak at home. Here, the main variable of interest $StrongFTR$ is coded using the language that the head of household asked to take the survey in.

The $F$ variables are sets of fixed effects that are jointly interacted to form groups similar to those in my analysis of the WVS. That is, households are compared only with others who are identical on every $F$ variable, but who asked to take the survey in a different language. $F_{it}^{ex}$ is my set of exogenous fixed effects; here it is the age of the head of household. $F_{it}^{en}$ is a set of fixed effects that are likely endogenous to a household's discount rate, such as income, education and family structure. $F_t^c$ is a set of country-wave fixed effects. Empirical estimates of equation 2 are presented in Table 4; all coefficients are reported in Euros.[9]

**Table 4: Household Retirement Assets (SHARE)**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | HHAssets | HHAssets | HHAssets | HHAssets | HHAssets |
| Strong FTR | -154,515 | -150,498 | -145,151 | -173,880 | -178,744 |
| | [68,481]* | [12,703]** | [15,656]** | [9,723]** | [25,300]** |
| Fixed Effects: | | | | | |
| Age | Yes | Yes | Yes | Yes | Yes |
| Country × Wave | No | Yes | Yes | Yes | Yes |
| Income | No | No | Yes | Yes | Yes |
| Education | No | No | No | Yes | Yes |
| Married × Num Chil | No | No | No | No | Yes |
| All FEs Interacted | Yes | Yes | Yes | Yes | Yes |
| Observations | 39,665 | 39,665 | 39,665 | 39,665 | 39,350 |
| F stat | 5.09 | 140.37 | 85.96 | 319.81 | 49.91 |

Regressions are fixed-effect OLS regressions where the dependent variable is net household retirement assets in Euros. Immigrant households are excluded from all regressions. Robust standard errors are reported in brackets; all regressions are clustered at the country level.
* significant at 5%; ** significant at 1%.

Regressions 2 through 5 identify only off of within-country variation in language. These regressions are identified almost entirely off the fact that Belgium has large Flemish (weak FTR) and French (strong FTR) speaking populations, and Switzerland has large German (weak FTR), and French, Italian, and Romansh (strong FTR) speaking populations.

Regressions 1 through 5 show our predicted effect; retired households that speak strong FTR languages have saved around 170 thousand Euros less by the time they retire. Looking at regressions 1 and 2, we see that the addition of country fixed effects does not significantly attenuate the effect of language. The differences in cross-country in savings attributable to language appear to be roughly the same size as the differences between different FTR groups within Belgium and Switzerland.[10]

[9] Details on variable construction: Age is coded in ten-year bins, Income is coded as an intra-country decile, and Education falls within one of 8 categories provided in the SHARE. For more details on the construction of variables and the measuring of household net-wroth int he SHARE, see Börsch-Supan and Jürges (2005).

[10] The average net-household assets in the SHARE is 347 thousand Euros, but the coefficients in Table 2 are estimated almost entirely off of Switzerland and Belgium, which are higher (695K and 374K Euros, respectively). Swiss household net assets were recorded in Francs, which I convert to Euros using the average rate in the year the survey was taken (1.534 and 1.621 in waves 1 and 2 of the SHARE).

Table 5 summarizes regressions that contain the same set of demographic fixed effects as in Regression 5 from Table 4, but increase the level of spatial control by including fixed effects for intra-country regions. This allows us to examine whether language may be proxying (even within country) for unobserved differences between regions, counties or even cities. If for example, families tend to segregate across regions by language, then I may be attributing institutional differences between regions to language.

**Table 5: Household Retirement Assets in Belgium and Switzerland**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | HHAssets | HHAssets | HHAssets | HHAssets | HHAssets |
| Strong FTR | -178,744 | -187,424 | -256,369 | -105,840 | -147,410 |
| | [44,038]** | [39,268]** | [318,346] | [338,223] | [744,983] |
| Full set of FEs from Table 3 | Yes | Yes | Yes | Yes | Yes |
| Region FEs | 2 (BE & CH) | 1 | 11 | 1 | 7 |
| All FEs Interacted | Yes | Yes | Yes | Yes | Yes |
| Country | BE & CH | Belgium | Belgium | Switzerland | Switzerland |
| Observations | 5,937 | 4,394 | 4,393 | 1,543 | 1,543 |
| F stat | 16.47 | 22.78 | 2.44 | 0.10 | 0.16 |

Regressions are fixed-effect OLS regressions where the dependent variable is net household retirement assets in Euros. Immigrant households are excluded from all regressions. Robust standard errors are reported in brackets; all regressions are clustered at the household level.
* significant at 5%; ** significant at 1%.

Comparing regressions 2 and 3 (in Belgium) and regions regressions 4 and 5 (in Switzerland) shows that the addition of finer spatial controls (in the form of region dummies) does not appear to attenuate the effect of language on retirement savings. This suggests that the language effect we are measuring is not explained by unobserved spatial differences, at least not on the level we are able to capture with the regions coded in the SHARE.

### 4.3 Language and Health

The SHARE, in addition to measuring household wealth, also asks each member of the household about their health behaviors and records several measures of physical health. I look at these measures next, since if a languages affect their speakers intertemporal choices, this should also have implications for their speakers' health behaviors and long-run health. More specifically, if an obligatory future-tense reduces the psychological importance of the future, we would predict that it would lead to more smoking, less exercise, and worse long-run health.

To investigate this, Table 6 summarizes regressions investigating the effect of FTR on health variables found in the SHARE. Some of these measures are binary, such as ever having smoked heavily, remaining physically active, and being medically obese. For these regressions I estimate fixed-effect logit model similar to equation 1. The other measures I examine, walking speed, grip strength, and peak expiratory flow, are commonly studied measures of long-run health. These measure the spread at which a person comfortably walks, the maximum among of force they can apply while squeezing a dynometer, and their maximum exhalatory air flow (lung strength). For these regressions I estimate fixed-effect OLS regressions similar to equation 2.

**Table 6: Health Behaviors and Measures of Health (SHARE)**

| | (1) Smoked | (2) Phy Act | (3) Obesity | (4) Walk Sp | (5) Grip Str. | (6) Peak Flow |
|---|---|---|---|---|---|---|
| Strong FTR | 1.241 | 0.709 | 1.131 | -0.028 | -0.899 | -16.083 |
| | [0.042]** | [0.025]** | [0.007]** | [0.101] | [0.049]** | [2.806]** |
| Full set of FEs from Table 3 | Yes | Yes | Yes | Yes | Yes | Yes |
| All FEs Interacted | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 15,750 | 9,135 | 11,958 | 6,038 | 51,571 | 26,836 |
| R-squared | | | | 0.85 | 0.84 | 0.73 |

Regressions 1, 2, and 3 are fixed-effect (or conditional) logistic regressions with coefficients reported as odds ratios. The dependent variables are having ever smoked daily for a year or more, engaging in regular physical activity, and being medically obese. Regressions 4, 5, and 6 are fixed-effect OLS regressions for measures of old-age health; walking speed (m/sec), grip strength (kg), and peak expiratory flow(L/min). Immigrants are excluded from all regressions. Robust standard errors are reported in brackets; all regressions are clustered at the country level.
* significant at 5%; ** significant at 1%.

Regression 1 indicates that a strongly grammaticalized FTR leads to a 24% higher probability of having ever smoked (daily for a year or more). This is consistent with our findings on savings if the decision to smoke trades off immediate benefits versus future health costs. Similarly, regression 2 indicates that strong FTR leads to a 29% lower probability of being physically active. Regressions 3, 4, 5, and 6 examine the effect of strong FTR on long-run measures of health. While there appears to be no effect on walking speed, speaking a strong FTR language is associated with a 13% higher probability of being medically obese, a reduction in grip strength of almost a kilogram, and a reduction in peak expiratory flow of 16 liters per minute.

## 4.4 Linguistic Effects on National Savings Rates in the OECD

The evidence on both individual and household behavior we have presented so far supports our hypotheses that strongly grammaticalized FTR languages are associated with less future-oriented choices by its speakers. If, as our previous results suggest, people who speak strong-FTR languages discount more heavily, then it seems natural to expect that the countries in which they live would have lower equilibrium savings rates. This prediction does not immediately follow from theory, however.

Samuelson (1937) showed that when the duration of a potential project is fixed, the value of that project may not be even weakly decreasing in the interest rate. Arrow and Levhari (1969) established that if an agent controls when a project terminates, then in deterministic settings the natural monotonic relationship must hold; the value of investment in projects must be monotonically decreasing in the interest rate. In Hick's book *Capital and Time* (1973), this is referred to as the *Fundamental Theorem of Capital*. Under the conditions for which this relationship holds then, it is natural to predict that countries with strong FTR languages will, on average, save less.

Table 7 summarizes a first set of regressions which test this prediction. These regressions closely follow Barrow and McDonald (1979), who run similar regressions on the same OECD national savings data that we investigate here. The basic functional form of these regressions is:

$$((Y-C)/Y)_{it} = \alpha_0 + \alpha_1(\text{Str. FTR})_i + \alpha_2(1/Y)_{it} + \alpha_3(Y_{t-1}/Y_t)_{it} + \alpha_4(CAGR)_i + \varepsilon_{it} \quad (3)$$

where annual observations for each country in the OECD are indexed by country $i = 1, ..., 35$ and year $t = 1970, ..., 2009$. Details on the construction of each variable can be found in the Data Appendix, most importantly: $C$ is total consumption while $Y$ is GDP, $CAGR$ is the average growth rate of the country from 1993 to 2009 (the earliest date for which data is available for all countries).and "Strong FTR" is weighted by the percent of the country's population reports speaking each of their major languages.

This form of this savings equation is a simple linear relation that is based on simple forms of the Life Cycle Hypothesis (LCH) of savings (see Modigliani 1986 for a good review of the theory). Notice that as equation 3 is written, all terms in the savings equation except $(1/Y)_{it}$ imply that a savings function that is homogeneous of degree 0, which is to say that the savings rate is independent of the level or unit of income. This assumption has theoretical support in the LCH model, and allows for a specification in which units of measurement do not need to be comparable across countries. It may be violated if, as Feldstein (1977) points out, higher incomes lead to a increase in the share of life spent in retirement. This leads to the presence of the $1/Y_{it}$ term, which can test for such effects as measured by a positive $\alpha_2$. Essentially this term allows the marginal propensity to consume out of income to differ by the level of development of a country.

In addition, OECD data allows for the inclusion of a number of important demographic controls:

$$\alpha_5(Unemployment)_{it} + \alpha_6(Old)_{it} + \alpha_7(Young)_{it} + \alpha_8(SocSec)_{it}$$

These control for the unemployment rate, the fraction of the population that are over 65, the fraction under 15, and the per-capita fraction of GDP spent on social security payments (defined as % GDP spent on disability, old age, and survivors benefits divided by the fraction of the population that are over 65). Empirical estimates of equation 3 are presented in Table 7.

**Table 7: Gross Domestic Savings Rates in the OECD**

| | (1) $GDSR_t$ | (2) $GDSR_t$ | (3) $GDSR_t$ | (4) $GDSR_t$ |
|---|---|---|---|---|
| Strong FTR | -8.035 | -5.518 | -5.309 | -4.046 |
| | [2.813]** | [1.503]** | [1.786]** | [1.305]** |
| 1 / $PCGDP_t$ | | 136.863 | 143.727 | 43.580 |
| | | [48.654]** | [57.394]* | [56.031] |
| $PCGDP_{t-1}$ / $PCGDP_t$ | -37.106 | -23.486 | -20.016 | -21.766 |
| | [10.179]** | [6.645]** | [7.423]* | [6.883]** |
| CAGR | -0.110 | -0.248 | -0.302 | 0.010 |
| | [0.096] | [0.039]** | [0.125]* | [0.143] |
| $Unemployment_t$ (%) | -0.061 | -0.344 | -0.163 | -0.141 |
| | [0.237] | [0.177] | [0.174] | [0.135] |
| $Old_t$ (%) | -1.186 | -1.077 | -1.222 | -0.969 |
| | [0.408]** | [0.327]** | [0.356]** | [0.222]** |
| $Young_t$ (%) | -0.464 | -0.856 | -0.993 | -0.378 |
| | [0.337] | [0.277]** | [0.313]** | [0.319] |
| Soc $Sec_t$ (%GDP / Old) | | | -4.184 | |
| | | | [2.872] | |
| Fixed Effects: | None | None | None | Region (7) |
| Observations | 904 | 904 | 614 | 904 |
| R-squared | 0.56 | 0.67 | 0.66 | 0.75 |

Regressions are OLS regressions where the dependent variable is a country's Gross Domestic Savings Rates in year $t$. Observations are for OECD countries from 1970 to 2009. All regressions are weighted by the population of the country in that year. Robust standard errors are reported in brackets and clustered at the country level.
* significant at 5%; ** significant at 1%.

Regression 1 estimates a version of equation 3 that is fully homogeneous of degree 0; regressions 2 and 3 add a term which allows savings rates to vary with the size of the economy, and not just its short and long-run growth rates. These regressions suggests that countries with a strong FTR save on average around 5% (percentage points) less per year than do countries with weak FTR, a result consistent with our earlier results on household savings and health measures. Regression 4 includes region fixed-effects, where the OECD countries are apportioned into 7 regions: Australia, E & W Europe, the Middle East, N & S America, and SE Asia.

## 4.5 Linguistic Effects in the OECD: Robustness Checks and an Aggregation Reversal

One concern with the result that strong-FTR countries tend to save more is that the FTR strength of countries is spatially correlated. In Western Europe for example, most strong-FTR countries are in the northern half of the continent. This leads to the possibility that (at least in Western Europe), the effects I attribute to strong FTR could actually be due to a correlated spatial factor (like climate or distance from Mediterranean trade routes) which leads northern-European countries to save more than southern-European countries. Similar stories might also invalidate our results on other continents.

To examine whether these types of spatial confounds are a concern, I re-estimate equation 3 with an additional control variable, "distance from equator". This is the distance from a country's capital to the equator in thousands of miles. If the effects of language reported in Table 7 were actually due to a spatial factor correlated with Latitude, then we would expect the inclusion of "distance from equator" to attenuate or eliminate the coefficient on language. Table 8 reports the results of these regressions.

**Table 8: Gross Domestic Savings Rates in the OECD by Region**

| | (1) $GDSR_t$ | (2) $GDSR_t$ | (3) $GDSR_t$ | (4) $GDSR_t$ |
|---|---|---|---|---|
| Strong FTR | -5.578 | -7.343 | -8.951 | -16.310 |
| | [1.456]** | [1.814]** | [4.634] | [5.560]* |
| 1 / $PCGDP_t$ | 135.863 | -163.861 | 19.524 | 127.156 |
| | [49.985]* | [87.815] | [37.737] | [79.621] |
| $PCGDP_{t-1}$ / $PCGDP_t$ | -24.360 | -35.846 | -2.111 | -23.717 |
| | [5.504]** | [4.326]** | [8.143] | [8.045]* |
| CAGR | -0.246 | 0.117 | -0.191 | 1.169 |
| | [0.040]** | [0.624] | [0.085] | [0.608] |
| $Unemployment_t$ (%) | -0.329 | 0.070 | -0.642 | -0.433 |
| | [0.185] | [0.119] | [0.120]** | [0.234] |
| $Old_t$ (%) | -1.061 | -0.157 | -1.158 | -1.103 |
| | [0.335]** | [0.379] | [0.698] | [0.315]** |
| $Young_t$ (%) | -0.859 | 0.607 | -1.017 | -0.798 |
| | [0.277]** | [0.334] | [0.443] | [0.350] |
| Dist from Equator | -0.277 | -5.007 | -2.300 | 9.766 |
| (1K miles) | [0.983] | [2.050]* | [3.485] | [4.260] |
| Region: | All | W EU | E EU & Mid. East | All others |
| Observations | 904 | 539 | 109 | 256 |
| R-squared | 0.67 | 0.41 | 0.73 | 0.85 |

Regressions are OLS regressions where the dependent variable is a country's Gross Domestic Savings Rates in year $t$. Observations are for OECD countries from 1970 to 2009. Regression 1 includes: Austria, Belgium, Denmark, Finland, France, Germany, Great Britain, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, and Switzerland. Regression 2 includes: Czech Republic, Estonia, Hungary, Israel, Poland, Russian Federation, Slovak Republic, Slovenia, and Turkey. Regression 3 includes: Australia, Canada, Chile, Japan, Mexico, New Zealand, South Korea, and the United States. All regressions are weighted by the population of the country in that year. Robust standard errors are reported in brackets and clustered at the country level.
* significant at 5%; ** significant at 1%.

The results in Tables 8 suggest that this type of spatial confound seems unlikely. Regressions 1 through 4 demonstrate that the effects I attribute to language are not attenuated by the addition of "dist from equator", neither in Western Europe nor in any other major OECD region. Comparing regression 2 from Table 7 to regression 1 in Table 8, we see that the effect of language on savings is unchanged ($-5.518$ vs. $-5.578$). If anything, the inclusion of north-south spatial controls strengthen the measured effect of language in every region of OECD.

Interestingly, the coefficient on "dist from equator" in regression 2 is the opposite sign of the common observation that northern-European countries tend to save more than their southern counterparts. Quite the contrary, I find that when language controls are included, European countries save on average 5 percentage of their GDP *less* for every thousand miles further north they lie. To further investigate this finding, I re-estimate equation 3 restricted to Western Europe, examining what effect the inclusion and removal of language controls have on the measured effect of distance from the equator. Table 9 details these regressions.

**Table 9: Aggregation Reversal in Western Europe by FTR Strength**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | $GDSR_t$ | $GDSR_t$ | $GDSR_t$ | $GDSR_t$ | $GDSR_t$ |
| Dist from Equator | 0.980 | 1.510 | -5.007 | -2.582 | -4.786 |
| (1K miles) | [1.999] | [1.675] | [2.050]* | [2.002] | [2.095] |
| Strong FTR | | | -7.343 | | |
| | | | [1.814]** | | |
| 1 / $PCGDP_t$ | | -71.439 | -163.861 | -106.322 | -100.616 |
| | | [98.879] | [87.815] | [67.106] | [155.671] |
| $PCGDP_{t-1}$ / $PCGDP_t$ | | -29.947 | -35.846 | -34.836 | -34.282 |
| | | [6.735]** | [4.326]** | [4.986]** | [7.738]** |
| CAGR | | -0.130 | 0.117 | 1.266 | -0.576 |
| | | [0.707] | [0.624] | [0.381]** | [0.961] |
| $Unemployment_t$ (%) | | -0.084 | 0.070 | -0.391 | 0.209 |
| | | [0.173] | [0.119] | [0.148]* | [0.157] |
| $Old_t$ (%) | | -1.103 | -0.157 | -0.036 | -0.455 |
| | | [0.444]* | [0.379] | [0.321] | [0.688] |
| $Young_t$ (%) | | -0.539 | 0.607 | -0.035 | 0.254 |
| | | [0.387] | [0.334] | [0.288] | [0.289] |
| FTR of Country: | All | All | All | Weak | Strong |
| Observations | 720 | 539 | 539 | 323 | 216 |
| R-squared | 0.01 | 0.19 | 0.41 | 0.58 | 0.24 |

Regressions are OLS regressions where the dependent variable is a country's Gross Domestic Savings Rates in year $t$. Observations are for Western-European OECD countries from 1970 to 2009. Regression 4 includes: Austria, Belgium, Denmark, Finland, Germany, Iceland, Luxembourg, Netherlands, Norway, Sweden, and Switzerland. Regression 5 includes: France, Great Britain, Greece, Ireland, Italy, Portugal, and Spain. Regressions 1, 2, and 3 include both sets of countries. All regressions are weighted by the population of the country in that year. Robust standard errors are reported in brackets and clustered at the country level.
* significant at 5%; ** significant at 1%.

These regressions suggest that what is often thought of as a north-versus-south divide in European savings rates may be better explained by language than geography. In specific, language patterns appear to induce an *aggregation reversal* in savings rates. That is, while northern-European countries tend to save more that southern-European countries; after controlling for language the opposite is true (countries save more the further South they are).Regressions 1 and 2 demonstrate the commonly held wisdom that countries appear to have higher savings rates the further north they lie in Europe, both without and with economic controls. The coefficient in regression 2 can be interpreted as saying that holding economic conditions constant, a western-European country saves

1.5% of GDP more per year for every one thousand miles more north their capital lies. However after controlling for "strong FTR" in regression 3 the sign flips: a country saves on average 5% *less* for every thousand miles it lies further north. Regressions 4 and 5 demonstrate this aggregation reversal directly. Within both sets of western European countries (strong and weak FTR), countries that lie further north save less than their southern counterparts.

# 5 Discussion

## 5.1 Intertemporal Choice and the Determinants of Discounting

At least since Samuelson (1937) introduced the discounted-utility model, rates of time preference have been seen as a central part of almost all important economic decisions. Despite this centrality, most economic analysis takes the level of time discounting for granted as exogenous. Notable exceptions include Barro & Becker (1989) which models discount rates as a function of fertility, and Becker & Mulligan (1997) which models a consumer who invests in lowering their own discounting of future utility. The determinants of time-preference has also been investigated in related fields. In sociobiology, Rogers (1994) models the effect of natural selection on time preferences. He finds that if evolution sets the discount rate equal to the rate of substitution of Darwinian fitness, then people will discount the future at a rate of ln(2) per generation, which is about 2% per year.

A large literature in child development studies the acquisition of language in children, and several papers have studied the specific question of how children acquire the ability to conceptualize and speak about time. Most notably, Harner (1981) finds that among English-speaking children the use of the future tense begins by age 3 and is relatively developed by age 5. Szagun (1978) finds that the time-path of this development is identical in matched pairs of English and German children, with these pairs of children showing no discernible difference in the rate at which they acquire and use the future tense. This is despite the fact that English is a strong-FTR language while German a weak-FTR language. This distinction is reflected in Szagun's study, but only among adults: the German-speaking parents of the children Szagun studied used the future tense much less than the English-speaking parents did. That is, differences in the use of the future tense across languages do not seem to manifest themselves in early language acquisition, they emerge later in life. While far from dispositive, this suggests that the differences that I study between weak and strong-FTR languages do not reflect either innate cognitive nor cultural differences between speakers of different languages, at least as reflected in the development of children through age 5.

Some empirical findings suggest that individual's time preferences are closely linked with other characteristics. Warner & Pleeter (2001) found large amounts of variation in personal discount rates among military personal who were offered either a lump-sum payment or an annuity upon leaving the military. Suggestively, these discount rates were highly correlated with age, race, sex, and scores on an IQ-like test. Similarly, Frederick (2005) finds that even at elite universities, students who score high on an IQ-like "cognitive-reflection test" showed much lower discount rates.

Closest to this paper, a large literature in psychology and behavioral economics studies numerous well-documented departures from the discounted-utility model. For example, Lowenstein (1988) finds a reference-point effect: people demand much more compensation to delay receiving a good by one year, (from today to a year from now), than they are willing to pay to move up consumption of that same good (from a year from now to today). Read & Frederick et al. (2005) show that discount rates are lower when time is described using calendar dates (e.g., on October 17) than when described in terms of the corresponding delay (e.g., in six months).[11]

[11] See Frederick, Lowenstein, and O'Donoghue (2002) for an excellent review of this literature.

Many of these departures are framing effects: instances in which different descriptions of the same intertemporal choice lead people to make different decisions. This paper can be seen as asking whether the structure of a person's language can subtly act as such a frame at the moment of decision, with cumulatively large effects. Alternatively, my findings are consistent with languages which force individuals to mark the future as distinct from present leading to a mindset in which the future is discounted more heavily.

## 6 Conclusion

Overall, my findings are largely consistent with the hypothesis that languages with obligatory future-time reference lead their speakers to engage in less future-oriented behavior. On savings, the evidence is consistent on multiple levels: at an individual's propensity to save, to long-run effects on retirement wealth, and in the aggregate with national savings rates. These findings also extend to health behaviors ranging from smoking to exercise, and reflect in several measures of long-run health. All of these results survive after comparing only individuals who are identical on numerous demographic levels, and who were born and raised in the same country.

One important issue in interpreting these results is the possibility that language is not *causing* but rather *reflecting* deeper differences that drive savings behavior. These available data provide preliminary evidence that much of the measured effects I find are causal, for several reasons that I have outlined in the paper. Mainly, self-reported measures of savings as a cultural value appear to drive savings behavior, and are completely uncorrelated with the effect of language on savings. In addition, differences in the use of FTR do not seem to correspond to cognitive or developmental differences in the acquisition of language. This suggests that the effect of language that I measure occurs through a channel that is independent of either cultural or cognitive differences between linguistic groups.

Nevertheless, the possibility that language acts only as a powerful marker of some deeper driver of intertemporal preferences cannot be completely ruled out. This possibility is intriguing in itself, as the differences I identify off of in languages are themselves very old. In Europe for example, Dahl (2000) notes that Germanic and Finno-Ugrian languages have been futureless for at least two-thousand years. Additional evidence of language's role in shaping intertemporal choice is the goal of ongoing experimental work (Boroditsky and Chen, 2011), which hopes to isolate the channel through which this linguistic-savings effect occurs.

# 7 Data Statements

This paper uses data from SHARE release 2.3.1, as of July 29th 2010. SHARE data collection in 2004-2007 was primarily funded by the European Commission through its 5th and 6th framework programmes (project numbers QLK6-CT-2001- 00360; RII-CT- 2006-062193; CIT5-CT-2005-028857). Additional funding by the US National Institute on Aging (grant numbers U01 AG09740-13S2; P01 AG005842; P01 AG08291; P30 AG12815; Y1-AG-4553-01; OGHA 04-064; R21 AG025169) as well as by various national sources is gratefully acknowledged (see http://www.share-project.org for a full list of funding institutions).

# References

[1] Adu-Amankwah, David. 2003. *Asante-Twi Learners' Reference Grammar*. Madison: NALRC Press

[2] Arrow, Kenneth, David Levhari. 1969. "Uniqueness of the Internal Rate of Return with Variable Life of Investment". *Economic Journal*, 79(315): 560-566.

[3] Awobuluyi, Qladele. 1978. *Essentials of Yoruba Grammar*. Ibadan: Oxford University Press

[4] Barro, Robert, and Gary Becker. 1989. "Fertility Choice in a Model of Economic Growth". *Econometrica*, 57(2): 481-501.

[5] Barro, Robert, and Glenn MacDonald. 1979. "Social Security and Consumer Spending in an International Cross Section". *Journal of Public Economics*, 11(3): 275-289.

[6] Barro, Robert, and Xavier Sala-i-Martin, 1995. *Economic Growth*. New York:McGraw-Hill.

[7] Boroditsky, Lera, and M. Keith Chen. 2011. *Language and Time: Experimental Evidence for a Linguistic-Savings Effect*. Working Paper.

[8] Börsch-Supan, Anette Reil-Held, Ralf Rodepeter, Reinhold Schnabel, and Joachim Winter. 2001. "The German Savings Puzzle". *Research in Economics*, 55(1): 15-38.

[9] Becker, Gary and Barro, Robert. 1988. "A Reformulation of the Economic Theory of Fertility". *The Quarterly Journal of Economics*, 103(1): 1-25.

[10] Becker, Gary and Casey B. Mulligan. 1997. "The Endogenous Determination of Time Preference". *The Quarterly Journal of Economics*, 112(3): 729-758.

[11] Börsch-Supan, Axel, and Hendrik Jürges. ed. 2005. *The Survey of Health, Aging and Retirement in Europe – Methodology*. Mannheim: Research Institute for the Economics of Aging.

[12] Bybee, Joan, Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar*. Chicago: University of Chicago Press.

[13] Carrell, Patricia L. 1970. *A Transformational Grammar of Igbo*. Cambridge: Cambridge University Press.

[14] Chamberlain, Gary. 1980. "Analysis of Covariance with Qualitative Data". *Review of Economic Studies*, 47(1): 225-238.

[15] Cyffer, Norbert, Erwin Ebermann, and Georg Ziegelmeyer. 2009. *Negation Patterns in West African Languages and Beyond*. Amsterdam: John Benjamins Publishing Company.

[16] Dahl, Östen, and Dora Kós-Dienes. 1984. "Selected Working Papers from the Tense-Mood-Aspect Project". Stockholm: *Institute of Linguistics Working Paper*, University of Stockholm.

[17] Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Basil Blackwell.

[18] Dahl, Östen. 2000. "The Grammar of Future Time Reference in European Languages". In *Tense and Aspect in the Languages of Europe*, ed. Östen Dahl, 309-328. Berlin: Mouton de Gruyter.

[19] Dahl, Östen and Viveka Velupillai. 2011. "The Future Tense". In *The World Atlas of Language Structures Online*, ed. Matthew Dryer & Martin Haspelmath. Munich: Max Planck Digital Library. chapter 67. accessed June 1, 2011. http://wals.info/chapter/67.

[20] Dryer, Matthew, and Martin Haspelmath (eds.). 2011. "The World Atlas of Language Structures Online". Munich: Max Planck Digital Library. accessed June 1, 2011. http://wals.info/.

[21] Emenanjo, E. Nolue. 1978. *Elements of Modern Igbo Grammar*. Ibadan: Oxford University Press

[22] Feldstein, Martin. 1980. "International Differences in Social Security and Saving". *Journal of Public Economics*, 14(2): 225-244.

[23] Fisher, Irving. 1930. *The Theory of Interest*. New York: Macmillan.

[24] Frederick, Shane, George Loewenstein, and Ted O'Donoghue. 2002. "Time Discounting and Time Preference: A Critical Review". *Journal of Economic Literature,* 40(2): 351-401.

[25] Frederick, Shane. 2005. "*Cognitive Reflection and Decision Making*". Journal of Economic Perspectives. 19(4): 25-42.

[26] Gaye, J.A. and W.S. Beecroft. 1964. *Yoruba Grammar*. London: Routledge.

[27] Harner, Lorraine. 1981. "Children Talk about the Time and Aspect of Actions". *Child Development,* 52(2): 498-506.

[28] Hicks, John. 1973 *Capital and Time*. Oxford: Clarendon Press.

[29] Lehr, Marianne, James E. Redden, and Adam Balima. 1966. *More, a Basic Course.* Washington D.C: United States State Department, Foreign Service Institute.

[30] Loewenstein, George. 1988. Frames of Mind and Intertemporal Choice. *Management Science*, 34(2): 200-214.

[31] Modigliani, Franco. "Life Cycle, Individual Thrift, and the Wealth of Nations". *The American Economic Review*, 76(3): 297-313.

[32] Ndimele, Ozo-mekuri. 2009. "Negation Markings in Igbo". In *Negation Patterns in West African Languages and Beyond*, ed. Cyffer, Ebermann and Ziegelmeyer, 121-137. Amsterdam: John Benjamins Publishing Co.

[33] Newman, Paul. 2000. *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven: Yale University Press.

[34] Nurse, Derek. 2008. *Tense and Aspect in Bantu*. New York: Oxford University Press.

[35] Olawsky, Knut J. 1999. *Aspects of Dagbani Grammar: With Special Emphasis on Phonology and Morphology.* Munich: LINCOM Europa.

[36] OECD. 2010. *National Accounts of OECD Countries.* http://stats.oecd.org (accessed September 15th, 2010).

[37] Poterba, James, ed. 1994. *International Comparisons of Household Saving.* Chicago: NBER, University of Chicago Press.

[38] Poterba, James, Steven Venti, and David Wise. 1995. *Do 401(k) contributions Crowd Out Other Personal Saving?". Journal of Public Economics*, 58(1): 1-32.

[39] Read, Daniel, Shane Frederick, Burcu Orsel, and Juwaria Rahman. "Four Score and Seven Years From Now: The Date/Delay Effect in Temporal Discounting". *Management Science*, 51(9): 1326-1335.

[40] Rogers, Alan. 1994. "Evolution of Time Preference by Natural Selection". *The American Economic Review*, 84(3): 460-481.

[41] Samuelson, Paul. 1937. "Some Aspects of the Pure Theory of Capital". *Quarterly Journal of Economics*, 51(3): 469-496.

[42] Szagun, Gisela. 1978. "On the Frequency of Use of Tenses in English and German Children's Spontaneous Speech". *Child Development,* 49(3): 898-901.

[43] Thieroff, Rolf. 2000. "On the Areal Distribution of Tense-Aspect Categories in Europe". In *Tense and Aspect in the Languages of Europe*, ed. Östen Dahl, 309-328. Berlin: Mouton de Gruyter.

[44] Uwalaka, Mary Angela. 1997. *Igbo Grammar*. Ibadan: The Pen Services

[45] Warner, John T. and Saul Pleeter. 2001. "The Personal Discount Rate: Evidence from Military Downsizing Programs". The American Economic Review, 91(1): 33-53.

[46] World Values Survey Association. 2009. "World Values Survey 1981-2008, Official Aggregate". v.20090901. Aggregate File Producer: ASEP/JDS, Madrid. www.worldvaluessurvey.org.

[47] Whorf, Benjamin Lee. 1940. "Science and Linguistics". *MIT Technology Review,* 42(6): 229-231, 247-248.

# 8 Data Appendix

## 8.1 Wording of Questions in the WVS

FAMSAVED: During the past year, did your family (read out and code one answer):
Save money (23%)
Just get by (51%)
Spent some savings and borrowed money (14%)
Spent savings and borrowed money (12%)
TRUST: Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people? (Code one answer):
Most people can be trusted. (26%)
Need to be very careful. (74%)
CHILDSAVE: Here is a list of qualities that children can be encouraged to learn at home. Which, if any, do you consider to be especially important? Please choose up to five! (Code five mentions at the maximum):

| | |
|---|---|
| Independence | Hard work |
| Feeling of responsibility | Imagination |
| Tolerance and respect for other people | Thrift, saving money and things (37%) |
| Determination, perseverance | Religious faith |
| Unselfishness | Obedience |

## 8.2 Variables in the SHARE

HHNETWORTH: A household net worth in the SHARE "HHNetWorth" is attempt to measure all real assets net of any debts on them. It is equal to the estimated value of a household's: main residence, real estate other than the main residence, businesses, cars, bank accounts, bonds, stocks, mutual funds, life insurance, minus mortgage and other debt.
SMOKED: This codes whether an individual reports: "Have you ever smoked cigarettes, cigars, cigarillos or a pipe daily for a period of at least one year?"
PHYSICALLY ACTIVE: Physical inactivity is defined as "never or almost never engaging in neither moderate nor vigorous physical activity." Being physically active is not being inactive.
OBESITY: This is defined as a body-mass-index of 30 or greater.
WALKING SPEED: This was measured only among individuals aged 76 years and older. Walking speed was averaged over two tests of walking speed, as measured in meters per second.
GRIP STRENGTH: Grip strength is measured with a dynamometer at the interview (in kg).
PEAK FLOW: Peak expiratory flow measures a person's maximum exhalation air-flow, as measured with a peak-flow meter (in L/min).

## 8.3 OECD Variables

All GDP-based measures are computed using the expenditure method, with constant PPPs using the OECD base year (2000). CAGR is the average growth rate of the country from 1993 to 2009 (the earliest date for which data is available for all countries). $\text{Old}_t$ and $\text{Young}_t$ are the percent of the population that are older than 65 and younger than 15 in year $t$. Social Security is the percent of a country's GDP spent in a given year on disability, old age, and survivors benefits, divided by the percent of the population over 65.

**Appendix Table: Coded Languages and FTR Values**

| Language | Family | Genus | FTR |
|---|---|---|---|
| Afrikaans | Indo-European | Germanic | Strong |
| Akan | Niger-Congo | Kwa | Strong |
| Alawa | Australian | Maran | Strong |
| Albanian | Indo-European | Albanian | Strong |
| Amharic | Afro-Asiatic | Semitic | Weak |
| Arabic | Afro-Asiatic | Semitic | Strong |
| Armenian | Indo-European | Armenian | Strong |
| Azari | Altaic | Turkic | Strong |
| Azerbaijani | Altaic | Turkic | Strong |
| Bandjalang | Australian | Pama-Nyungan | Strong |
| Bambara | Niger-Congo | Western Mande | Weak |
| Basque | Basque | Basque | Strong |
| Belorussian | Indo-European | Slavic | Strong |
| Bemba | Niger-Congo | Bantoid | Strong |
| Bengali | Indo-European | Indic | Strong |
| Beja | Afro-Asiatic | Beja | Weak |
| Bosnian | Indo-European | Slavic | Strong |
| Bulgarian | Indo-European | Slavic | Strong |
| Cantonese | Sino-Tibetan | Chinese | Weak |
| Catalan | Indo-European | Romance | Strong |
| Cebuano | Western Malayo-Polynesian | Meso-Philippine | Weak |
| Chaha | Afro-Asiatic | Semitic | Strong |
| Chichewa | Niger-Congo | Bantoid | Strong |
| Croatian | Indo-European | Slavic | Strong |
| Czech | Indo-European | Slavic | Strong |
| Dagbani | Niger-Congo | Gur | Strong |
| Danish | Indo-European | Germanic | Weak |
| Dutch | Indo-European | Germanic | Weak |
| Dyula | Niger-Congo | Western Mande | Weak |
| English | Indo-European | Germanic | Strong |
| Estonian | Finno-Ugric | Finnic | Weak |
| Ewe | Niger-Congo | Kwa | Strong |
| Finnish | Finno-Ugric | Finnic | Weak |
| Flemish | Indo-European | Germanic | Weak |
| French | Indo-European | Romance | Strong |
| Frisian | Indo-European | Germanic | Weak |
| Fula | Niger-Congo | Northern Atlantic | Strong |
| Gamo | Afro-Asiatic | North Omotic | Strong |
| Galician | Indo-European | Romance | Strong |
| Georgian | Kartvelian | Kartvelian | Strong |
| German | Indo-European | Germanic | Weak |
| Greek | Indo-European | Greek | Strong |
| Guarani | Tupian | Tupi-Guarani | Strong |
| Gujarati | Indo-European | Indic | Strong |

**Appendix Table: Coded Languages and FTR Values (Continued)**

| Language | Family | Genus | FTR |
|---|---|---|---|
| Hakka | Sino-Tibetan | Chinese | Weak |
| Hausa | Afro-Asiatic | West Chadic | Strong |
| Hawaiian | Eastern Malayo-Polynesian | Oceanic | Weak |
| Hebrew | Afro-Asiatic | Semitic | Strong |
| Hindi | Indo-European | Indic | Strong |
| Hungarian | Finno-Ugric | Ugric | Strong |
| Icelandic | Indo-European | Germanic | Weak |
| Igbo | Niger-Congo | Igboid | Strong |
| Irish | Indo-European | Celtic | Strong |
| Isekiri | Niger-Congo | Defoid | Strong |
| Indonesian | Western Malayo-Polynesian | Sundic | Weak |
| Italian | Indo-European | Romance | Strong |
| Japanese | Japanese | Japanese | Weak |
| Javanese | Western Malayo-Polynesian | Sundic | Weak |
| Kammu | Austro-Asiatic (Mon-Khmer) | Palaung-Khmuic | Strong |
| Kannada | Dravidian | Southern Dravidian | Strong |
| Karaim | Altaic | Turkic | Strong |
| Korean | Korean | Korean | Strong |
| Kikuyu | Niger-Congo | Bantoid | Weak |
| Kurdish | Indo-European | Iranian | Strong |
| Latvian | Indo-European | Baltic | Strong |
| Lithuanian | Indo-European | Baltic | Strong |
| Lozi | Niger-Congo | Bantoid | Strong |
| Luganda | Niger-Congo | Bantoid | Strong |
| Luxembourgish | Indo-European | Germanic | Weak |
| Malay | Western Malayo-Polynesian | Sundic | Weak |
| Maltese | Afro-Asiatic | Semitic | Weak |
| Macedonian | Indo-European | Slavic | Strong |
| Mandarin | Sino-Tibetan | Chinese | Weak |
| Maori | Western Malayo-Polynesian | Oceanic | Weak |
| Moldavian | Indo-European | Romance | Strong |
| Montenegrin | Indo-European | Slavic | Strong |
| Moore | Niger-Congo | Gur | Strong |
| Norwegian | Indo-European | Germanic | Weak |
| Oromo | Afro-Asiatic | Cushitic | Weak |
| Panjabi | Indo-European | Indic | Strong |
| Persian | Indo-European | Iranian | Strong |
| Polish | Indo-European | Slavic | Strong |
| Portuguese | Indo-European | Romance | Strong |
| Quechua | Quechuan | Quechuan | Strong |
| Romanian | Indo-European | Romance | Strong |
| Romansh | Indo-European | Romance | Strong |
| Russian | Indo-European | Slavic | Strong |

**Appendix Table: Coded Languages and FTR Values (Continued)**

| Language | Family | Genus | FTR |
|---|---|---|---|
| Serbian | Indo-European | Slavic | Strong |
| Slovak | Indo-European | Slavic | Strong |
| Slovene | Indo-European | Slavic | Strong |
| Soddo | Afro-Asiatic | Cushitic | Weak |
| Sotho (Northern) | Niger-Congo | Bantoid | Strong |
| Seraiki | Indo-European | Indic | Strong |
| Sesotho | Niger-Congo | Bantoid | Strong |
| Sidamo | Afro-Asiatic | Cushitic | Weak |
| Spanish | Indo-European | Romance | Strong |
| Sumatranese | Western Malayo-Polynesian | Sundic | Weak |
| Sundanese | Western Malayo-Polynesian | Sundic | Weak |
| Swati | Niger-Congo | Bantoid | Strong |
| Swedish | Indo-European | Germanic | Weak |
| Swahili | Niger-Congo | Bantoid | Strong |
| Swiss French | Indo-European | Romance | Strong |
| Swiss German | Indo-European | Germanic | Weak |
| Swiss Italian | Indo-European | Romance | Strong |
| Tagalog | Western Malayo-Polynesian | Meso-Philippine | Strong |
| Tamil | Dravidian | Southern Dravidian | Strong |
| Tenyer | Niger-Congo | Gur | Strong |
| Thai | Tai-Kadai | Kam-Tai | Strong |
| Tigrinya | Afro-Asiatic | Semitic | Strong |
| Tsonga | Niger-Congo | Bantoid | Strong |
| Tswana | Niger-Congo | Bantoid | Strong |
| Turkish | Altaic | Turkic | Strong |
| Ukrainian | Indo-European | Slavic | Strong |
| Urdu | Indo-European | Indic | Strong |
| Uzbek | Altaic | Turkic | Strong |
| Venda | Niger-Congo | Bantoid | Strong |
| Vietnamese | Austro-Asiatic (Mon-Khmer) | Viet-Muong | Weak |
| Wolaytta | Afro-Asiatic | North Omotic | Strong |
| Wolof | Niger-Congo | Northern Atlantic | Strong |
| Xhosa | Niger-Congo | Bantoid | Strong |
| Yoruba | Niger-Congo | Defoid | Weak |
| Zulu | Niger-Congo | Bantoid | Strong |