Ca' Foscari University of Venice

Department of Economics

Working Paper

Juliana Bernhofer, Francesco Costantini and Matija Kovacic,

Risk attitudes, investment behavior and linguistic variation

ISSN: 1827-3580
No. 34/WP/2015

Risk attitudes, investment behavior and linguistic variation

*Juliana Bernhofer, Francesco Costantini and Matija Kovacic**

February 12, 2019

Abstract

This paper explores the relationship between linguistic variation and individual attitudes toward risk and uncertainty. We propose an innovative marker that classifies languages according to the number of non-indicative moods in the grammatical contexts involving uncertainty. We find that speakers of languages that use these moods more intensively are on average more risk averse. Our marker is then used to instrument risk aversion in the model for financial asset accumulation. In addition, by using Chen (2013)'s FTR linguistic marker as a proxy for the subjective discount rate, we disentangle the effects of risk aversion and time preferences on asset accumulation.

Keywords: Language, Risk Aversion, Time Preferences, Assets Accumulation.

JEL Classification: D14, D81, D91.

*M. Kovacic: Ca' Foscari University of Venice, Department of Economics, Cannaregio 873, 30121 Venice, E-mail: matija.kovacic@unive.it, Phone: +39 041 234 9138. J. Bernhofer: Faculty of Economics and Management, Free University of Bozen-Bolzano, Piazza Universita' 1, 39100 Bolzano, E-mail: juliana.bernhofer@unibz.it; F. Costantini: University of Udine, Department of Humanities, Via Petracco 8, 33100 Udine, E-mail: francesco.costantini@uniud.it

1 Introduction

Consider a situation in which you have to take a decision about something that has an uncertain prospect, may it be related to sports, health or financial choices. Your personal characteristics and preferences impact on how you evaluate the potential outcomes, but would you also argue that the language you speak may influence the way you perceive risk? In this paper we propose an innovative approach to analyze individual attitudes toward uncertainty and risky behavior based on the hypothesis of linguistic relativity. The concept stems from the idea that differences in grammatical structures and the vocabulary may induce speakers of distinct languages to experience the world differently (Hill and Mannheim, 1992). Research in support of this hypothesis has mainly focused on conceptual contents of languages linking for instance individual perceptions of space, color and even cross-country differences in gender political quota, as well as female labor force participation to specific linguistic features (see Majid et al. (2004); Davies and Corbett (1997), Roberson et al. (1999) and Winawer et al. (2007); Santacreu-Vasut et al. (2013) and Gay et al. (2017), respectively).

If speakers of different languages vary in their *worldview* depending on the language they use, some dimensions of linguistic structures may also shape individual preferences and their economic decision-making. In a recent paper on the effect of language on economic behavior, Chen (2013) tests a linguistic-savings hypothesis: when people are grammatically required to speak in a distinct way about future events, they take fewer future-oriented actions. The author adopts a criterion which separates languages into two broad categories: weak and strong Future Time Reference (FTR henceforth) according to how they induce speakers to mark the timing of events. Some languages require an explicit verb conjugation in order to distinguish between present and future events (strong FTR languages), while others allow their speakers to talk about the future by using the same verb form as for present events (weak FTR languages). The author then examines how these differences correlate with future-oriented behavior and finds that speakers of weak FTR languages save more, accumulate more wealth by retirement, smoke less frequently and are more physically active. This evidence remains reasonably robust even after controlling for geographic and historical relatedness of languages (Roberts et al., 2015).

On the other hand, Galor and Özak (2016) and Sarid et al. (2017) argue that ancestral characteristics from the parental country of origin might have affected the formation of time preferences and triggered the gradual emergence of grammatical forms that fostered the transmission of these traits across generations. Indeed, the authors show that higher pre-industrial crop yield potential experienced by ancestral populations had a positive impact on the descendants' future orientation. Nevertheless, the language still had an independent effect on time preferences and economic behavior.

Our approach is conceptually in line with Chen (2013) since we rely on a weak version of the linguistic relativity hypothesis. However, it departs from Chen (2013) for two reasons. First, we propose to consider the linguistic relativity hypothesis on the background of a different grammatical property and in a different economic context, namely *mood* and *uncertainty*. We develop a new linguistic marker based on the number of grammatical contexts concerned with the expression of uncertainty in which specific non-indicative moods (*i.e.* subjunctive, conditional, etc.) are used. Since indicative moods are generally used to assert that a fact or a situation is true as of the actual world, we conjecture that the perceived degree of uncertainty is larger with a non-indicative mood compared to an indicative one. Therefore, based on a weak version of the linguistic relativity hypothesis, speakers of languages where these specific grammatical forms are used more often, should experience the world as being more mutable and uncertain compared to speakers of languages where these forms are less frequent. Our mapping offers a rigorous and, to the best of our knowledge, the first linguistic mapping related to the grammatical treatment of uncertainty. Second, we analyze the correlation between our linguistic marker and individual self-declared risk aversion and use the marker to instrument the individual attitudes towards risk in order to quantify a direct effect of risk aversion on the probability of holding risky financial assets. In addition, we estimate the separate effects of time preferences and risk attitudes by using the FTR parameter as a proxy for intertemporal choice behavior, alongside our instrument for individual risk preferences.

Using data on a large set of individuals from the Survey of Health, Aging and Retirement in Europe (SHARE), we show that a more intensive use of non-indicative moods in grammatical contexts involving uncertainty strongly correlates with individual risk preferences. The likelihood of individuals being highly risk averse increases with the frequency of use of these forms in their respective languages. This evidence holds both within linguistically heterogeneous countries, and among the population of first-generation immigrants when they are assigned their original language. Intensive users are on average 43% more likely to report a high level of risk aversion compared to individuals speaking languages where these forms are used less frequently or where they are not required at all. This evidence is robust to the inclusion of agro-climatic variables for the parental country of origin from Galor and Özak (2016), as well as linguistic family controls. Moreover, the effect of language is not attenuated by controls for the degree of trust in others, family composition and occupational status, which suggest that these effects are independent.

As for the separate effects of risk and time preferences on the propensity to invest in financial risky assets, we find that highly risk averse individuals are, on average, 13 percentage points less likely to hold stocks or bonds with respect to intermediate risk takers and to individuals with a low level of risk aversion. When compared to individual time preferences, the effect of risk aversion is four times larger than the effect of the individual discount rate.

The paper is structured as follows. In the next section we introduce the issue of linguistic relativity and mood, as well as a discussion of the typological distinction used in Chen (2013). In Section 3 we exploit the relationship between our linguistic marker and individual attitudes toward risk and estimate a direct effect of risk aversion and time preferences on the probability of investing in risky financial assets. Section 4 concludes.

2 Linguistic Relativity and Economic Behavior

The idea that language categories can influence thought has come to be known as *Sapir-Whorf hypothesis* after Sapir (1921) and Whorf and Carroll (1964) and boasts a long history in the philosophy of language and linguistics which can be traced back at least to Humboldt's (1836) idea of *Innere Sprachform*. Following Geeraerts and Cuyckens (2010), the hypothesis of linguistic relativity encompasses two basic notions: the first being that languages are relative as they vary in their expression of concepts, and the second being that the semantic expression of concepts influences, at least to some extent, conceptualization at the cognitive level. Therefore, speakers of distinct languages may perceive reality differently.

The linguistic relativity hypothesis has generally been interpreted according to two versions. The strong one, also known as *linguistic determinism*, states that linguistic categories control general cognitive processes. This version of the hypothesis, however, has generally been refuted (Pinker, 1994). The weak version claims that linguistic categories have some effect on cognitive habits, particularly with respect to memory and categorization. The latter version of the Sapir-Whorf hypothesis was taken to be more feasible and has inspired research on topics such as color perception, shape classification, conditional reasoning, number, space, and time categorization.

If speakers of different languages tend to think and behave differently depending on the language they use, some dimensions of linguistic structures may also shape individual preferences. Chen (2013) represents the first attempt to analyze the impact of language differences on the cognitive domain and consequently on several aspects of individual economic behavior. The empirical analysis in Chen (2013) uses a typological distinction discussed in Dahl (2000) and Thieroff (2000) whereby there are languages that employ specific verb morphology for FTR, whereas other languages do not. By adopting the weak version of the Sapir-Whorf hypothesis, Chen (2013) hypothesized that this typological divide affects how speakers conceive time. Specifically, speakers of languages that separate the future from the present tense ("strong FTR" languages) are more prone to dissociate the future from the present compared to speakers of languages that do not employ that specific verb morphology when referring to future events ("weak FTR" or "futureless" languages). As a consequence, this may induce people to perceive the future as being more distant and, hence, to undertake fewer future-oriented actions such as saving, smoking, using condoms, accumulating wealth before retirement, and taking initiatives to enhance long-run health. The association between weak FTR and future-oriented behavior in Chen (2013) is strong: speakers of weak FTR languages save more, accumulate more wealth by retirement, smoke less frequently and are more physically active.

We propose a reconsideration of Chen (2013)'s idea of linking language features to economic behavior through the linguistic relativity hypothesis. Following a weak version of the hypothesis, we conjecture that individual levels of risk aversion are influenced by differences in the intensity of use of indicative versus non-indicative (*i.e., irrealis*) moods as they assign a different degree of uncertainty to possible situations. In other words, when describing possible or hypothetical situations, the displacement of the actual from the alternative state of facts is perceived as larger when an *irrealis* mood is used. According to this conjecture, in sentences (1) and (2), for example, the leaving event should be perceived as less uncertain by an English speaker than by an Italian speaker, even though they describe the same possible situation.

(1) *I think he has left.* (English)

(2) *Penso sia partito.* (Italian)
Think-1SG is-SUBJ left
"I think he has left"

The former expresses the leaving situation by resorting to the indicative mood, while the latter has to use a subjunctive (*irrealis*) mood. In general, by using *irrealis* more intensively, speakers move from the region of certainty to that of uncertainty, in other words, their latent area of the unknown is greater with respect to their peers who speak a less *irrealis*-intensive language. As a consequence, they are expected to be more risk averse as the semantic salience of their region of uncertainty increases.

For this purpose, we develop a specific linguistic marker defined on the number of non-indicative moods used in *irrealis* contexts, *i.e.*, contexts that involve grammatical categories concerned with the expression of uncertainty and we relate it to the individual's perception of risk and risky behavior. In what follows we describe the definitions of displacement, modality and mood more in depth, providing also some applied examples and contexts that define our linguistic marker.

2.1 Displacement and Modality

By displacement semanticists mean the specific characteristic of human language whereby language expressions do not only refer to the here and now, but are able to range over future, past, potential, possible and even impossible situations (Hockett, 1960; Hockett and Altmann, 1968). In that sense, futurity is an instance of displacement within the temporal dimension. Another crucial dimension of displacement is modality, the grammatical category that indicates whether a sentence expresses a fact, a command, a condition, an opinion or a desire. Consider for instance the following sentences:[1]

(1)	*Wenn*	*es*	*sonnig*	*wäre,*	*ginge*	*ich*	*spazieren.*	(German)
	If	it	sunny	be-**KONJ**	go-**KONJ**	I	walk	
	"If it were sunny, I would go for a walk."							
(2)	*Penso*	*che*	*la*	*riunione*	*sia*	*finita.*		(Italian)
	Think-**1SG**	that	the	meeting	is-**SUBJ**	finished		
	"I think the meeting has finished."							
(3)	*Chodżmy*	*do*	*mnie*	*na*	*kawę.*			(Polish)[2]
	Go-**IMP**	to	me	for	coffee			
	"Let's go to my place for a coffee."							

By observing sentences (1), (3), and the embedded clause in (2), we notice that they do not describe actual facts. The truth or falsity of the expressions cannot be decided simply by considering whether the state of facts described in the sentences is true (or false). Sentence (1) does not assert that it is sunny and that the speaker is having a walk, sentence (2) does not assert that the meeting is finished. It may be finished, and the speaker probably believes that it has, but one's belief may turn out to be wrong when actual states of facts are taken into consideration. Sentence (3) does not assert that the speaker is at home having a coffee with the hearer. Sentences (1) to (3) do not refer to actual facts, differently from sentences like "It is sunny today", "I am having a walk", "The meeting has finished", "I'm having a coffee at home with a friend". They refer to possible situations or "possible worlds" (Carnap, 1947), not to real ones. Possible worlds represent alternative states of facts, which cannot be asserted as of the world we actually live in (the "actual world"), and as such they involve the notion of uncertainty.

2.2 Mood and *Irrealis* context of use

Mood is the grammatical category concerned with the expression of situations involving the "world" parameter. What grammarians call *indicative* is the mood generally used to assert that a proposition is true as of the actual world.[3] To express possible situations

1 "KONJ" stems for German *Konjunktiv*; "1SG" for *First Singular*, and "IMP" for *Imperative*.
2 Swan (2002), pp 242.
3 This does not exclude that the indicative may have modal functions, too.

languages can use moods other than the indicative. In sentence (1), for instance, the verbs are in the so-called *Konjunktiv II*. The embedded clause in (2) is in the *Subjunctive*. Sentence (3) is in the *Imperative*. In sentence (2), the English language uses an indicative while Italian uses a non-indicative mood (subjunctive).

Some languages have a wide range of morphological moods, some, the most in fact, have a limited number of grammatical categories concerning mood, which are basically the indicative, the imperative and the subjunctive/conditional and others do not have any specific morphological markers for mood[4]. Most importantly, languages may vary as for the contexts of use of different moods. While in all languages the indicative is the mood used to assert a state of fact and imperative the one to command, the other moods (*i.e.*, subjunctive, conditional, etc.) have different functions and may be used in contexts that vary from language to language. The contexts where *irrealis* moods are used more consistently from a cross-linguistic viewpoint include the following:

– complements of modal predicates (e.g. to be possible, to be likely, to be necessary, to be probable);

  It is probable that these events were coincidences.

– complements of volitional predicates (e.g. to want, to wish, to desire);

  I wish I hadn't been late for school.

– complements of epistemic (non-factive) predicates (e.g. to think, to believe, to doubt);

  I think we should keep a diverse energy portfolio.

– complements of emotive factive predicates (e.g. to regret, to be happy, to be sad);

  I regret that this joke has garnered so much attention.

– complements of declarative predicates (e.g. to say, to tell, to announce);

[4] With the exclusion of nonfinite moods, like the infinitive or the gerund, most Romance languages have four moods according to traditional grammars: the indicative, the subjunctive, the conditional and the imperative. Most Slavic languages have three moods: the indicative, the conditional and the imperative. German has three moods, too: the indicative, the "Konjunktiv" and the imperative. Northern Germanic languages have only two moods: the indicative and the imperative. Subjunctive mood is also mentioned in some traditional grammars, but it has only residual uses and is no longer productive.

I said that one day in my career bad results will come.

– the protasis (the if-clause) in conditional sentences;

If he had studied harder, he would have passed the exam.

– the apodosis (the main clause) in conditional sentences.

If he had studied harder, he would have passed the exam.

For the purpose of our index, we take the extent of use of different *irrealis* moods in these syntactic contexts. We assign a value of 1 to the occurrence of a non-indicative mood in a particular syntactic environment and 0 otherwise. By addition we obtain an indicator (IRR henceforth) of how frequently *irrealis* forms are used in a language, so that languages can be ranked according to the intensity of use of *irrealis* moods[5]. Finally, languages that do not require *irrealis* moods in any of the context above are called "moodless" languages.

Our linguistic mapping covers 38 languages as listed in Table 13 in Appendix A. Data on grammatical moods were mainly collected from Rothstein and Thieroff (2010) (RT henceforth) as it is the most comprehensive typological survey on grammatical moods in the languages of Europe. Since not all the data we needed were included in RT, we also collected some additional primary data. For this purpose, we worked out a questionnaire compiled by a number of linguists throughout Europe[6]. We contacted linguistic experts who were asked to provide a translation of various sentences into their native language and to produce, for each sentence, explanations on which mood they were using in their versions (indicative versus other non-indicative moods to be described).

Table 13 in Appendix A shows that six languages are moodless, whereas three languages use *irrealis* moods in all of the six contexts. The remaining 29 languages range from two to four contexts in which they employ *irrealis* moods. Thus, significant variation of IRR across languages may represent a good platform for testing the linguistic relativity hypothesis in the context of economic behavior involving risk and uncertainty.

[5] Since there is no qualitative difference between contexts in defining IRR, the index sum was calculated using a uniform weighting function.

[6] Full version available upon request.

3 Linguistic Variation, Risk Attitudes and Investment Behavior

Our main hypothesis is based on the assumption that speakers of languages in which non-indicative moods are used more intensively to express potential situations (H-type) perceive the world as being more mutable and hence more uncertain. As a consequence, they are expected to be more risk averse than those speaking a low-intensity IRR language (L-type), and to engage less in activities with an uncertain outcome.

In this section we empirically validate the following hypotheses:

Hypothesis 1 ***Direct Effect of Irrealis on Risk Aversion***
Speakers of languages characterized by a more pronounced displacement into uncertainty (H-type) are more risk averse with respect to speakers of languages with a weaker displacement into uncertainty (L-type), ceteris paribus.

Hypothesis 2 ***Indirect Effect of Irrealis on Investment Behavior***
Speakers of languages with a more pronounced displacement into uncertainty invest less in risky assets, ceteris paribus.

The intensity of displacement into uncertainty, hence, directly influences the individuals' attitudes toward risk, and indirectly their propensity to invest in risky assets.

The empirical examination of the proposed mechanism proceeds in two steps. We first estimate a direct association between our linguistic marker and the individuals' attitudes toward risk and then specify a two stage empirical model and use our linguistic marker to instrument risk aversion in the equation for the probability of holding risky assets. In addition, we disentangle the effects of risk and time preferences by using the FTR linguistic marker as a proxy for the individuals' subjective discount rate.

3.1 Data Description

Our empirical analysis is run on the pool of individuals in the Survey on Health, Ageing and Retirement in Europe (SHARE henceforth)[7], Wave 2, Wave 4 and Wave 5 - release

[7] SHARE is a multidisciplinary and cross-national panel database of micro data on health, socio-economic status and social and family networks of individuals aged 50 or older.

5.0.0.[8] We consider individuals for which we have complete information on the self-declared level of risk aversion, asset holdings, as well as on demographic, socio-economic, family, cognitive and health conditions. In addition to SHARE, we run some of our baseline regression models also on individuals from the European Social Survey (ESS henceforth).[9] We use SHARE as a primary source of data mainly for two reasons: (*i*) the features of risk preferences elicited in SHARE fit particularly well with the nature of our research question since they rely on financial and not on adventure risk taking like in ESS,[10] (*ii*) differently from SHARE, the ESS survey does not contain any information on the respondents' asset holdings.

3.1.1 Sample selection

In order to analyze the relationship between our linguistic marker and individual attitudes toward risk, we consider the full sample of individuals living in linguistically heterogeneous and the population of first-generation immigrants separately. Within their host countries, non-native individuals have experienced and face the same economic and political environment and institutions but differ in their linguistic heritage. If we assume that such linguistic traits are persistent, then according to the hypothesis of linguistic relativity, different linguistic backgrounds will make them behave and perceive the world differently in a similar environment. The pool of first-generation immigrants in SHARE contains 11076 individuals for which we have complete information on their country of origin, linguistic background, the total amount of time spent living in the host country, as well as on other explanatory and control variables.[11] The full sample of individuals, on the other hand, counts 41022 in linguistically heterogeneous ones.[12]

[8] The variable of financial risk preferences in SHARE is not present in Wave 1. Wave 3 is a retrospective survey with a different methodology.

[9] See Appendix C and Appendix D for a comprehensive list of countries and languages in SHARE and ESS.

[10] The same question related to adventure risk taking is also available in the World Value Survey (WVS). However, we do not consider this additional source of data because the information on both the immigrant status and origin, and adventure risk taking is available only in Wave 6 for a very restricted sample of individuals unevenly distributed across countries. See Appendix D for the question in ESS and Weber et al. (2002) for the definition of different risk-taking domains.

[11] The distribution of first-generation immigrants across countries can be found in Appendix A, Table 16.

[12] As an additional robustness check we also consider the subset of native individuals separately. These results are available upon request.

3.1.2 Language assignment

Regarding the language variable treatment in SHARE, respondents are not asked to declare the language they normally speak at home (as it was the case in ESS). As a consequence, we assume that the language in which the questionnaire is compiled is also the native individuals' primary language. Those living in countries with two or more official languages were given the choice between one language or the other.

As for the language assignment to non-native individuals, we use three different criteria: (*i*) the language of the immigrants' country of origin, (*ii*) the immigrant father's language of origin, and (*iii*) the immigrant mother's language of origin. As regards the immigrant's original language we follow Hicks et al. (2015) and consider the official language spoken in the country of birth, if available, or the official language spoken by more than 80% of the population (in all those cases where the country of birth has more than one official language). Immigrants from linguistically heterogeneous countries, such as Switzerland, Belgium, Spain or Canada are excluded from the analysis since we are not able to track their original language.[13] In addition, we exclude the languages spoken by less than 100 individuals.[14] The pool of first-generation immigrants in SHARE originates from 20 different linguistic backgrounds coming from 82 different countries for which the information on IRR is available. For the language assignment to the respondent's parents, we follow the same criteria as for the language assignment to the respondent's country of origin.

3.1.3 Data on risk attitudes and IRR

As for risk attitudes in SHARE, the respondents were asked to answer a simple question on financial risk tolerance:

[13] In Spain the main official language is Spanish, but a significant portion of the population speaks Catalan (17%), Basque (2%) or Galician (5%). In Belgium there are two official languages, French and Dutch, none of them spoken by more than 80% of the population. In Canada, the English language is the mother tongue for 60% of the population while 22% of all Canadians speaks French.

[14] The list of languages spoken by less than 100 immigrants in our sample includes: Albanian, Bulgarian, Danish, Greek, Hebrew, Icelandic, Latvian, Lithuanian, Macedonian, Slovenian, Swedish, and Norwegian.

When people invest their savings, they can choose between assets that give low return with little risk to lose money, for instance a bank account or a safe bond, or assets with a high return but also a higher risk of losing, for instance stocks and shares. Which of the statements on the card comes closest to the amount of financial risk that you are willing to take when you save or make investments?

(1) Take substantial financial risk expecting to earn substantial returns;

(2) Take above average financial risks expecting to earn above average returns;

(3) Take average financial risk expecting to earn average returns;

(4) Not willing to take any financial risk.

Individuals who answered (1) and (2) have a greater tolerance of volatility of return and hence a lower level of risk aversion. Intermediate risk takers are those who answered (3) while individuals who answered (4) are highly averse to financial risk. In our sample, 75.88% of individuals declare not to be willing to take financial risks, 19.93% are ready to take average financial risk, and only 4.19% are willing to take above average or substantial financial risk. The distribution of individual attitudes towards risk across first-generation immigrants is similar, with 81.53% of the respondents classifying as highly risk averse, 15.01% being intermediate risk-takers and 3.46% exhibiting a low level of risk aversion.

Figure 1 shows the distribution of individuals by IRR for the entire set of countries (left-hand side figure) and for a restricted set of linguistically heterogeneous countries (right-hand side figure).[15] Roughly 12% of the respondents are moodless speakers, 55% are intermediate IRR users, while 33% are intensive and very intensive IRR users.
Similarly, the majority of individuals in linguistically heterogeneous countries are intermediate IRR users, there are no moodless speakers, while 31% classify as intensive and very intensive IRR users.[16]

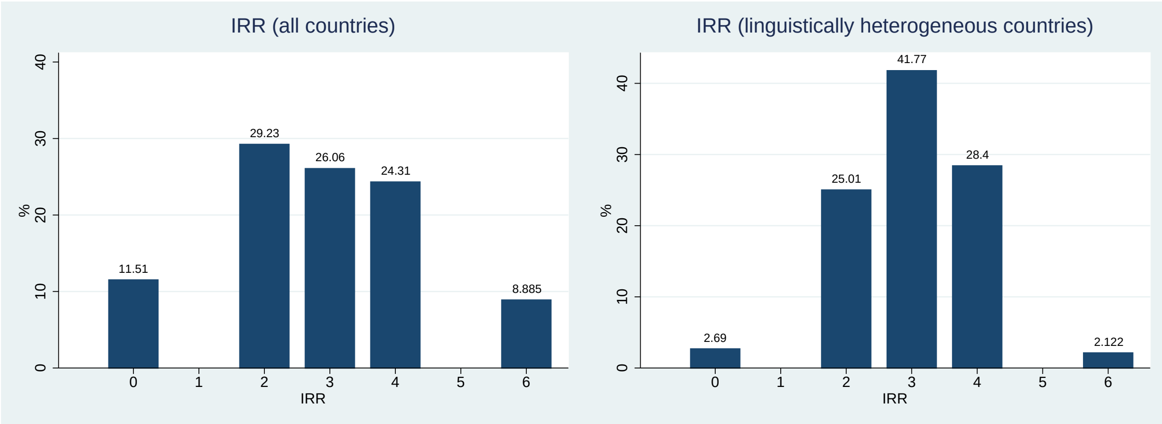
The distribution of IRR for the population of first-generation immigrants is presented in Figure 2. The majority of non-native individuals are intensive and very intensive IRR users, more than 27% are intermediate IRR users and only 3.2% are moodless speakers.

Finally, in Tables 1 and 2 we report the distribution of the three levels of risk aversion over the three different categories of IRR: CatIRR0 contains no IRR (0 IRR), CatIRR1

[15] There are in total 6 linguistically heterogeneous countries: Spain, Luxembourg, Belgium, Switzerland, Estonia and Israel. Distribution of IRR by country in Appendix A, Table 14.

[16] Italian and Portuguese are the only two European languages in our sample with the maximum number of IRR (6). Another European language not included in our sample with IRR=6 is Icelandic.

Figure 1: Distribution of IRR users in all countries and in linguistically heterogeneous countries.



Source: SHARE, Wave 2, Wave 4 and Wave 5. N. Observations: 118148 (all countries); 41229 (linguistically heterogeneous countries).

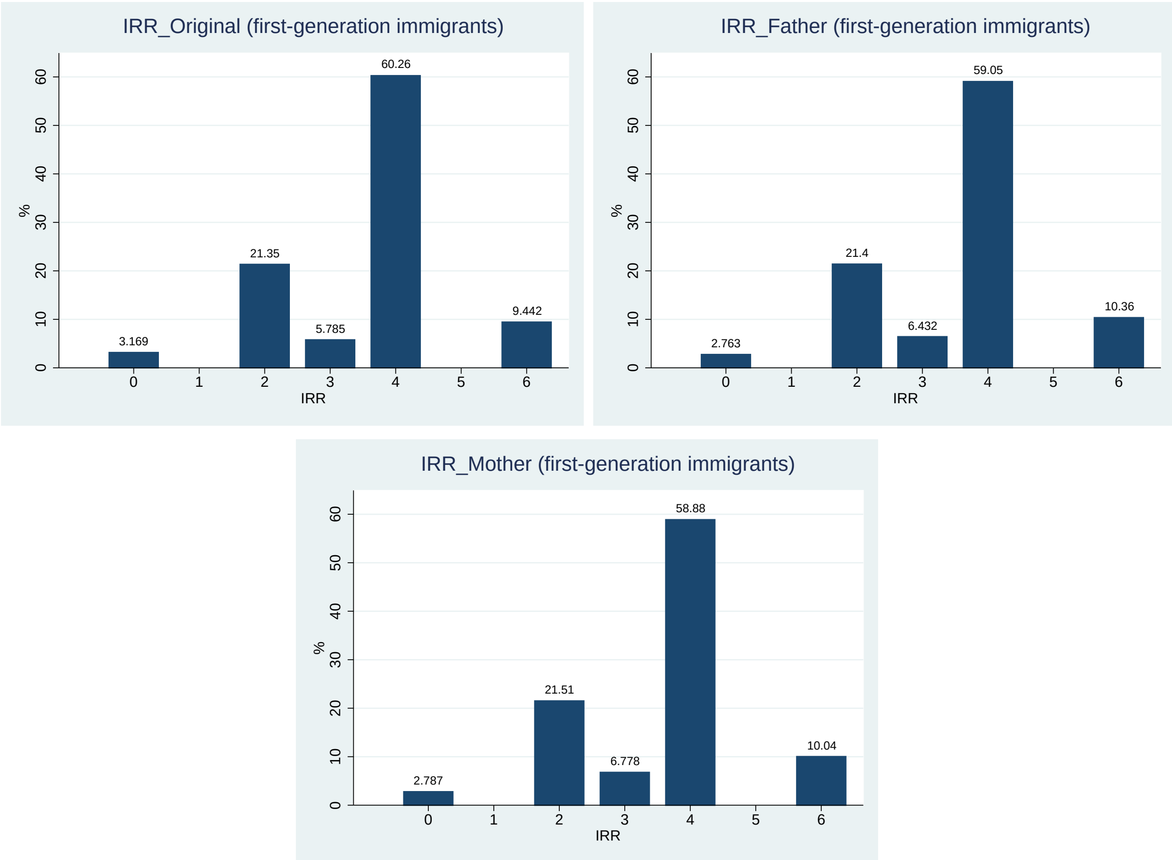
refers to intermediate IRR usage (2 or 3 IRR) and CatIRR2 represents an intensive and very intensive IRR usage (4 or 6 IRR).[17] Table 1 considers the full sample while Table 2 refers only to the population of non-native individuals. Roughly 38% of individuals with a low level of risk aversion also speaks a moodless language (CatIRR0). Half of the respondents with an intermediate risk attitude also classify as intermediate IRR users (CatIRR1). Finally, those who declare to be averse to any financial risk are mostly either intermediate (CatIRR1) or intensive IRR users (CatIRR2). As for the immigrant population, the entire distribution is skewed towards the second category of IRR (CatIRR2) which is a direct consequence of the general features of the IRR distribution across non-native speakers (Figure 2). However, more than 70% of those who declare to be highly risk averse are intensive IRR speakers (CatIRR2).

It is important to note that the distribution of IRR does not follow any marked territorial pattern (Table 3).[18] For instance, we find IRR = 4 users in the Mediterranean area, central, eastern and northern Europe, and in the Middle East. However, the moodless category (IRR = 0) contains only four countries. Given the low number of countries in this category, a comparison between individuals in these countries to those in any other

17 None of the languages in our dataset has IRR=1 and IRR=5. However, generally these values are admissible.

18 Linguistically heterogeneous countries may appear in more than one category whenever the languages spoken in these countries differ in terms of IRR.

Figure 2: Distribution of IRR users across first-generation immigrants.



Source: SHARE, N. Observations: 11046 (Original Language), 8505 (Father's Original Language), 8646 (Mother's Original Language).

IRR category will require some caution. We will turn to this point in the next section. The absence of a clear segregation of IRR along territorial, political, historical or any other path-dependent line represents an important feature of our marker, since we want to estimate a direct effect of language net of other unobserved and potentially overlapping non-linguistic factors.

3.2 Empirical Strategy: Irrealis and Risk Aversion

In order to test the association between IRR and individual attitudes toward risk (*Hypothesis 1*), we run two separate blocks of regressions. The first block considers the full sample of individuals in linguistically heterogeneous countries.[19] The second set of regressions follows the epidemiological approach to the identification of the effects of language

[19] We do not include linguistically homogeneous countries since all the effect of language in the pooled sample would be identified off of within-country variation in IRR that comes from linguistically heterogeneous countries.

Table 1: Risk Aversion by CatIRR (%). Full Sample.

	CatIRR0	CatIRR1	CatIRR2	Total
Low Risk Aversion	38.44	40.14	21.42	100.00
Intermediate Risk Aversion	19.00	54.50	26.50	100.00
High Risk Aversion	8.06	56.33	35.61	100.00
Total	11.51	55.29	33.20	100.00

Source: SHARE, Wave 2, Wave 4 and Wave 5.

Table 2: Risk Aversion by CatIRR (%). Non-Native Individuals.

	CatIRR0	CatIRR1	CatIRR2	Total
Low Risk Aversion	12.38	40.78	46.84	100.00
Intermediate Risk Aversion	9.32	37.86	52.81	100.00
High Risk Aversion	3.16	25.69	71.15	100.00
Total	4.42	28.06	67.52	100.00

Source: SHARE, Wave 2, Wave 4 and Wave 5.

and considers the subpopulation of first-generation immigrants. By moving from the first to the second block of regressions, we increase the variability of IRR within countries, and reduce potential concerns regarding omitted variables at the host country level. As a further attempt to overcome the problem of identification bias due to omitted variables, in line with Galor and Özak (2016) and Sarid et al. (2017) we consider a set of ancestral characteristics in the parental country of origin to control for a range of confounding factors that might have influenced the formation of risk preferences, independently from the effect of language. In all regression models we calculate the robust standard errors clustered at the country level.

The dependent variable RA_i is equal to 1 for an individual declaring to be averse to risk taking and 0 otherwise. The empirical problem consists in estimating the following logistic model:

$$P(RA_i) = \frac{exp(r_i)}{1 + exp(r_i)}$$

where:

$$r_i = \alpha + \beta IRR_i + \gamma X_i + \theta Z_i + \rho CW_i + \lambda FE_i + \eta_i. \quad (1)$$

Table 3: Sorting of Countries in IRR Classes

IRR=0	IRR=2	IRR=3	IRR=4	IRR=6
Sweden	Austria	Spain	Spain	Italy
Denmark	Germany	France	Israel	Switzerland
Israel	Netherlands	Switzerland	Czech Republic	Portugal
Ireland	Greece	Belgium	Poland	
	Switzerland	Luxembourg	Hungary	
	Belgium	Slovenia	Estonia	
	Luxembourg	Estonia		

Source: SHARE, Wave 2, Wave 4 and Wave 5.

Our main variable of interest IRR_i denotes the number of non-indicative moods in *irrealis* contexts in the individual $i's$ language. X_i is the vector of demographic and socio-economic characteristics, *i.e.,* age, gender, marital status (married or in registered partnerships, divorced or separated, widowed), number of children, occupational status (employed, unemployed, retired, homemaker or disabled), educational attainment (low, medium, high), household income, ownership (housing) and, for first-generation immigrants, the continent of origin and the time spent living in the host country. Z_i contains controls for cognitive ability and literacy (reading and writing abilities), a dummy variable whether the respondent thinks others are generally trustworthy, and (objective) health conditions (ability to perform daily activities). CW_i contains country-wave fixed effects and, for the subsample of first-generation immigrants, also a complete set of geographical and agricultural factors from Galor and Özak (2016), immigrants' continent of origin, and controls for linguistic families.[20] In addition, in line with Chen (2013) and Roberts et al. (2015), we also consider an additional set of fixed-effects, FE_i, for individual demographic and socio-economic characteristics divided into a set of exogenous determinants (Age and Gender, Country and Wave) and several groups of factors potentially endogenous to individual preferences toward risk and risky behavior (Income and Education, Marital Status and Number of Children). By considering this additional set of controls, we are able to compare individuals living in the same country and who are identical on

[20] See Appendix C for the summary statistics and the description of selected explanatory and control variables.

these dimensions but differ in their IRR usage. These regressions are estimated by means of a fixed-effect (conditional) logistic model.

3.2.1 Results: Full sample

Empirical estimates of equation (1) for the full sample of individuals are presented in Tables 4 and 5. Since the intensity of IRR usage represents a prominent feature of our linguistic marker, we mainly concentrate on IRR as a limited discrete (*i.e.*, non-categorized) variable ranging from 0 to 6 on the *irrealis* intensity scale. This allows us to estimate the average effect of a gradual increase in IRR and not merely the effects of one particular IRR category (for instance, high intensity IRR speakers) *versus* another. However, for an easier interpretation of the estimated coefficients, we also consider two alternative categorizations of IRR. All the coefficients are reported as odds ratios.

Table 4 shows the estimation of the probability of being highly risk averse for the full sample of individuals living in 5 European linguistically heterogeneous countries and Israel. Models RA1 - RA4 consider a non-categorized version of IRR. In model RA5 we group the respondents into three different IRR intensity classes and use intensive and very intensive IRR speakers (IRR = 4 and IRR = 6) as a reference category. Finally, in model RA6 we compare intensive and very intensive IRR users with the pool of moodless (IRR = 0) and intermediate IRR speakers (IRR = 2 and IRR = 3). We do not consider IRR = 0 as a reference category since the incidence of moodless speakers in our sample is low (3.21%).

The correlation between IRR and risk aversion is significant in all model specifications. The average effect of a gradual increase in IRR on risk aversion in model RA4 is 12.7% and is significant at the 1% level. Intermediate and moodless IRR speakers are respectively 35% and 37% less likely to be highly risk averse than intensive IRR users (model RA5) which have a 55% higher probability of reporting risk aversion compared to the entire pool of individuals speaking languages with lower IRR intensities (model RA6).

Table 4: Logit Model: probability of being Highly Risk Averse. Linguistically Heterogeneous Countries. Full Sample.

Highly Risk Averse (d)	RA 1	RA 2	RA 3	RA 4	RA 5	RA 6
IRR	1.239***	1.165***	1.151***	1.127***		
	(0.026)	(0.022)	(0.020)	(0.017)		
Cat_IRR0					0.627***	
					(0.058)	
Cat_IRR1					0.649***	
					(0.049)	
Strong_IRR						1.553***
						(0.098)
Age	1.036***	1.025***	1.018***	1.014***	1.014***	1.014***
	(0.008)	(0.008)	(0.005)	(0.004)	(0.004)	(0.004)
Female	1.978***	1.824***	1.882***	1.902***	1.905***	1.905***
	(0.077)	(0.061)	(0.090)	(0.091)	(0.089)	(0.090)
Low Education		1.797***	1.771***	1.638***	1.638***	1.637***
		(0.072)	(0.059)	(0.061)	(0.065)	(0.065)
High Education		0.558***	0.580***	0.622***	0.623***	0.623***
		(0.054)	(0.054)	(0.058)	(0.057)	(0.057)
Income		0.909***	0.913***	0.917***	0.918***	0.918***
		(0.013)	(0.013)	(0.012)	(0.012)	(0.012)
Owner		0.668***	0.678***	0.688***	0.688***	0.688***
		(0.030)	(0.030)	(0.027)	(0.027)	(0.027)
Married			1.111	1.115	1.106	1.106
			(0.078)	(0.081)	(0.080)	(0.080)
Num. Children			0.964	0.962	0.964	0.964
			(0.024)	(0.024)	(0.023)	(0.023)
Trust			0.948***	0.953***	0.952***	0.952***
			(0.004)	(0.004)	(0.005)	(0.005)
Retired			1.322***	1.306***	1.299***	1.299***
			(0.142)	(0.121)	(0.115)	(0.115)
Unemployed			1.383***	1.366**	1.350**	1.350**
			(0.158)	(0.166)	(0.165)	(0.164)
Disabled			1.580***	1.361***	1.353***	1.353***
			(0.139)	(0.085)	(0.084)	(0.084)
Homemaker			1.144**	1.103	1.085	1.085
			(0.060)	(0.062)	(0.061)	(0.061)
Country x Wave	Yes	Yes	Yes	Yes	Yes	Yes
Cognitive, Health	No	No	No	Yes	Yes	Yes
N. Observations	42376	42376	41041	41022	41022	41022
R Sq.	0.105	0.152	0.155	0.159	0.160	0.160
N. Countries	6	6	6	6	6	6

Notes: *The dependent variable is "Highly Risk Averse (d)". The method of estimation is Logit with the coefficients reported as odds ratios. Robust standard errors are clustered at the country level. Reference categories for dichotomous variables: IRR = 4 or IRR = 6 (model RA5), IRR < 4 (model RA6), Male, Not Married (divorced, separated, widowed), Medium Education, Employed or Self-Employed, Do not own their house/flat.*

The effects of other determinants of risk aversion remain significant even after controlling for the effect of language. In line with the empirical literature in the field (Hartog et al. (2002), Bellante and Green (2004), Cohen and Einav (2007), Lin (2009), Dohmen et al. (2011), we find that individuals who generally believe others to be trustworthy are on average less risk averse, even though this effect is not very strong. Females are on average more risk averse than men, higher levels of education are associated with lower risk aversion, while wealthier individuals are on average less risk averse than poorer ones. Regarding the occupational status, employment correlates significantly with the level of individual risk attitudes with unemployed and retired individuals being significantly more risk averse than employed and self-employed ones.

In Table 5 we estimate the conditional logit model with a set of additional individual-specific fixed-effect controls. Models RA1 and RA2 include fully interacted fixed effects for age, gender, income, education, country and wave, while Models RA3 - RA5 also add fixed effects for family structure. The results show that even when comparing individuals that are identical on every dimension defined by our fixed effects, a higher intensity of IRR is associated with a significantly higher probability of being averse to taking risks. For instance, individuals identical in all aspects except for the intensity of IRR in their respective languages, differ significantly in terms of their risk preferences, with speakers of intensive IRR languages being 46% more likely to be highly averse to risk taking compared to identical moodless and intermediate IRR users (model RA5).

The inclusion of the fixed effect for family structure in model RA3, and cognitive and health controls in models RA3 and RA4 slightly reduces the estimated effect of IRR on risk aversion. On the other hand, by comparing models RA3 and RA4, we see that the inclusion of trust does not alter the estimated coefficient of IRR, which suggests that the attitude of trusting other people is independent of the specific linguistic features captured by IRR.

The effects of the employment status are significant and relatively large in magnitude. Being unemployed, for instance, is associated with 59% higher probability of being averse to risk taking, which is roughly 13 percentage point more than the effect of intensive and very intensive IRR.

Table 5: Conditional Logit model: probability of being Highly Risk Averse. Linguistically Heterogeneous Countries. Full Sample.

Highly Risk Averse (d)	RA 1	RA 2	RA 3	RA 4	RA 5
IRR	1.168***	1.162***	1.123***	1.121***	
	(0.021)	(0.022)	(0.033)	(0.030)	
Strong_IRR					1.463***
					(0.083)
Retired	1.346***	1.359***	1.309***	1.304***	1.300***
	(0.097)	(0.096)	(0.054)	(0.057)	(0.055)
Unemployed	1.632***	1.572***	1.610**	1.595**	1.588**
	(0.160)	(0.159)	(0.317)	(0.319)	(0.316)
Disabled	1.624***	1.570***	1.427***	1.425***	1.419***
	(0.120)	(0.124)	(0.130)	(0.132)	(0.131)
Homemaker	1.129***	1.157***	1.063	1.057	1.044
	(0.035)	(0.034)	(0.064)	(0.064)	(0.060)
Owner		0.698***	0.721***	0.725***	0.722***
		(0.019)	(0.041)	(0.040)	(0.041)
Trust				0.952***	0.950***
				(0.006)	(0.006)
Fixed Effects:					
Sex x Age	Yes	Yes	Yes	Yes	Yes
Country x Wave	Yes	Yes	Yes	Yes	Yes
Income x Education	Yes	Yes	Yes	Yes	Yes
MarStatus x Num.Child	No	No	Yes	Yes	Yes
All FE Interacted	Yes	Yes	Yes	Yes	Yes
Additional Controls:					
Cognitive x Health	No	No	Yes	Yes	Yes
N. Observations	29569	29569	13664	13526	13526
N. Countries	6	6	6	6	6

Notes: *The dependent variable is "Highly Risk Averse (d)". The method of estimation is Conditional Logit Model with coefficients reported as odds ratios. Robust standard errors are clustered at the country level. Reference categories: IRR $<$ 4, Employed or Self-Employed, Do not own their house/flat.*

3.2.2 Results: First-Generation Immigrants

Our second block of regressions represents a further attempt to isolate the effect of language on individuals' risk preferences by analyzing the differences in risk attitudes across first-generation immigrants. In this way we compare individuals with different cultural and linguistic backgrounds living in the same economic, political and institutional context independent of their country of origin. Moreover, by considering this particular subset of individuals, we significantly increase the pool of languages and the variability of IRR within each host country (Table 15, Appendix A).

This identification strategy, also known as the epidemiological approach to the identification of the effects of culture and other unobserved factors (Fogli and Fernandez, 2009; Giuliano, 2007), allows us to overcome the potential bias due to omitted variables. In order to control for a range of confounding factors that might have influenced the formation of risk preferences, independently from the effect of language, we consider a set of ancestral characteristics from the parental country of origin. In addition, we include a set of individual specific continental and country fixed effects to control for unobserved characteristics in the respondent's host country and continent of origin.[21]

As for the language treatment, we apply alternative assignments of an immigrant's native language by using either the language of the respondent's country of origin, the language of the respondent father's country of origin, or the language of the respondent mother's country of origin. In order to account for the effect of length of exposure to the host country's language and/or the level of integration into host societies, we include the duration of residence in the receiving country.

Table 6 reports the estimation coefficients from our baseline regression models with the probability of being highly risk averse as the dependent variable. The IRR variable refers to the language of the respondent's country of origin.[22] Immigrants with original linguistic backgrounds characterized by an intermediate IRR intensity (Cat_IRR_Or1) are on average 25% less likely to be highly risk averse when compared to intensive and very intensive IRR users (model RA5). Moreover, the probability of risk aversion is 60% lower among moodless speakers with respect to intensive and very intensive IRR users. Finally, the coefficient on Strong_IRR_Or in model RA6 indicates that intensive and very intensive IRR users are on average 43% more likely to be highly risk averse with respect to intermediate and moodless speakers. The inclusion of trust and occupational status in model RA3 does not alter the measured effect of language. The effects of these

[21] In existing literature on risk attitudes of immigrants and migrants in general, there does not seem to be a clear consensus on whether or not they differ with respect to the native population. Some findings indicate that first-generation immigrants are more risk averse than natives (Bonin et al., 2009), whereas others find migrants to have a more positive attitude towards risk (Jaeger et al., 2010). It is likely that the risk taking behavior of migrants depends on their specific historical, political and social background and could thus result in either a higher or a lower level of risk aversion compared to the local population. In our analysis including ancestral, continental and country fixed effects allows us to keep the potential selection bias to a minimum.

[22] Estimation results of the baseline models on individuals from ESS are available in Appendix D, Table 19.

variables, again, appear to be independent of the effect of IRR.

Table 6: Logit Model: probability of being Highly Risk Averse. First-generation immigrants. Baseline specification.

Highly Risk Averse (d)	RA 1	RA 2	RA 3	RA 4	RA 5	RA 6
IRR_Or	1.264***	1.165***	1.161***	1.140***		
	(0.043)	(0.030)	(0.031)	(0.033)		
Cat_IRR_Or0					0.389***	
					(0.062)	
Cat_IRR_Or1					0.746***	
					(0.076)	
Strong_IRR_Or						1.427***
						(0.140)
Age	1.032***	1.021***	1.016**	1.013	1.012	1.013
	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Female	1.999***	1.727***	1.789***	1.807***	1.804***	1.796***
	(0.160)	(0.128)	(0.127)	(0.133)	(0.131)	(0.133)
Low Education		1.781***	1.743***	1.609***	1.630***	1.627***
		(0.175)	(0.155)	(0.142)	(0.143)	(0.145)
High Education		0.717**	0.738**	0.781**	0.777**	0.770**
		(0.104)	(0.093)	(0.094)	(0.092)	(0.092)
Income		0.872***	0.870***	0.876***	0.876***	0.877***
		(0.021)	(0.020)	(0.019)	(0.019)	(0.019)
Owner		0.702***	0.696***	0.722***	0.728***	0.726***
		(0.063)	(0.063)	(0.062)	(0.061)	(0.061)
Married			1.256***	1.255***	1.257***	1.247**
			(0.101)	(0.108)	(0.105)	(0.107)
Num. Children			0.997	0.996	0.998	0.996
			(0.038)	(0.038)	(0.037)	(0.038)
Trust			0.960***	0.965***	0.966***	0.964***
			(0.008)	(0.008)	(0.008)	(0.008)
Retired			1.292**	1.271	1.266	1.266
			(0.160)	(0.159)	(0.157)	(0.159)
Unemployed			1.504**	1.462**	1.453	1.448
			(0.285)	(0.280)	(0.278)	(0.274)
Disabled			1.295	1.127	1.119	1.122
			(0.246)	(0.227)	(0.222)	(0.224)
Homemaker			1.137	1.098	1.078	1.086
			(0.146)	(0.142)	(0.134)	(0.140)
Host Country x Wave	Yes	Yes	Yes	Yes	Yes	Yes
Continent	Yes	Yes	Yes	Yes	Yes	Yes
Cognitive, Health	No	No	No	Yes	Yes	Yes
N. Observations	11511	11511	11085	11076	11076	11076
N. Countries	20	20	20	20	20	20

Notes: *The dependent variable is "Highly Risk Averse (d)". The method of estimation is Logit with the coefficients reported as odds ratios. Robust standard errors are clustered at the country level. Reference categories: IRR = 4 or IRR = 6 (model RA5), IRR < 4 (model RA6), Male, Not Married (divorced, separated, widowed), Medium Education, Employed or Self-Employed, Do not own their house/flat.*

Even though the identification strategy specified above reduces the concerns regarding potential biases due to omitted variables, it still fails to account for factors that have been transmitted across generations, as well as for the ancestral characteristics from the parental country of origin. These factors might have influenced both the formation of preferences and triggered the gradual emergence of particular grammatical forms that fostered the transmission of these traits across generations. Linguistic structures in turn might have reinforced the effect of transmitted preferences on economic choices and behavior. Indeed, Galor and Özak (2016) show that geographical variation in the pre-industrial return to agriculture in the parental country of origin had a persistent effect on time preferences of second-generation immigrants. In particular, higher historical crop yield potential experienced by ancestral populations had a positive effect on the descendants' long-term orientation. In line with this evidence, Sarid et al. (2017) argue that these traits are at the root of existing variations in the presence of the future tense across languages. Nevertheless, their empirical findings suggest that linguistic structures still have an independent effect on time preferences and economic choices and behavior.

Initial conditions experienced by ancestral populations might also have influenced the co-evolution of other individual specific traits, such as the perception of risk, and linguistic structures concerned with the expression of uncertainty. In order to control for potentially confounding historical characteristics of ancestral populations, in Table 7 we account for a set of additional controls from Galor and Özak (2016) for the immigrant father's country of origin.[23] This set of variables includes: ancestral pre - 1500CE potential crop yield and growth cycle, and their change in the post-1500CE period, absolute latitude, mean elevation above sea level, terrain roughness, distance to coast or river, and landlocked variables. Even though these controls have been originally thought as direct and exogenous drivers of individual time preferences, some of them may also be relevant in the context of risk and uncertainty, especially the changes in potential crop yield and growth cycle in the post - 1500 period. For any given level of agricultural productivity, higher oscillations in crop productivity may have affected the individuals' level of uncertainty regarding future rewards. Furthermore, we follow Roberts et al. (2015) and correct standard errors for the relatedness between languages by including controls for language family.

[23] We consider only the respondents from Wave 5 since in other waves information on parental backgrounds was not available. The total number of countries reduces to 15 since Poland, Ireland, Hungary, Portugal and Greece are not included in Wave 5.

Table 7: Logit Model: probability of being Highly Risk Averse. First-generation immigrants. Baseline specification with additional controls from Galor and Özak (2016) referring to the respondent father's country of origin.

Highly Risk Averse (d)	RA 1	RA 2	RA 3	RA 4	RA 5	RA 6
IRR_Or	1.147**			1.155**		
	(0.069)			(0.077)		
Cat_IRR_Or0		0.340***			0.292***	
		(0.094)			(0.107)	
Cat_IRR_Or1		0.729**			0.680**	
		(0.113)			(0.119)	
Strong_IRR_Or			1.465**			1.564**
			(0.251)			(0.311)
Absolute Latitude	0.805	0.888	0.795	0.774	0.880	0.756
	(0.177)	(0.162)	(0.188)	(0.221)	(0.235)	(0.225)
Mean Elevation	1.266	1.187	1.220	1.307	1.275	1.322
	(0.244)	(0.197)	(0.242)	(0.286)	(0.263)	(0.309)
Terrain Roughness	0.943	1.022	0.985	0.913	0.958	0.932
	(0.101)	(0.087)	(0.099)	(0.114)	(0.095)	(0.108)
Distance to Coast or River	1.020	1.003	1.019	1.007	0.964	0.998
	(0.054)	(0.049)	(0.057)	(0.060)	(0.063)	(0.060)
Land-Locked	0.882**	0.887*	0.871**	0.886**	0.897**	0.864**
	(0.052)	(0.056)	(0.054)	(0.046)	(0.049)	(0.053)
Pct. Land in Tropics	0.973	1.054	0.985	0.943	1.049	0.945
	(0.239)	(0.260)	(0.252)	(0.238)	(0.273)	(0.252)
Neolithic Transition Timing	0.869	0.855	0.887	0.873	0.852	0.897
	(0.084)	(0.084)	(0.086)	(0.104)	(0.105)	(0.106)
Crop Yield	1.022	0.984	1.035	1.000	0.936	1.005
(Anc., pre-1500)	(0.049)	(0.040)	(0.050)	(0.058)	(0.053)	(0.058)
Crop Yield Change	1.153	1.224	1.207	1.089	1.115	1.097
(post-1500)	(0.206)	(0.233)	(0.219)	(0.181)	(0.187)	(0.175)
Crop Growth Cycle	0.997	1.000	0.994	0.997	1.002	0.996
(Anc., pre-1500)	(0.004)	(0.002)	(0.004)	(0.004)	(0.003)	(0.004)
Crop Growth Cycle Change	0.997	0.992	0.996	1.038	1.081	1.054
(post-1500)	(0.091)	(0.094)	(0.090)	(0.079)	(0.084)	(0.086)
Full set of regressors						
from Table 6	Yes	Yes	Yes	Yes	Yes	Yes
Linguistic Family	No	No	No	Yes	Yes	Yes
Host Country x Wave	Yes	Yes	Yes	Yes	Yes	Yes
Continent Origin	Yes	Yes	Yes	Yes	Yes	Yes
N. Observations	8246	8246	8246	8246	8246	8246
N. Clusters	15	15	15	15	15	15

Notes: *The dependent variable is "Highly Risk Averse (d)". The method of estimation is Logit with the coefficients reported as odds ratios. Robust standard errors are clustered at the country level. Reference categories: IRR = 4 or IRR = 6 (models RA2 and RA5), IRR < 4 (models RA3 and RA6), Male, Not Married (divorced, separated, widowed), Medium Education, Employed or Self-Employed, Do not own their house/flat.*

The results clearly show that linguistic structures embodied in the IRR linguistic marker still have an independent effect on risk preferences. Moreover, compared to Table 6, the probability of reporting risk aversion is now even higher for intensive IRR speakers, in particular when we control for the relatedness between languages in models RA4-RA6. The coefficients on potential crop yield, growth cycles and their change in the post-1500 period are not statistically different from zero, while among additional geographic and climatic controls, only the variable land-locked is statistically significant at the 5% level. Individuals whose parents come from countries entirely enclosed by land (or whose only coastlines lie on closed seas) have, on average, 12% less chance to be highly risk averse. This evidence may be due to the fact that land-locked countries were safer by means of the probability of attacks and incursions through sea, and/or were more self-sufficient in terms of resources.

The correlation between our linguistic marker and risk aversion remains significant even after controlling for IRR of the host country, the amount of time spent in the host country and under alternative language assignments (Table 8). The interaction between IRR and the duration of residence in model RA1 indicates that the effect of IRR weakens with time, *i.e.*, the longer the time spent living in the host country, the less pronounced the effect of the original linguistic background. The coefficient of the interaction term suggests that the effect of the original language decreases by 6% every 10 years, which runs out roughly one half of the total effect of original IRR after 40 years. Finally, in models RA2 and RA3 we consider the father's and the mother's linguistic backgrounds respectively. The estimated effect of the immigrant mother's IRR is similar in magnitude to IRR_Or in model RA1, while the IRR marker associated to the immigrant father's language of origin is lower but still significant at the 5% level.

In addition to parental linguistic backgrounds, in models RA4 - RA6 we also control for their educational attainments. The parental level of education may either determine the respondent's educational choices or influence the formation of preferences during childhood. The estimated effects of IRR are very similar to those obtained in models RA1 - RA3. The interaction term capturing the speed of deterioration of the original linguistic backgrounds loses some significance (it remains significant at the 10% level), while it becomes insignificant under alternative language assignments. The effects of language and parental backgrounds, hence, flow through different and mutually independent channels.

Table 8: Probit Model: probability of being Risk Averse. First - Generation Immigrants. Length of Residence and Alternative Language Assignments.

	RA 1	RA 2	RA 3	RA 4	RA 5	RA 6
IRR_Or	1.461***			1.448**		
	(0.206)			(0.221)		
IRR_Host	1.332***	1.360***	1.365***	1.388***	1.422***	1.428***
	(0.119)	(0.129)	(0.125)	(0.081)	(0.074)	(0.072)
Yrs. Host	1.121	1.103	1.145	1.171	1.146	1.164
	(0.100)	(0.106)	(0.095)	(0.121)	(0.130)	(0.127)
IRR_Or*Yrs.Host	0.935**			0.937		
	(0.030)			(0.035)		
IRR_F		1.386**			1.382**	
		(0.203)			(0.228)	
IRR_F*Yrs.Host		0.942			0.945	
		(0.033)			(0.038)	
IRR_M			1.488***			1.458**
			(0.219)			(0.243)
IRR_M*Yrs.Host			0.931**			0.940
			(0.031)			(0.038)
Full set of regressors from Table 7	Yes	Yes	Yes	Yes	Yes	Yes
Controls from Galor and Özak (2016)	Yes	Yes	Yes	Yes	Yes	Yes
Linguistic Family	Yes	Yes	Yes	Yes	Yes	Yes
Continent Origin	Yes	Yes	Yes	Yes	Yes	Yes
N. Observation	8204	8040	7911	4777	4679	4609
N. Countries	15	15	15	15	15	15

Notes: *The dependent variable is "Highly Risk Averse (d)". The method of estimation is Logit with the coefficients reported as odds ratios. Robust standard errors are clustered at the country level. Reference categories: Medium Education Father, Medium Education Mother*

Finally, in Table 9 we also check whether immigrants identical in age, gender, income and educational attainments, but who use IRR more or less intensively, have different probabilities of being highly risk averse. In order to facilitate the interpretation of the coefficients, we consider only a dichotomous variable for the highest category of IRR (*i.e.*, IRR = 4 or IRR = 6). The results confirm our previous findings: the level of IRR strongly correlates with individual risk attitudes, independently of their socio-economic and family characteristics. The coefficient on Strong_IRR_Or in model RA8, for instance, suggests that when comparing only individuals identical in all dimensions defined by our fixed effects, those speaking intensive IRR languages are on average 49% more likely to report risk aversion with respect to other identical individuals speaking “moodless” or intermediate IRR languages.

Table 9: Conditional Logit model: probability of being Highly Risk Averse. First-Generation Immigrants.

Highly Risk Averse (d)	RA 1	RA 2	RA 3	RA 4	RA 5	RA 6	RA 7	RA 8
Strong_IRR_Or	1.523***	1.506***	1.483**	1.398**	1.606***	1.585***	1.561**	1.486**
	(0.227)	(0.226)	(0.229)	(0.206)	(0.274)	(0.274)	(0.275)	(0.250)
Retired	1.074	1.076	1.076	1.048	1.086	1.089	1.085	1.058
	(0.179)	(0.178)	(0.179)	(0.173)	(0.184)	(0.183)	(0.183)	(0.176)
Unemployed	1.490	1.412	1.410	1.345	1.497	1.418	1.414	1.346
	(0.438)	(0.414)	(0.421)	(0.409)	(0.448)	(0.422)	(0.429)	(0.418)
Disabled	1.257	1.199	1.214	1.047	1.254	1.196	1.210	1.041
	(0.460)	(0.445)	(0.452)	(0.394)	(0.454)	(0.439)	(0.447)	(0.390)
Homemaker	1.018	1.037	1.010	0.953	1.014	1.030	1.007	0.948
	(0.317)	(0.319)	(0.306)	(0.298)	(0.304)	(0.303)	(0.295)	(0.285)
Yrs. in Host	0.865	0.878	0.875	0.888	0.863	0.876	0.874	0.886
	(0.084)	(0.082)	(0.077)	(0.074)	(0.082)	(0.080)	(0.075)	(0.073)
Owner		0.772**	0.781**	0.779**		0.767**	0.777**	0.772**
		(0.086)	(0.089)	(0.094)		(0.086)	(0.088)	(0.093)
Trust			0.969**	0.976			0.968**	0.975
			(0.015)	(0.016)			(0.015)	(0.017)
Married				1.257				1.265
				(0.211)				(0.209)
Num. Children				0.980				0.984
				(0.038)				(0.037)
Fixed Effects:								
Sex x Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country x Wave	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income x Education	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
All FE Interacted	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls:								
Controls from								
Galor and Özak (2016)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linguistic Family	No	No	No	No	Yes	Yes	Yes	Yes
Continent Origin	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cognitive, Health	No	No	No	Yes	No	No	No	Yes
N. Observations	2278	2278	2239	2237	2278	2278	2239	2237
N. Observations	15	15	15	15	15	15	15	15

Notes: *The dependent variable is "Highly Risk Averse (d)". The method of estimation is Conditional Logit Model with the coefficients reported as odds ratios. Robust standard errors are clustered at the country level. Reference categories: IRR < 4, Employed or Self-Employed, Not Married (divorced, separated, widowed), Do not own their house/flat.*

The empirical results presented so far strongly support the idea that the intensity by which individuals use specific grammatical forms to describe potential or non-actual situations correlates with their perception of risk and uncertainty. The estimated effects of our linguistic marker are strong and robust to alternative model specifications, the inclusion of a rich set of individual and country fixed effect controls, as well as within different population subsamples. Moreover, since the effects of IRR remain strong even after controlling for a set of geographical and agro-climatic characteristics for the parental country of origin, the problem of possible confounding effects between IRR and some unobserved cultural and/or social characteristics, is significantly reduced.

3.3 Empirical Strategy: Risk Aversion and Asset Accumulation

According to our hypotheses, linguistic differences directly influence the individual perception of risk and uncertainty (*Hypothesis 1*) and indirectly their investment decisions (*Hypothesis 2*). In other words, language (IRR) affects investment decisions through its direct impact on risk aversion. In light of the empirical evidence presented in the previous section which strongly supports *Hypothesis 1*, the IRR linguistic marker may represent a suitable instrument for individual risk preferences in the equation for the propensity of holding risky assets.

The empirical problem of linking risk preferences to investment decision making, hence, consists in estimating the following causal relationship:

$$AS_i = \alpha + \beta RA_i + \gamma X_i + \theta Z_i + \rho CW_i + \eta_i \tag{2}$$

where RA_i denotes the individual i's risk aversion, while X_i, Z_i and CW_i are the vectors of explanatory and control variables as in Section 3.2. In the first stage we estimate the effects of IRR and other covariates on individual self-declared risk aversion:

$$RA_i = \alpha + \pi_{i1} IRR_i + \pi_{i2} X_i + \pi_{i3} Z_i + \pi_{i4} CW_i + \zeta_i \tag{3}$$

where IRR_i denotes the number of non-indicative moods in IRR contexts in the individual $i's$ language. By plugging the first stage fitted values in the second stage equation we obtain the reduced form model for asset accumulation:

$$AS_i = \alpha + \beta \widehat{RA_i} + \gamma X_i + \theta Z_i + \rho CW_i + error_i \tag{4}$$

Since more risk averse individuals are less prone to take risk and cope with outcomes that are potentially uncertain, the empirical validation of (4) should yield a negative coefficient of $\widehat{RA_i}$.

In addition to risk aversion, individual time preferences represent another fundamental driver of intertemporal decision-making. In order to disentangle the effects of time and risk preferences on asset accumulation, we extend (3) and (4) and consider the FTR linguistic marker from Chen (2013) as a proxy for the individual subjective discount rate. Regardless of the nature of investment choices (investment in risky assets versus savings), separating the effects of risk aversion and intertemporal preferences is not an easy task. Epstein and Zin (1989), for instance, develop a theoretical model flexible enough to allow for the separation between the intertemporal preferences and the attitudes toward risk. The authors propose a class of utility functions that allows each dimension to be parameterized separately and show that this utility representation is equivalent to the CRRA utility whenever the agent's coefficient of risk aversion is inversely related to the time preference parameter. However, Andreoni and Sprenger (2012a) and Andreoni and Sprenger (2012b) agree that these two aspects of individual preferences cannot be considered as perfect substitutes but also claim that they cannot be completely separated.[24]

Even though the separability of risk and time preferences remains an open question from a theoretical point of view, an empirical implementation of FTR and IRR, the first as a proxy for the individual discount rate and the second as an instrument for risk aversion, is an attempt to disentangle the effects of these fundamental aspects of individuals preferences on the propensity to invest in risky assets.

3.3.1 Results: Risk Aversion and Asset Accumulation

The empirical estimation of the causal relationship in (2) is run on the population of first-generation immigrants.[25] For the two-stage empirical model in (3) and (4) to work, the IRR linguistic marker must satisfy two basic requirements: a) it must be correlated with the endogenous variable (instrument relevance), and b) it must be uncorrelated with the error term (independence). The exclusion restriction requires that it must not have any direct impact on the probability of holding risky assets other than through its first stage impact on risk aversion.

Table 10 reports the coefficients from the first stage estimation. According to the Stock, Wright and Yogo's rule of thumb, the F-statistics in models RA1-RA3 confirm the

[24] In a similar manner, Andersen et al. (2008) stress that the assumption of risk neutrality for individuals that are instead risk averse, result in upward-biased discount rate estimates. They also point out that the parameter values for risk and time preferences must come from the same population.

[25] Results for the full-sample of individuals and for native individuals are available upon request.

strength of our instrument. However, when we control for a set of pre-industrial agro-climatic characteristics of the parental country of origin from Galor and Özak (2016), the correlation between the instrument and the endogenous variable becomes somewhat weaker. The reduction in the F-statistic in models RA4-RA6 is due to a relatively large number of additional controls and, to a lesser extent, to smaller sample size.

Table 10: First Stage Linear Estimation and Test Statistics, First-Generation Immigrants.

Highly Risk Averse (d)	RA 1	RA 2	RA 3	RA 4	RA 5	RA 6
IRR_Or	0.023***		0.021***	0.021**		0.023**
	(0.006)		(0.006)	(0.009)		(0.008)
Strong_IRR_Or		0.059***			0.055**	
		(0.018)			(0.022)	
Strong_FTR_Or			0.013			-0.016
			(0.013)			(0.020)
Controls and Expl. Var:						
Full set of regressors from RA4, Table 6	Yes	Yes	Yes	Yes	Yes	Yes
Geographic, Crop Yield from Galor and Özak (2016)	No	No	No	Yes	Yes	Yes
Origin Immigrants	Yes	Yes	Yes	Yes	Yes	Yes
Years in Host	Yes	Yes	Yes	Yes	Yes	Yes
N. Observations	7971	7971	7971	5776	5776	5776
N. Countries	20	20	20	15	15	15
F-Statistic	12.54	10.16	11.78	5.96	6.44	7.42
Endogenous RA	0.002	0.004	0.002	0.020	0.018	0.010

Notes: *The dependent variable is "Highly Risk Averse (d)". The method of estimation is ivreg2 (only the first stage estimates reported). Robust standard errors are clustered at the country level. Reference categories: IRR $<$ 4 (models RA2 and RA5), Weak FTR (FTR = 0).*

Even though the exogeneity of the instrument cannot be directly tested, there is no reason to suspect that there is any reverse effect of the propensity to invest in risky assets on the instrument. Since we control for country fixed effects (which capture institutional and other country-specific heterogeneities), trust, education, income, occupational status and health conditions, the exclusion restriction should not be violated. In other words, it seems reasonable to assume that there are no direct effects of linguistic variation on the propensity to invest in risky assets through omitted variables. Furthermore, the coefficients in models RA3 and RA6 suggest that the two linguistic markers can be considered as independent since the inclusion of Strong_FTR_Or does not alter the estimated effect of IRR.

Table 11 shows the second stage estimates from a recursive bivariate probit model. The dependent variable (asset accumulation) equals 1 whenever individuals hold some money in stocks or shares (listed or unlisted on the stock market), and 0 otherwise. Only marginal effects are reported. In all regression models we control for country and wave fixed effects, cognitive abilities, individual health conditions, years spent living in the host country and the continent of origin. We report the estimated coefficients for a non-categorized version of the instrument only, since it proves to be a stronger instrument than Strong_IRR_Or.

To obtain a direct effect of individual time preferences on asset accumulation, we run a separate regression using the FTR linguistic marker as a proxy for the individual subjective discount rate. In order to estimate the separate effects of risk aversion and time preferences on the propensity to invest in risky assets, we re-estimate a recursive bivariate probit model using the FTR parameter as a proxy for intertemporal choice preferences and the IRR marker as an instrument for risk aversion. Furthermore, for comparison purposes, in Table 12 we report the coefficients from a simple Probit estimation of Table 11 where the individuals' risk preferences are not instrumented.

The instrumented risk aversion is highly significant and larger in magnitude than the non-instrumented one. Without controlling for time preferences, for an individual with average characteristics of the population, being highly risk averse reduces the probability of holding risky assets by approximately 13 percentage points (model AS1). As for the individual time preferences, the estimated effect of FTR in model AS2 shows that individuals with a high subjective discount rate are 3 percentage points less likely to invest in risky assets. The separate effects of risk aversion and time preferences are shown in model AS3. The effect of the individual discount rate is negative and highly significant while the coefficient of risk aversion reduces by 2 percentage points. The effect of risk aversion is almost four times larger than the effect of the individual discount rate (model AS3). Moreover, the coefficients suggest that once the individual risk preferences are taken into account, the effect of FTR reduces only by 0.5 points, which suggests that risk attitudes and time preferences cannot be treated as substitutes. This result is in line with the recent empirical evidence found in the experimental research in the field (Andersen et al., 2008; Andreoni and Sprenger, 2012b). Interestingly, the inclusion of additional controls for the characteristics of the parental country of origin does not alter significantly the effect of risk aversion while it cancels out the effect of time preferences embodied in the FTR linguistic marker. The latter result is in line with some evidence from Sarid et al. (2017) which confirms the fact that historical agro-climatic factors might have influenced the individuals' future orientation and fostered the use of grammatical forms (future reference) reflecting these particular traits.

Table 11: IV Risky Assets (Stocks): Bivariate Probit, First-Generation Immigrants.

Pr(Risky Assets)	AS 1	AS 2	AS 3	AS 4	AS 5	AS 6
Risk Aversion	-0.127***		-0.100**	-0.115**		-0.115**
	(0.042)		(0.046)	(0.047)		(0.045)
Strong_FTR_Or		-0.031***	-0.026***		-0.007	-0.007
		(0.010)	(0.008)		(0.011)	(0.013)
Owner	0.038***	0.045***	0.039***	0.042***	0.049***	0.042***
	(0.007)	(0.006)	(0.007)	(0.009)	(0.008)	(0.010)
Income	0.008***	0.012***	0.009***	0.007***	0.011***	0.007***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Low Edu.	-0.018***	-0.028***	-0.018***	-0.014*	-0.025***	-0.014*
	(0.006)	(0.005)	(0.006)	(0.008)	(0.007)	(0.008)
High Edu.	0.013*	0.027***	0.017**	0.008	0.017	0.009
	(0.007)	(0.010)	(0.007)	(0.009)	(0.014)	(0.010)
Married	0.002	-0.004	0.002	0.009	-0.000	0.010
	(0.005)	(0.006)	(0.004)	(0.006)	(0.009)	(0.006)
Num. Children	-0.006**	-0.006**	-0.005**	-0.005	-0.005*	-0.005
	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Age	0.001	0.000	0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	-0.005	-0.021***	-0.008	-0.007	-0.023***	-0.007
	(0.004)	(0.006)	(0.005)	(0.006)	(0.006)	(0.005)
Retired	0.008	0.002	0.006	0.007	0.005	0.007
	(0.008)	(0.008)	(0.008)	(0.009)	(0.010)	(0.010)
Unemployed	-0.012	-0.025***	-0.013	-0.014	-0.022*	-0.013
	(0.014)	(0.009)	(0.014)	(0.015)	(0.012)	(0.015)
Disabled	-0.028***	-0.034***	-0.028***	-0.025*	-0.029**	-0.025*
	(0.009)	(0.011)	(0.010)	(0.014)	(0.014)	(0.014)
Homemaker	0.019	0.018	0.017	0.016	0.019	0.016
	(0.013)	(0.013)	(0.013)	(0.015)	(0.015)	(0.014)
Trust	0.001	0.002	0.001	0.003*	0.003**	0.002
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Controls:						
Country x Wave	Yes	Yes	Yes	Yes	Yes	Yes
Cognitive, Health	Yes	Yes	Yes	Yes	Yes	Yes
Origin Immigrants	Yes	Yes	Yes	Yes	Yes	Yes
Years in Host	Yes	Yes	Yes	Yes	Yes	Yes
Geographic, Crop Yield from Galor and Özak (2016)	No	No	No	Yes	Yes	Yes
N. Observations	7972	7912	7972	5776	5744	5776
N. Countries	20	20	20	15	15	15
Estimation Method	Bi-Probit	Probit	Bi-Probit	Bi-Probit	Probit	Bi-Probit

Notes: *The dependent variable is "Has Stocks (d)". The method of estimation is Recursive Bivariate Probit (only second stage marginal effects reported). Robust standard errors are clustered at the country level. Reference categories: Male, Not Married (divorced, separated, widowed), Medium Education, Employed or Self-Employed, Do not own their house/flat, Weak FTR (FTR = 0). The difference in the number of observations (52) between models AS2 and AS3 is due to the presence of collinearity.*

Table 12: Probit model with a non-instrumented risk aversion, First-Generation Immigrants.

Pr(Risky Assets)	RA 1	RA 2	RA 3	RA 4	RA 5	RA 6
Strong_RA	-0.092***		-0.091***	-0.092***		-0.092***
	(0.004)		(0.005)	(0.007)		(0.006)
Strong_FTR		-0.031***	-0.027***		-0.007	-0.007
		(0.010)	(0.009)		(0.011)	(0.013)
Controls and Expl. Var:						
Full set of regressors						
from RA4, Table 6	Yes	Yes	Yes	Yes	Yes	Yes
Geographic, Crop Yield						
from Galor and Özak (2016)	No	No	No	Yes	Yes	Yes
Origin Immigrants	Yes	Yes	Yes	Yes	Yes	Yes
Years in Host	Yes	Yes	Yes	Yes	Yes	Yes
N. Observations	7912	7912	7912	5744	5744	5744
N. Countries	20	20	20	15	15	15

Notes: *The dependent variable is "Has Stocks (d)". The method of estimation is Probit. Robust standard errors are clustered at the country level. Reference categories: Intermediate and Low Risk Aversion (Strong_RA = 0), Weak FTR (FTR = 0).*

4 Conclusions

In this paper we propose an innovative approach to analyze individual attitudes toward uncertainty and asset accumulation based on the *Sapir-Whorf* hypothesis of linguistic relativity. We develop a specific linguistic marker defined on the basis of the number of non-indicative moods used in *Irrealis* contexts, *i.e.,* contexts that involve grammatical categories related to the expression of uncertainty. Our empirical exercise consists in testing the hypothesis that speakers of languages in which non-indicative moods are used more frequently perceive the world as being more mutable and uncertain with respect to speakers of languages where these forms are less common, or do not exist at all. The association between our linguistic markers and risk aversion is strong and robust to alternative model specifications, as well as to the inclusion of a rich set of additional controls at the individual and country level.

Individuals speaking languages where non-indicative moods are used more often to describe uncertain situations, have a significantly higher probability of being averse to risk. Even when we compare individuals that are identical on a number of dimensions, such as gender, education, age, income, marital status, number of children, and cognitive and health conditions, a more intensive use of non-indicative moods is still associated with significantly higher levels of risk aversion. These effects are robust across individuals living in linguistically heterogeneous countries, and across first-generation immigrants. As for the sub-population of non-native individuals, we find some evidence that the duration of residence in the host country attenuates the effects of the original linguistic backgrounds.

The approach adopted in this paper is, to the best of our knowledge, also the first non-experimental attempt to measure a direct effect of risk aversion and individual time preferences on the propensity to invest in risky financial assets. Using our linguistic marker as an instrument for the individuals' self-declared risk aversion we show that being highly risk averse reduces the probability of holding risky financial assets by 13 percentage points. In addition to risk preferences, we run separate regressions using the FTR linguistic marker (Chen (2013)) as a proxy for the individual subjective discount rate. We find that both measures are relevant determinants in the decision of investing in stocks. In line with our hypotheses, the level of risk aversion and the preference for current consumption have a negative impact on risky asset holdings. Moreover, by exploiting the orthogonality of the FTR marker and the *Irrealis* marker we were able to show that the impact of risk aversion is four times larger than the impact of the individual discount rate. Since linguistic variation is a trait of individual identity it can be used as a cultural marker, not only at the individual but also at the group level. The results obtained in this paper therefore also shed light on the importance of non-economic factors in shaping individual risk and time preferences, and consequently their economic behavior.

References

Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, 76(3):583–618.

Andreoni, J. and Sprenger, C. (2012a). Estimating Time Prefernces from Convex Budgets. *American Economic Review*, 102(7):3333–3356.

Andreoni, J. and Sprenger, C. (2012b). Risk Preferences Are Not Time Preferences. *American Economic Review*, 102(7):3357–3376.

Bellante, D. and Green, C. A. (2004). Relative risk aversion among the elderly. *Review of Financial Economics*, 13(3):269–281.

Bonin, H., Constant, A., Tatsiramos, K., and Zimmermann, K. F. (2009). Native-migrant differences in risk attitudes. *Applied Economics Letters*, 16(15):1581–1586.

Carnap, R. (1947). *Meaning and Necessity: A Study in Semantics and Modal Logic*. University of Chicago Press.

Chen, M. K. (2013). The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets. *American Economic Review*, 103(2):690–731.

Cohen, A. and Einav, L. (2007). Estimating Risk Preferences from Deductible Choice. *American Economic Review*, 97(3):745–788.

Dahl, Ö. (2000). The grammar of future time reference in European languages. In Dahl, Ö., editor, *Tense and Aspect in the Languages of Europe*, pages 309–328. Walter de Gruyter, Berlin.

Davies, I. R. L. and Corbett, G. G. (1997). A cross-cultural study of colour grouping: evidence for weak linguistic relativity. *British journal of psychology*, 88:493–517.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the European Economic Association*, 9(3):522–550.

Epstein, L. G. and Zin, S. E. (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica*, 57(4):937–969.

Fogli, A. and Fernandez, R. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, 1(1):146–77.

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

Gay, V., Hicks, D. L., Santacreu-Vasut, E., and Shoham, A. (2017). Decomposing culture: An analysis of gender, language, and labor supply in the household. *Review of Economics of the Household*, pages 1–30.

Geeraerts, D. and Cuyckens, H. (2010). *The Oxford handbook of cognitive linguistics*. Oxford University Press.

Giuliano, P. (2007). Living arrangements in western europe: Does cultural origin matter? *Journal of the European Economic Association*, 5(5):927–952.

Hartog, J., Ferrer-i Carbonell, A., and Jonker, N. (2002). Linking Measured Risk Aversion to Individual Characteristics. *Kyklos*, 55(1):3–26.

Hill, J. H. and Mannheim, B. (1992). Language and World View. *Annual Review of Anthropology*, 21(1):381–406.

Hockett, C. F. (1960). The origin of speech. *Scientific American*, 203:88–96.

Hockett, C. F. and Altmann, S. A. (1968). A note on design features. *Animal communication: Techniques of study and results of research. Indiana University Press.*, 203:61–72.

Jaeger, D. A., Dohmen, T., Falk, A., Huffman, D., Sunde, U., and Bonin, H. (2010). Direct evidence on risk attitudes and migration. *The Review of Economics and Statistics*, 92(3):684–689.

Lin, F.-T. (2009). Does the risk aversion vary with different background risk of households? *International Research Jounal of Finance and Economics*, 34(34):69 – 82.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., and Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, 8(3):108–114.

Pinker, S. (1994). *The Language Instinct: The New Science of Language and Mind.* Penguin Books.

Roberson, D., Davidoff, J., and Braisby, N. (1999). Similarity and categorisation: neuropsychological evidence for a dissociation in explicit categorisation tasks. *Cognition*, 71(1):1–42.

Roberts, S. G., Winters, J., and Chen, K. (2015). Future tense and economic decisions: Controlling for cultural evolution. *PLoS ONE*, 10(7).

Rothstein, B. and Thieroff, R. (2010). *Mood in the Languages of Europe.* Studies in language companion series. John Benjamins.

Santacreu-Vasut, E., Shoham, A., and Gay, V. (2013). Do female/male distinctions in language matter? Evidence from gender political quotas. *Applied Economics Letters*, (September 2012):37–41.

Sapir, E. (1921). *Language: An Introduction to the Study of Speech.* Harvest books. Harcourt, Brace.

Sarid, A., Galor, O., and Özak, m. (2017). Geographical origins and economic consequences of language structures. (11917).

Swan, O. E. (2002). *A Grammar of Contemporary Polish.* Slavica.

Thieroff, R. (2000). On the Areal Distribution of Tense-Aspect Categories in Europe. In Dahl, Ö., editor, *Tense and Aspect in the Languages of Europe*, pages 265–308. Walter de Gruyter, Berlin.

Weber, E. U., Blais, A.-r. E., and Betz, N. E. (2002). A Domain-specific Risk-attitude Scale: Measuring Risk Perceptions and Risk Behaviors. *Journal of Behavioral Decision Making J. Behav. Dec. Making*, 15(August):263–290.

Whorf, B. L. and Carroll, J. B. (1964). *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. MIT language antropology. M.I.T. Press.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., and Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences of the United States of America*, 104(19):7780–7785.

Appendix A: Linguistic Mapping and Distribution of IRR (SHARE)

Table 13: Number of non-indicative moods (IRR) by language

Language	Family	Sub-Family	#Moods	a	b	c	d	e	f	g	IRR
Albanian	Indo-Euro	—	>2	1	1	0	0	0	0	1	3
Arabic (IL, AC)	Semitic	—	2	1	1	1	1	0	0	0	4
Basque	Isolate	—	2	1	1	0	0	0	0	1	3
Belorussian	Indo-Euro	Slavic	1	1	1	0	0	0	1	1	4
Bulgarian	Indo-Euro	Slavic	1	1	1	0	0	0	0	0	2
Catalan	Indo-Euro	Romance	1	1	1	0	0	0	1	0	3
Croatian	Indo-Euro	Slavic	1	0	0	0	0	0	1	1	2
Czech	Indo-Euro	Slavic	0	1	1	0	0	0	1	1	4
Danish	Indo-Euro	Germanic	0	0	0	0	0	0	0	0	0
Dutch	Indo-Euro	Germanic	0	0	0	0	0	0	1	1	2
Dutch (BE)	Indo-Euro	Germanic	0	0	0	0	0	0	1	1	2
English (GB)	Indo-Euro	Germanic	0	0	0	0	0	0	0	0	0
English (CA, USA)	Indo-Euro	Germanic	0	0	0	0	0	0	0	0	0
Estonian	Uralic	Finno-Ugric	2	0	1	0	0	0	1	1	3
Finnish	Uralic	Finno-Ugric	2	0	0	0	0	0	1	1	2
French	Indo-Euro	Romance	1	1	1	0	1	0	0	0	3
French (CA, AC)	Indo-Euro	Romance	1	1	1	0	1	0	0	0	3
German	Indo-Euro	Germanic	1	0	0	0	0	0	1	1	2
German (AU, CH)	Indo-Euro	Germanic	1	0	0	0	0	0	1	1	2
Greek	Indo-Euro	—	1	0	1	0	0	0	0	1	2
Hebrew	Semitic	—	0	0	0	0	0	0	0	0	0
Hungarian	Uralic	Finno-Ugric	2	1	1	0	0	0	1	1	4
Icelandic	Indo-Euro	Germanic	1	1	1	1	0	1	1	1	6
Irish	Indo-Euro	Celtic	2	1	1	0	0	0	1	1	4
Italian	Indo-Euro	Romance	2	1	1	1	1	0	1	1	6
Latvian	Indo-Euro	Baltic	1	1	1	0	0	0	1	1	4
Lithuanian	Indo-Euro	Baltic	1	1	1	0	0	0	1	1	4
Macedonian	Indo-Euro	Slavic	1	1	1	0	0	0	0	0	2
Maltese	Semitic	—	0	0	0	0	0	0	0	0	0
Norwegian	Indo-Euro	Germanic	0	0	0	0	0	0	0	0	0
Polish	Indo-Euro	Slavic	1	1	1	0	0	0	1	1	4
Portuguese	Indo-Euro	Romance	2	1	1	1	1	0	1	1	6
Portuguese (BR)	Indo-Euro	Romance	2	1	1	1	1	0	1	1	6
Romanian	Indo-Euro	Romance	1	1	1	0	0	0	1	1	4
Russian	Indo-Euro	Slavic	1	1	1	0	0	0	1	1	4
Russian (IL, EE)	Indo-Euro	Slavic	1	1	1	0	0	0	1	1	4

Language	Family	Sub-Family	# Mood	a	b	c	d	e	f	g	IRR
Serbian	Indo-Euro	Slavic	1	0	0	0	0	0	1	1	2
Slovak	Indo-Euro	Slavic	1	1	1	0	0	0	1	1	4
Slovenian	Indo-Euro	Slavic	1	0	1	0	0	0	1	1	3
Spanish	Indo-Euro	Romance	1	1	1	0	1	0	1	0	4
Spanish (LA)	Indo-Euro	Romance	1	1	1	0	1	0	1	0	4
Swedish	Indo-Euro	Germanic	0	0	0	0	0	0	0	0	0
Turkish	Ural-Altaic	Turkic	>2	1	1	1	0	0	1	0	4
Ukrainian	Indo-Euro	Slavic	1	1	1	0	0	0	1	1	4
Welsh	Indo-Euro	Celtic	1	1	0	0	0	0	1	1	3

Notes: *Contexts: a = Modal; b = Desire; c = Attitude (non factive); d = Attitude (factive); e = Declarative; f = Protasis (counterfactual conditional); g = Apodosis (counterfactual conditional). LA stays for Latin American countries, IL for Israel, EE for Estonia, AU for Austria, CH for Switzerland, BE for Belgium, BR for Brazil, CA for Canada, AC for Arab countries and for French speaking North-African countries, USA for the United States of America.*

Table 14: % Irrealis by Country: Full Sample (LH countries in bold)

Country	IRR=0	IRR=2	IRR=3	IRR=4	IRR=6	Total
Austria	0.3	95.9	0.4	3.2	0.2	100.0
Germany	0.3	91.7	0.3	7.1	0.6	100.0
Sweden	93.8	4.3	0.0	1.8	0.1	100.0
Netherlands	0.4	98.3	0.0	1.1	0.2	100.0
Spain	0.3	0.4	21.8	77.4	0.2	100.0
Italy	0.1	0.2	0.2	0.5	98.9	100.0
France	0.5	0.8	91.2	5.5	1.9	100.0
Denmark	98.5	0.8	0.1	0.6	0.1	100.0
Greece	0.0	98.0	0.1	1.8	0.0	100.0
Switzerland	0.8	70.2	20.8	2.3	5.8	100.0
Belgium	0.3	49.4	45.8	2.1	2.3	100.0
Israel	35.3	3.3	0.5	60.6	0.4	100.0
Czech Republic	0.1	0.4	0.2	99.2	0.0	100.0
Poland	0.0	0.7	0.2	99.1	0.0	100.0
Ireland	99.6	0.0	0.0	0.4	0.0	100.0
Luxembourg	1.3	47.7	31.8	2.5	16.7	100.0
Hungary	0.0	0.3	0.0	99.6	0.0	100.0
Portugal	0.1	0.0	0.1	0.2	99.6	100.0
Slovenia	0.0	9.9	89.5	0.1	0.5	100.0
Estonia	0.0	0.2	75.7	24.2	0.0	100.0
Total	11.5	29.2	26.1	24.3	8.9	100.0

Table 15: % Irrealis by Country: First-Generation Immigrants

Country	IRR=0	IRR=2	IRR=3	IRR=4	IRR=6	Total
Austria	4.18	47.42	4.60	41.00	2.79	100.0
Germany	2.73	18.82	3.30	69.68	5.46	100.0
Sweden	18.99	55.49	0.42	24.05	1.05	100.0
Netherlands	10.78	55.60	0.43	28.88	4.31	100.0
Spain	6.32	9.77	10.06	69.25	4.60	100.0
Italy	8.05	20.69	17.24	47.13	6.90	100.0
France	5.54	8.71	6.79	58.26	20.70	100.0
Denmark	45.24	26.79	3.57	21.43	2.98	100.0
Greece	0.00	8.70	6.52	82.61	2.17	100.0
Switzerland	5.01	44.71	16.16	13.99	20.13	100.0
Belgium	3.62	20.87	28.65	22.15	24.71	100.0
Israel	2.82	5.98	0.92	89.56	0.72	100.0
Czech Republic	1.21	9.90	5.25	83.23	0.40	100.0
Poland	0.00	26.19	9.52	64.29	0.00	100.0
Ireland	93.02	0.00	0.00	6.98	0.00	100.0
Luxembourg	3.98	19.47	15.49	7.96	53.10	100.0
Hungary	1.82	18.18	0.00	78.18	1.82	100.0
Portugal	1.79	0.00	3.57	7.14	87.50	100.0
Slovenia	0.00	90.39	3.71	1.01	4.89	100.0
Estonia	0.04	0.69	0.04	99.20	0.04	100.0
Total	4.42	21.57	6.49	58.41	9.11	100.0

Table 16: % Non-Native Population by Country

Country	Non-Native	Country	Non-Native
Austria	7.86	Belgium	9.35
Germany	10.21	Israel	54.74
Sweden	7.68	Czech Republic	4.52
Netherlands	3.72	Poland	2.56
Spain	3.97	Ireland	5.97
Italy	1.15	Luxembourg	31.54
France	9.40	Hungary	1.88
Denmark	2.83	Portugal	3.01
Greece	2.17	Slovenia	10.91
Switzerland	16.43	Estonia	23.36

Appendix B: Summary Statistics (SHARE)

Table 17: Summary statistics (Full Sample)

Variable	Mean	Std. Dev.	Min.	Max.	N
Risk aversion	3.707	0.575	1	4	118148
Strong_RA	0.759	0.428	0	1	118148
No Irrealis Moods	0.115	0.319	0	1	118148
2 Irrealis Moods	0.292	0.455	0	1	118148
3 Irrealis Moods	0.261	0.439	0	1	118148
4 Irrealis Moods	0.243	0.429	0	1	118148
6 Irrealis Moods	0.089	0.285	0	1	118148
No Irrealis Moods	0.115	0.319	0	1	118148
2 or 3 Irrealis Moods	0.553	0.497	0	1	118148
4 or 6 Irrealis Moods	0.332	0.471	0	1	118148
Strong_IRR	0.332	0.471	0	1	118148
Income (deciles)	4.554	2.876	0	9	118148
Owner (housing)	0.75	0.433	0	1	118148
Low Education	0.401	0.49	0	1	118148
Medium Education	0.38	0.485	0	1	118148
High Education	0.219	0.413	0	1	118148
Trust People	5.782	2.406	0	10	118148
Married	0.692	0.462	0	1	118148
Number Children	2.158	1.371	0	17	118148
Age (categorized)	60.797	10.295	40	90	118148
Female	0.556	0.497	0	1	118148
Retired	0.551	0.497	0	1	118148
Employed	0.296	0.456	0	1	118148
Unemployed	0.032	0.177	0	1	118148
Disabled	0.037	0.188	0	1	118148
Homemaker	0.084	0.278	0	1	118148
adl	0.209	0.762	0	6	118148
iadl	0.327	0.957	0	7	118148
reading	3.718	1.104	1	5	118148
writing	3.589	1.146	1	5	118148
# chronic diseases	1.733	1.567	0	14	118148

Table 18: Summary statistics (First-generation Immigrants)

Variable	Mean	Std. Dev.	Min.	Max.	N
Risk aversion	3.771	0.534	1	4	11111
Strong_RA	0.816	0.388	0	1	11111
No Irrealis Moods	0.032	0.175	0	1	11111
2 Irrealis Moods	0.214	0.41	0	1	11111
3 Irrealis Moods	0.058	0.233	0	1	11111
4 Irrealis Moods	0.603	0.489	0	1	11111
6 Irrealis Moods	0.094	0.292	0	1	11111
No Irrealis Moods	0.032	0.175	0	1	11111
2 or 3 Irrealis Moods	0.271	0.445	0	1	11111
4 or 6 Irrealis Moods	0.697	0.459	0	1	11111
Income (deciles)	4.261	2.87	0	9	11111
Owner (housing)	0.673	0.469	0	1	11111
Low Education	0.34	0.474	0	1	11111
Medium Education	0.385	0.487	0	1	11111
High Education	0.276	0.447	0	1	11111
Trust People	5.757	2.489	0	10	11111
Married	0.684	0.465	0	1	11111
Number Children	2.175	1.42	0	14	11111
Age (categorized)	61.202	10.399	40	90	11111
Female	0.571	0.495	0	1	11111
Retired	0.552	0.497	0	1	11111
Employed	0.289	0.453	0	1	11111
Unemployed	0.048	0.214	0	1	11111
Disabled	0.046	0.209	0	1	11111
Homemaker	0.066	0.247	0	1	11111
adl	0.275	0.874	0	6	11111
iadl	0.413	1.076	0	7	11111
reading	3.605	1.12	1	5	11111
writing	3.437	1.177	1	5	11111
# chronic diseases	1.975	1.726	0	12	11111
Time Host Country (TH)	4.18	1.178	1	5	11046
F: No Irrealis Moods	0.027	0.164	0	1	8549
F: 2 Irrealis Moods	0.214	0.41	0	1	8549
F: 3 Irrealis Moods	0.064	0.245	0	1	8549
F: 4 Irrealis Moods	0.591	0.492	0	1	8549
F: 6 Irrealis Moods	0.103	0.304	0	1	8549
M: No Irrealis Moods	0.028	0.164	0	1	8689
M: 2 Irrealis Moods	0.215	0.411	0	1	8689
M: 3 Irrealis Moods	0.068	0.252	0	1	8689
M: 4 Irrealis Moods	0.589	0.492	0	1	8689
M: 6 Irrealis Moods	0.1	0.3	0	1	8689

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Absolute Latitude	0.752	0.652	-1.913	1.75	8332
Mean Elevation	-0.309	0.463	-1.131	3.126	8325
Terrain Roughness	-0.282	0.952	-1.352	7.855	8325
Distance to Coast or River	0.859	2.35	-0.621	5.181	8302
Landlocked	-0.257	0.689	-0.451	2.193	8345
Pct. Land in Tropics	-0.504	0.348	-0.558	2.202	8302
Neolithic Transition Timing	0.36	0.679	-2.05	2.201	8323
Crop Yield (Anc., pre-1500)	7.093	2.099	0	10.116	8293
Crop Yield Change (post-1500)	-0.395	0.43	-0.919	2.691	8345
Crop Growth Cycle (Anc., pre-1500)	122.991	19.163	0	199.206	8293
Crop Growth Cycle Change (post-1500)	0.113	0.62	-7.139	2.083	8345
L. Family: Semitic	0.102	0.303	0	1	11111
L. Family: Ural-Altaic	0.018	0.132	0	1	11111
L. Family: Indo-European	0.851	0.356	0	1	11111
L. Family: Uralic	0.029	0.169	0	1	11111

Appendix C: List of Countries (and Languages) and Selected Variable Description (SHARE)

List of countries - SHARE: Austria, Germany, Sweden, Netherlands, Spain (LH), Italy, France, Denmark, Greece (Wave 2 only), Switzerland (LH), Belgium (LH), Israel (LH), Czech Republic, Poland (Wave 2 only), Luxembourg (Wave 5 only) (LH), Slovenia (Wave 5 only), Estonia (Wave 5 only) (LH). Ireland not considered as no information on income is available for release 2.6.0. "LH" stems for *linguistically heterogeneous countries*.

List of languages - SHARE questionnaires: Arabic, Catalan, Czech, Danish, Dutch, Estonian, French, German, Greek, Hebrew, Italian, Polish, Russian, Slovenian, Spanish, Swedish.

List of languages - SHARE (first-generation immigrants): Albanian, Arabic, Belorussian, Bulgarian, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hebrew, Hungarian, Icelandic, Italian, Latvian, Lithuanian, Macedonian, Norwegian, Polish, Portuguese, Romanian, Russian, Serbian, Slovak, Slovenian, Spanish, Swedish, Turkish, Ukrainian.

Selected Explanatory and Control Variables Definition:

AGE (categorized): category 40 (40 - 49 years); category 50 (50 - 59 years); category 60 (60 - 69 years); category 70 (70 - 79 years); category 80 (80 - 89 years); category 90 ($>$90 years).

EDUCATION (categorized): Low Education (none or elementary), Medium Education (secondary and upper-secondary), High Education (tertiary).

INCOME: Annual income from employment in the year, net of taxes and contributions. Includes additional or extra or lump sum payment, such as bonuses, 13 month, Christmas or Summer pays.

OWNER: A binary variable equal to 1 when the respondent's household is the owner of the dwelling it currently resides in, and 0 otherwise.

HEALTH CONDITIONS:

(i) total number of limitations in doing everyday activities. The list of limitations includes: Walking 100 meters; Sitting for about two hours; Getting up from a chair after sitting for long periods; Climbing several flights of stairs without resting; Climbing one flight of stairs without resting; Stooping, kneeling, or crouching; Reaching or extending your arms above shoulder level; Pulling or pushing large objects like a living room chair; Lifting or carrying weights over 10 pounds/5 kilos, like a heavy bag of groceries; Picking up a small coin from a table.

(ii) total number of difficulties due to a physical, mental, emotional or memory problem. The list of difficulties includes: Dressing, including putting on shoes and socks; Walking across a room; Bathing or showering; Eating, such as cutting up your food; Getting in or out of bed; Using the toilet, including getting up or down; Using a map to figure out how to get around in a strange place; Preparing a hot meal; Shopping for groceries; Making telephone calls; Taking medications; Doing work around the house or garden; Managing money, such as paying bills and keeping track of expenses.

(ii) number of chronic diseases: Number of chronic diseases and disturbances certified by a doctor. The respondents are offered a list with 17 different diseases and disturbances.

COGNITIVE ABILITIES: Reading and Writing skills: Poor, Fair, Good, Very Good, Excellent.

TRUST: Level of trust in other people. Range from 0 to 10, where 0 means you can't be too careful and 10 means that most people can be trusted.

TIME IN HOST COUNTRY (TH): This variable represents the total amount of time passed between the year in which the respondent moved into host country and the interview year. We consider a categorized version of this variable (in years): $TH \leq 10$; $10 < TH \leq 20$; $20 < TH \leq 30$; $30 < TH \leq 40$; $TH > 40$.

Appendix D: Questionnaire Details, List of Counties (and Languages) and Additional Regression Results (ESS)

Risk preference question from ESS:

> *Now I will briefly describe some people. Please listen to each description and tell me how much each person is or is not like you. Use this card for your answer. She/he looks for adventures and likes to take risks. She/he wants to have an exciting life.*

(1) Very much like me;

(2) Like me;

(3) Somewhat like me;

(4) A little like me;

(5) Not like me;

(6) Not at all like me.

List of countries (ESS):

Austria, Belgium, Bulgaria, Czech Republic, Cyprus, Switzerland, Germany, Denmark, Estonia, Spain, Finland, France, UK, Greece, Croatia, Hungary, Ireland, Israel, Iceland, Italy, Lithuania, Luxembourg, Netherlands, Norway, Poland, Portugal, Romania, Sweden, Slovenia, Slovakia, Turkey, Ukraine.

List of languages (ESS):

Albanian, Arabic, Bulgarian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hebrew, Croatian, Hungarian, Italian, Lithuanian, Norwegian, Polish, Portuguese, Romanian, Russian, Slovak, Slovenian, Spanish, Serbian, Swedish, Turkish, Ukrainian.

Table 19: Probit Model: probability of being Highly Risk Averse (Adventure Risk Taking). First-generation immigrants. Baseline specification. European Social Survey.

Highly Risk Averse (d)	RA 1	RA 2	RA 3	RA 4	RA 5	RA 6
IRR	1.035**	1.042***	1.042***	1.039**		
	(0.016)	(0.017)	(0.016)	(0.017)		
Cat_IRR0					0.811***	
					(0.053)	
Cat_IRR1					0.955	
					(0.060)	
Strong_IRR					.	1.105**
					.	(0.056)
Female	1.291***	1.293***	1.319***	1.310***	1.309***	1.310***
	(0.096)	(0.088)	(0.089)	(0.083)	(0.083)	(0.083)
Age	1.098**	1.101**	1.104**	1.090***	1.090***	1.090***
	(0.041)	(0.042)	(0.045)	(0.035)	(0.036)	(0.035)
Income		0.977	0.967	0.970	0.970	0.971
		(0.014)	(0.019)	(0.018)	(0.018)	(0.018)
Low Education		1.168**	1.118	1.091	1.093	1.093
		(0.087)	(0.093)	(0.094)	(0.093)	(0.094)
High Education		0.872	0.879	0.891	0.894	0.888
		(0.076)	(0.081)	(0.085)	(0.085)	(0.084)
Trust			0.983	0.989	0.989	0.988
			(0.040)	(0.038)	(0.038)	(0.038)
Married			1.298***	1.290***	1.291***	1.288***
			(0.125)	(0.121)	(0.120)	(0.120)
Num. Children			0.978	0.987	0.987	0.986
			(0.030)	(0.032)	(0.031)	(0.031)
Unemployed				0.823	0.824	0.819
				(0.128)	(0.130)	(0.128)
Retired				1.094	1.091	1.094
				(0.104)	(0.106)	(0.106)
Disabled				1.092	1.092	1.088
				(0.112)	(0.111)	(0.111)
Country - Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Origin Immigrants	Yes	Yes	Yes	Yes	Yes	Yes
Health	No	No	No	Yes	Yes	Yes
N. Observations	22151	16014	15514	15496	15496	15496
N. Countries	32	32	32	32	32	32

Notes: *The dependent variable is "Highly Risk Averse (d)". The method of estimation is Logit. Robust standard errors are clustered at the individual level. Reference categories: IRR = 0 (model RA5), IRR < 4 (model RA7), Male, Not Married (divorced, separated, widowed), Medium Education, Employed or Self-Employed.*