



Language learning: human capital investment or consumption?

Matthias Huber^{1,2} · Ann-Marie Sommerfeld³ · Silke Uebelmesser^{1,4} 

Accepted: 29 June 2022 / Published online: 21 August 2022
© The Author(s) 2022, corrected publication 2022

Abstract

This paper focuses on foreign language learning as human capital investment or consumption. We apply the human capital investment framework to foreign language learning and enlarge it by the consumption motive. Based on a novel dataset of close to 5000 language course participants in 14 countries worldwide, we estimate individual and country-level determinants of the different motives for language learning and of the expected use of language skills in the labour market. We highlight possible spillovers from the consumption motive to a professional use, which emerge mostly in a “tied-mover” context. This provides guidance for targeted language policies.

Keywords Language learning · New dataset · Human capital investment · Consumption

JEL classification C82 · E21 · E22 · F15 · F22

Responsible Editor: Jesus Crespo Cuaresma.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, UE 124/2-2–270886786). The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Federal Office for Migration and Refugees.

✉ Silke Uebelmesser
silke.uebelmesser@uni-jena.de

Matthias Huber
matthias.huber@uni-jena.de

Ann-Marie Sommerfeld
ann-marie.sommerfeld@uni-goettingen.de

¹ Friedrich Schiller University Jena, Carl-Zeiss-Str. 3, 07743 Jena, Germany

² Federal Office for Migration and Refugees, Nuremberg, Germany

³ University of Göttingen, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany

⁴ CESifo, Munich, Germany

1 Introduction

Foreign language skills have a productive value in two different contexts. First, one can think of language skills of natives and immigrants, which are foreign relative to the main language of the country of residence. Whereas studies have found no or only very small returns to foreign language skills in the US (Saiz and Zoido 2005), high returns to those skills show up in the labour market of some European countries for immigrants (Toomet 2011; Isphording 2013) as well as for natives (Ginsburgh and Prieto-Rodriguez 2011). As the latter authors point out, these returns often depend on the relative scarcity of specific language skills. In Germany, there are large returns to expert-level English for natives and even more so for immigrants (Stöhr 2015).

Second, foreign language skills can also be viewed in the context of migration. The host country's main language is often foreign relative to the main language of the immigrants. Researchers and policy-makers alike emphasize the importance of immigrants' skills of the host country's language for integration into the labour market of the host country. More specifically, language skills improve labor market outcomes of migrants by increasing earnings (see e.g. Dustmann and van Soest 2001; Chiswick and Miller 1995) and employment probabilities (Dustmann and Fabbri 2003; Budría et al. 2019) and by improving occupational choices (Aldashev et al. 2009).

All this evidence can be put in the light of the human capital theory (Schultz 1960; Becker 1964). According to this theory, individuals choose the human capital that maximizes their expected net present value of income. Acquiring human capital is a costly investment which is expected to lead to monetary returns via increased wages or increased employment probabilities by fostering the individual productivity. This framework has been enlarged to include migration and therefore expected returns which can realise in the domestic and the foreign labour market (Sjaastad 1962); it can be further extended easily to comprise foreign language skills as a specific type of human capital.

When looking empirically at individual choices, however, the human capital model is not able to fully explain the data. Individuals often choose more education or other types than would be optimal according to the human capital theory (Oosterbeek and Webbink 1995; Oosterbeek and van Ophem 2000; Canton and de Jong 2005). If we ignore irrational choices, expected labour market returns do not seem to be the only determinant of the choices. One explanation for the observed pattern is that education or, broadly speaking, learning has a consumption value and generates direct utility (Schaafsma 1976; Lazear 1977; Kodde and Ritzen 1984). This consumption value can be defined as “the private, intended, non-pecuniary return to education” (see Alstadsaeter 2011). Individuals may then choose a quantity or type of education which leads to lower monetary returns than other possible choices (Alstadsaeter 2011; Arcidiacono 2004).

In this paper, we focus on different motives of foreign language learning. Language learning leads to a particular form of skills that can be acquired in many different contexts, e.g. at school, university, but also in language courses. While

choices about language acquisition at school are often determined by the school's curriculum or parents' preferences, participating in a language course offered by a university or a private provider as an adult is more directly related to the individual's human capital or consumption motive. For a better understanding of the different motives, in a first step, we study the determinants of the human capital motive of language learning. In a second step, we look at the determinants of a professional use of German in the labour market. While we expect a positive relation between the human capital motive and a professional use, we are particularly interested in possible spillovers from a given consumption motive to a professional use on the one hand and possible barriers that might hinder a professional use despite a given investment motive on the other hand.

To the best of our knowledge, we are the first to apply the human capital framework enlarged by the consumption motive to foreign language learning. For this, we use unique survey data of almost 5000 language course participants collected in 14 countries at institutes of the Goethe-Institut (GI).¹ The sample is very likely not representative for the populations in the respective countries, as the participants are relatively young and highly skilled. For policy-makers in the home country, and equally in Germany, this group might, however, be particularly interesting as those individuals are often more mobile and more open to international experiences, both professional and private. Furthermore, the dataset is very suitable for analysing our questions of interest about the motives of language learning and possible spillovers. We are thus able to contribute to a better understanding of foreign language acquisition in a cross-country perspective and to provide guidance for policy-makers for targeted language policies.

We use binary probit estimations to study individual and country-level determinants of the human capital investment motive and of the use of German language skills in a professional environment. In order to identify heterogeneities, we also have a closer look at subgroups based on age, gender and education and consider differences across countries. While a younger age and a job that is linked to more internationally applicable skills and a higher need for communication is positively related to the human capital motive, female gender, children and a native German partner make the consumption motive more likely. At the country-level, we find that larger linguistic and geographic distances increase the likelihood of language learning as human capital investment; a higher income level, on the contrary, allows more for language learning as consumption. We also show that a given human capital investment motive does not necessarily match with a high probability of professional use. We find spillovers from a given consumption motive to a professional use

¹ The GI is a German cultural institute, which offers language courses worldwide and is an important part of the foreign cultural policy of the German government. In addition, the GI is engaged in cultural exchange and provides information about German culture and society (Auswärtiges Amt and Goethe-Institut 2004). While the main funding is provided by the Federal Foreign Office, language courses are financed by fees (Goethe-Institut 2014). In 2021, the GI was present in 98 countries with a total of 158 institutes (Ginsburgh and Prieto-Rodriguez 2021). For more details on language learning at the GI, see Uebelmesser et al. (2018).

mostly in a “tied-mover” context, i.e. for women and in the presence of children and a German native partner.

The remainder of the paper is structured as follows. In Sect. 2, we describe the survey set-up and discuss selection issues and (limits to) the dataset’s representativity. Section 3 introduces the conceptual framework and provides graphical illustrations. Section 4 explains the empirical strategy. In Sect. 5, we present estimation results for the determinants of the human capital motive of language learning and the probability of a professional use and discuss spillover effects. Section 6 concludes.

2 Survey set-up and data

We address our questions of interest based on survey data which we collected from language course participants at the Goethe-Institut (GI). The survey was conducted between June and December 2018 and included questions on socio-demographic characteristics, education, professional background, language skills, previous migration experience and future migration plans as well as questions on the reasons of learning the German language. In the following, we explain the design of the survey, the data collection process and possible limitations.

2.1 Survey design

For our analysis, we selected 19 institutes in 14 countries. The choice of the countries was motivated by the wish to capture cross-country variations in several dimensions. Table 1 gives an overview of the selected countries and the main characteristics on which we based this selection to assure a heterogeneous sample of countries: geographic distance to the German-speaking region, linguistic distance to the German language, average income level as categorized by the World Development Indicator in 2018 (World Bank 2021) and the absence (or presence) of migration barriers vis-à-vis the German-speaking region. The presence of a large institute measured by the number of course participants was of further importance for the selection of countries. In Indonesia and South Korea, we had the opportunity to conduct the survey in more than one institute.

To reduce the issue of (non-)selection of participants into the survey, we undertook several measures to achieve a high response rate. First, the survey was translated into the main language of each country. In India, the questionnaire was in English. Additionally, we provided English and German questionnaires upon request in every country. Second, we opted for a pen-and-paper survey as this allowed for a more direct involvement with the participants. Third, each participant could take part in a lottery to win a free language course at the given institute (limited to one language course per country). Fourth, we encouraged participation in further ways, which differed between European and non-European countries. In European countries, a team member of the research project was present at the institutes for at least one unit of each course offered during a given week and handed out the questionnaire to all present participants. Most of the participants filled-in the questionnaire

Table 1 Country characteristics and response rates

Countries	Ling. close	Geogr. close	Income (GNI/capita)	Absence of migr. barriers	Notes	Participants	Response rate
Netherlands	Yes	Yes	High	Yes	Linguistically close country	139	0.67*
Great Britain	Yes	Yes	High	Yes/No (Brexit)	Linguistically close country	480	0.88
Spain	No	Yes	High	Yes	(Recently) high-unempl. c.	611	0.83
Italy	No	Yes	High	Yes	(Recently) high-unempl. c.	371	0.86
Czech Rep.	No	Yes	High	Yes	New EU country (since 2004)	481	0.83
Poland	No	Yes	High	Yes	New EU country (since 2004)	236	0.69
Romania	No	Yes	Upper-middle	Yes	New EU country (since 2007)	327	0.87
Bosnia	No	Yes	Upper-middle	No	Close, non-EU country	270	0.99
Ukraine	No	Yes	Lower-middle	No	Close, non-EU country	782	0.93
Japan	No	No	High	No	Developed, non-Europ. country	293	0.59*
South Korea	No	No	High	No	Developed, non-Europ. country	470	0.65*
Mexico	No	No	Upper-middle	No	Emerging markets	491	0.60
Indonesia	No	No	Lower-middle	No	Developing country	883	0.55*
India	Yes/No (English)	No	Lower-middle	No	Developing country	830	0.72*

* Response rates based on registered course participants, not actual attendance. Income levels are categorized as by the World Development Indicator in 2018 (World Bank 2021). Languages are linguistically close if they are Germanic languages. Countries are geographically close if the country is in Europe. Absence of migration barriers are defined by freedom of movement for workers within the European Union and the European Economic Area. For variable descriptions, see Table 9

during the course break or after the course unit, others took it home and returned it later to the team member. In non-European countries, team members were not present in person to conduct the survey. Instead, the printed questionnaires were sent by mail to the institutes and were then distributed by the course instructors. To reduce the time and effort of the instructors and other GI officials and to minimize the probability of errors in the distribution process, we prepared envelopes for each course containing the questionnaires. In Mexico, the questionnaires were distributed during the process of course inscription for the upcoming course term.

All those measures combined resulted in high participation numbers and high response rates. Table 1 gives an overview of the numbers and rates by country. In European countries, the response rates ranged from 67 to 99 %. In these countries (except the Netherlands), the response rate is based on the number of questionnaires distributed to all participants who were present during the language lesson. In non-European countries (and the Netherlands), on the contrary, the response rate is based on the number of registered course participants. In those countries, the response rates ranged from 59 to 72 %. It is not so straightforward to compare the response rates for the European and the non-European countries for the following reason: As not all registered course participants are present at every lesson, the number of registered participants is, by definition, equal or larger than the number of those who were present at the lesson when the survey took place. For the same number of collected questionnaires, therefore this leads to lower response rates in non-European institutes, which can be interpreted as lower-bounds, when compared to response rates in European institutes.² A further exception is Mexico where the response rate (60 %) is based on the number of distributed questionnaires during the process of course inscription.

2.2 Descriptives and (limits to) representativity

In total, 6664 language course participants submitted valid questionnaires. Of those, we excluded 1773 observations because of missing information in the variables utilized in our analysis. Our final sample therefore contains 4891 individuals. Table 8 explains the individual-level variables and Table 10 details the dropping of observations in a step-by-step way. While missings in individual characteristics only lead to a drop of roughly 220 observations, more missings are related to the questions on the international applicability of skills and the importance of communication skills in professional life, respectively. The most important drop is due to missing information about the main reasons of language learning, which is our main variable for constructing the human capital and the consumption motives. In order to see whether dropping observations due to missings introduces a bias, we present descriptives separately for our final sample in Table 2 and the sample before

² In those eight European institutes where a team member was present, we know the actual attendance numbers as well as the registration numbers. It turns out that on average 75% of registered participants were present.

Table 2 Descriptive statistics: means of individual characteristics by country-groups

Variable	EU	Non-EU		Total
		European	Non-European	
	<i>n</i> = 2040	<i>n</i> = 754	<i>n</i> = 2097	<i>n</i> = 4891
Age: under 35 years	0.67	0.81	0.87	0.78
Gender: male	0.40	0.38	0.40	0.40
Gender: female	0.58	0.59	0.56	0.57
Gender: n/a	0.03	0.03	0.04	0.03
Children	0.19	0.17	0.07	0.13
No partner	0.44	0.58	0.78	0.61
Partner (native German)	0.09	0.03	0.03	0.05
Partner (other native)	0.46	0.39	0.19	0.34
Occ.: low applicability	0.16	0.11	0.11	0.13
Occ.: high appl./low comm. skills	0.08	0.05	0.04	0.06
Occ.: high appl./high comm. skills	0.31	0.19	0.16	0.23
Occ.: in education	0.29	0.51	0.58	0.45
Occ.: other occ./no answer	0.15	0.13	0.11	0.13
University degree	0.74	0.56	0.51	0.61
Risk attitude	6.09	6.77	6.49	6.36
Patience	5.80	6.59	6.25	6.11
English speaker	0.48	0.35	0.22	0.35

For variable descriptions, see Table 8 and for the grouping of the countries see Table 9

the droppings in Table 11. T-tests for the means do not point towards significant differences.^{3,4}

Focusing on the means of the total sample in Table 2,⁵ we see that 78 % of the participants are younger than 35 years. The majority of course participants is female (57 %) and has no partner (61 %), with some variations across country-groups. While the share of partners with German as native language is very small in non-EU countries (3 %) (except Japan with 11 %), it is rather large in the EU (on average 9 %, but in particular due to Great Britain with 27 % and the Netherlands with 17 %).

The young average age of the participants, in particular, in the non-EU countries might be responsible for the overall low share of those with children (only 13 %). This might also explain the high share of those who indicate that they are in

³ The only exception is the variable “English speaker” with a share of 0.348 in the final sample and a share of 0.327 in the sample before the droppings (significant at the 5% level).

⁴ An analysis of the missings shows that the significant determinants of not answering the question on the main reason do not follow a pattern. We are therefore confident that this does not imply a selection issue. The results are available from the authors upon request.

⁵ See Table 12 for descriptives by countries.

education (45 % overall, over 50 % in non-EU countries, but only 29 % in EU countries). Most of the other participants are active in the labor market (42 %). Of those, 69 % state a high level of international applicability of their skills; and more than three quarter of them see a high importance of communication skills in their professional life. The majority of course participants has a university degree (61 %). In combination with those still in education, part of whom will likely receive a university degree in the future, the large majority of course participants is highly skilled.

Participants in the EU are on average slightly less risk prone, but also slightly less patient. Almost half of them have high English-language proficiency while the numbers are much lower for non-EU participants. Overall, it shows that individuals from non-EU countries (within or outside Europe) are more similar to one another than to individuals from EU countries with the exception of partner status and children where differences are more pronounced.

When looking at the descriptives, it becomes obvious that the sample is not representative for the populations at large. The participants in the survey are relatively young and highly skilled. For policy-makers in the home country, and equally in Germany, this group might, however, be particularly interesting as those individuals are often more mobile and more open to international experiences both related to their professional and their private activities. Furthermore, the dataset is very suitable for the purpose of our study given its focus on language learning. In the analyses, we will nevertheless consider subgroups based on age and on education (next to gender) to understand any potential differences and to assess their relevance for the results.

For a generalisation of our insights about motives of adult-age language-learning, a further issue concerns the question whether language course participants at the GI differ from those at other language institutes. Self-selection of participants, in general, could take place based on the following three characteristics (see also Uebelmesser et al. 2018): willingness or ability to pay, location, and age. Selection on willingness to pay could occur if the prices of courses at the GI differed significantly from prices of other equally suitable learning options. Prices could be higher if one considers the GI as a premium provider of language courses because of its semi-official status and its long tradition and good reputation. Prices could also be lower because of funding by the German government. Both arguments are not fully convincing, however: When looking at current prices of courses offered by the GI and by competitors, the prices do not indicate that the GI is usually the most expensive provider in the market. At the same time, language courses are priced to be self-financing, that is they are not financed by government funding. As to location, institutes are usually located in capitals and other major cities, which might lead to an under-representation of language learners from rural areas. However, the bias is likely attenuated by the fact that institutes also offer intensive courses taught en-block. Participants who do not live in the vicinity of the respective institute, may stay there for the duration of the course only. Still, we cannot rule out that other providers of language courses are more present in rural areas.

Admittedly, language services offered by the GI are only one way for adults to acquire skills in the German language. Naturally, there are a large number of alternatives, including universities, private language schools, and internet platforms.

This might give rise to additional concerns regarding the self-selection of language learners into courses offered by the GI, particularly based on age. Again, the bias might be less severe: On the one hand the GI has complemented its course offer by online and blended-learning courses, which combine traditional and online learning, since 2010. On the other hand, as our data show, courses offered by the GI are attractive to younger language learners; an over-representation of older participants cannot be observed in our sample. In sum, we conclude that there is no strong evidence that participants of language courses offered by the GI systematically differ from participants of language courses offered by other providers.

Despite all the caveats mentioned above, this sample suits well our purposes given our research interest in the motives of foreign language learning. It provides many individual-level information as well as information about language-learning motivations for a sample of close to 5000 individuals in 14 countries. At the same time, we are aware of the limitations as to its generalisation. In particular, we acknowledge that we cannot say anything about selection into language-courses or more generally, the decision of adults to learn a foreign language versus the decision not to do this and, despite the arguments brought forward above, we cannot fully rule out a selection bias relative to other providers.

3 Motives of language learning

Our dataset informs us about the reasons of language learning. In the following, we will explain how we derive the human capital motive and the consumption motive from these reasons, discuss the expected associations and provide some graphical illustrations.

3.1 Conceptual framework and operationalisation

In the survey among course participants, all respondents have in common that they decided to learn the German language. We now examine the motives behind this decision. All participants answered the following multiple-response question: *Why are you learning German?* Afterwards, participants were asked to choose the main reason among the stated reasons. Table 13 provides the reply options.

We categorize the main reasons according to Fig. 1 and use this categorization as the basis for our analysis. In a first step, we aggregate the 14 reasons presented in the questionnaire into the five categories *education, educational and labour migration, domestic labour market, personal reasons* and *cultural interest*. In a second step, we further aggregate these categories into the two motives *human capital investment* and *consumption*.

On the one hand, language learning can be an investment in human capital, i.e. language skills can be used in a productive way such that there are (expected) monetary returns to these skills. In our context, we use a broad definition of monetary returns and consider all categories which contain reasons related to the *domestic labour market* or the foreign labour market via *labour migration*.

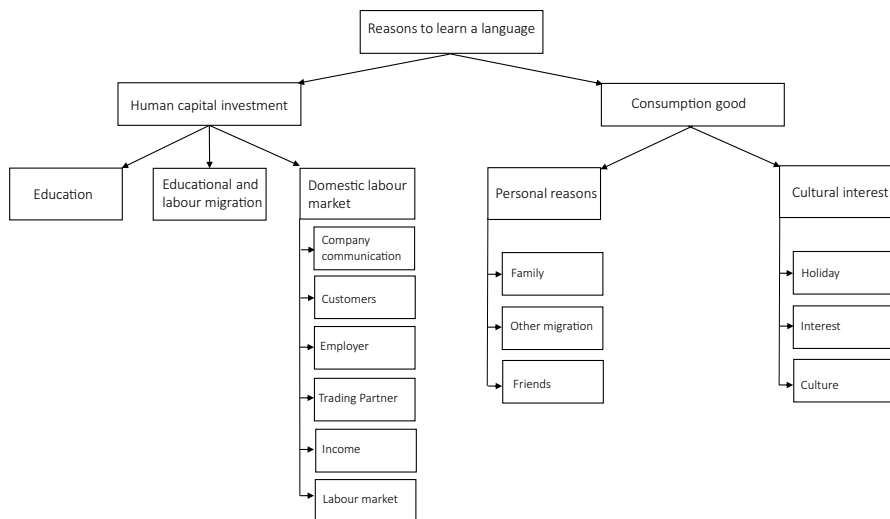


Fig. 1 Categorization of reasons to learn a language. See Table 13 in the Appendix for the exact wording of the question in the questionnaire and for country-specific details

On the other hand, language learning can be seen as a consumption good with non-monetary returns that leads to a direct increase in utility, either immediately or later. We define reasons as being related to the consumption good which belong to the categories *personal reasons* and *cultural interest*.

Additionally, learning a language can happen in the context of domestic *education* or foreign education via *educational migration*. This happens either directly by adding language skills to the human capital stock, or indirectly if language skills positively affect the accumulation of other human capital, e.g. by opening up better education possibilities in destination countries if language skills are a requirement for education. We see the human capital motive as the most relevant one when it comes to the link with education. This view is also supported by the observation that (higher) education is costly for the individuals in terms of opportunity costs and, even more important in many countries, tuition fees. Still, we acknowledge that education can also have a consumption aspect (we will come back to this point at the end of this section).

To base our categorization and the further aggregation to motives not only on the general considerations outlined above, we take the observed correlations between reasons and categories as further guidance to compensate for the lack of literature.

As can be seen in Table 14, all reasons belonging to the *personal reasons* and *cultural interest* categories are positively and significantly correlated. They are then further aggregated to the consumption motive. The reasons which are part of the *domestic labour market* category are also all significantly and positively correlated. To study the correlation of the single-reason categories *education* and *education and labour migration* category, we look at Table 15. We find support for relating both the *domestic labour market* category and the *education and labour migration* category to the human capital investment motive. As to the *education* category, the

picture is somewhat ambiguous. There is a clear positive correlation with the *education and labour migration* category, which induces us to aggregate it in the human capital motive as well. At the same time, we observe a negative, albeit smaller, correlation with the *domestic labour market* category. Therefore, when presenting our main regression results in Sect. 5, we complement them with specifications where we move the *education* category to the consumption motive and other specifications where we exclude all participants whose main reason is *education*.

3.2 Expected associations

In the following, we discuss what we expect for the association between individual-specific and country-specific explanatory variables on the one hand and the two motives on the other hand. Due to a lack of related research, our general approach and our choice of variables is guided by studies focusing on other forms of human capital investment, not language learning, based on the human capital theory (Becker 1964; Schultz 1960; Sjaastad 1962) or studies on the determinants of language proficiency of immigrants (Chiswick and Miller 2015).⁶

Looking first at individual characteristics, age is an important factor with an expected negative effect on human capital investment according to the human capital theory. To put it differently, the older the individual the less time there is to recoup the investment. Analogously, the older the participant in a language course, the less time for the monetary returns to realize. In addition, the costs of learning a language grow with age as the required effort increases. All of this makes the human capital motive less likely. This line of argument finds some support in the literature on language acquisition of immigrants in their host country (see Chiswick and Miller 2015, who stress “efficiency”, which relates to age, as an important determinant of language learning).⁷

A higher level of education, on the contrary, can be expected to decrease the cost of acquiring language skills by increasing the efficiency of learning (see again Chiswick and Miller 2015, and Footnote 7). Furthermore, foreign language skills might increase the productivity of other skills in the labour market in a complementary way and therefore positively affect the overall benefits. This makes the human capital investment more likely. Related to this, international applicability of education makes it also more likely that opportunities specifically on the foreign labour market emerge where language skills lead to benefits. This might be all the more important for occupations, where communication skills play an important role. When we focus on this complementary view, we therefore expect an overall positive correlation with

⁶ It is important to note that our analysis differs from these studies in one important way: in our case, the alternative to learning for investment purposes is learning for consumption purposes, and not “no learning” at all.

⁷ Chiswick and Miller (2015) focus on the three “Es”. They comprise exposure, which refers to the environment in which the migrants live and communicate, economic incentives, which cover a mix of internal and external factors such as planned duration of stay and expected earnings gains, and efficiency, which next to age at migration include the level of education and similar characteristics that enhance individuals’ abilities to learn.

the human capital investment motive for those in the labour market whose skills are internationally applicable and whose professional communication needs are high. There is, however, one caveat: It could be that, despite a high level of education coupled with internationally applicable skills in communication intensive jobs, German-language skills are not needed for a productive use because work relationships rely on a high proficiency in English. In this case, we would expect that German-language learning and English-language proficiency are substitutes making the human capital motive less likely for those with very good English-language skills.

For risk proneness and patience, we expect a positive correlation with the human capital motive following the investment literature starting with Becker (1964). Monetary returns to language-learning are uncertain and realise, if at all, in the future. So more risk-prone and more patient individuals should be more likely to learn a language with an investment motive.

We also predict that female participants have a lower probability of the investment motive than male participants. The situation of women on the domestic and foreign labour market is often worse in terms of labour market participation and wages. Furthermore, in the migration context, women are more likely the tied movers (Mincer 1978; Geist and McManus 2011) who join the male labour migrant with an a-priori lower own probability of labour market participation. This makes it more likely that the consumption motive dominates.

Closely related to the gender aspect, children might make it more difficult to realise benefits of language learning on the labour market. In particular, we expect that this is important for women who carry most often a larger burden of care work. If we consider the partnership status, we do not expect a significant association between a partner with a non-German native language and the investment motive. A partner with German as native language might be negatively correlated with the investment motive, as opportunities for consumption seem to be more likely, e.g. migration to the home country of the partner or communication with the partner as well as families and friends.

When looking at country-level determinants, a larger linguistic distance might mean larger costs of language-learning and therefore a smaller likelihood of observing an investment motive. Following a similar line of reasoning, a larger geographic distance – possibly linked with migration restrictions – might also make an investment motive less likely. Returns realise less easily as labor-market contacts to German-speaking countries are less frequent. At the same time, German-language skills might be less frequent in countries characterised by a larger linguistic or geographic distance allowing reaping higher returns. Predictions, therefore, are not clear. Cultural distance is, in many cases, related to linguistic and geographic distance. There are however exceptions, e.g. Australia, which justify a separate consideration with similar predictions as for linguistic and geographic distance, however. We expect individuals from higher income countries to have the means to see learning a foreign language as a consumption good and not as a way to reap monetary benefits. This should make the consumption motive more likely.

Indicating a reason for language learning, which we aggregate to the investment motive, is not exactly the same as using the German language at work with a high probability. We expect a positive relation between the two. But it is also possible

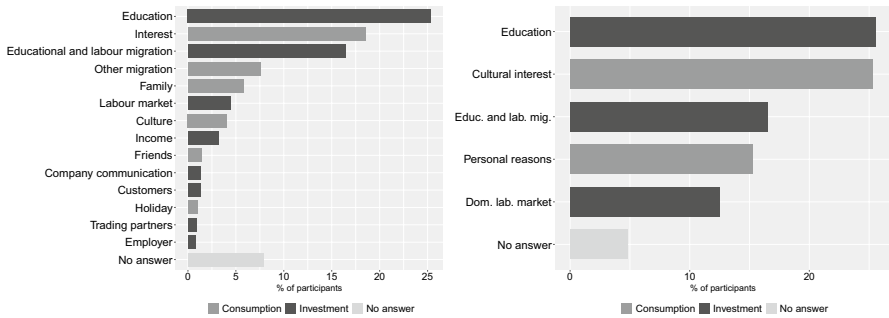


Fig. 2 Main reasons (left) and categories (right) for learning German ($n = 4891$)

that returns on the labor market will not realise in the near future – if at all. At the same time, it is possible that acquiring German language skills for consumption purposes opens up opportunities in the labor market.

For illustration, we take a look at the two migration-related reasons *educational and labour migration* and *other migration*, where the probability of a professional use of German might diverge – at least in the short-run. Let us start with *other migration*, which belongs to the consumption motive (see Fig. 1), and is often related to individuals who migrate as tied movers. At first, they acquire foreign language skills because a good command of the host country’s language increases utility by facilitating communication and integration in the new destination. Only later, the foreign language skills might be used in a professional environment, which we interpret as spillover from the consumption motive. On the other hand, individuals might prepare for *educational migration* with the purpose of using the language skills only to acquire other skills abroad before returning to their home country. That means that they do not have in mind a professional use despite their clearly given investment motive.⁸

When studying the determinants of a professional use of German, we will pay special attention to those individuals with a consumption motive. Following our reasoning from above, this group likely includes women, and individuals with a German native partner and with children.

⁸ In order to shed more light on those with a migration-related reason to learn German, we have made use of additional information in our dataset about migration intention and main reasons for migration in case of intention. Of those who stated *other migration* as main reason for language learning, 70% gave reasons related to the partner or cultural interest as their main migration reason. Considering those with *educational or labour migration* as main reason for language learning, 74% stated education or reasons related to the labour market as their main migration reason. It is interesting in our context that of those 74%, a bit more than one third (27.5%) indicated education as their main migration reason and a bit less than two thirds (46.5%) mentioned the labour market as their main migration reason.

3.3 Graphical illustrations

Figure 2 provides an overview of the relative importance of the different main reasons and categories.⁹ Looking at categories, we see that one quarter of participants each indicated either *education* or *cultural interest* as their motivation behind their decision to learn German, followed by the categories related to *educational and labour migration* (16.5 %), *personal reasons* (15.3 %) and the *domestic labour market* (12.5 %). The grey and black colors in Fig. 2, in addition, allow assessing the relative importance of the human capital motive and the consumption motive. Roughly 40% of the participants indicated a main reason belonging to the consumption motive, while for 55% the main reason is part of the human capital investment motive.¹⁰

Given our focus on cross-country differences with a special interest in possible differences between EU and non-EU countries and on heterogeneities based on age, gender and education, we also present graphical illustrations for subgroups.

In Fig. 5 in the Appendix, we present an overview of the main reasons by countries and show that there is a large heterogeneity. This is in particular obvious when looking at *education* and *family*. The share of participants that indicated *education* ranges from 4.2 to 57.4 % and the share that indicated *family* from 0.3 to 26.3 %.

When we take a look at the main reasons with the highest share in each country, we can see that there are four single reasons that make it to the top of at least one country. In India, Indonesia, Korea, Mexico and Ukraine, the largest share of participants indicated *education* as their main reason to study German. These countries have in common, that they are not members of the EU and their income is relatively low with Korea as an exception. The six countries where most of the participants indicated *interest* in languages as their main reason have the opposite in common: the Czech Republic, Great Britain, Italy, Japan, Poland and Spain are the countries with the highest income in the sample, and all are members of the EU, except for Japan. In Romania and Bosnia and Herzegovina the most important main reason is *educational and labour migration* and in the Netherlands it is *family*.

Aggregating the reasons to categories according to Fig. 1, we can again see the heterogeneity across countries similar to what we observed for the main reasons. This means that the main reason with the largest share often translates into the category with the largest share. That is the case for the Czech Republic, Spain, Poland, Japan and Italy, where the large share of *interest* translates into the category *cultural*

⁹ Given our main interest in motives (see the estimations in Sect. 5.1), we impute motives and categories if the respondents gave reasons that belong only to one motive or category, respectively. (As an example, suppose that someone indicates “Family” and “Culture” as reasons. We do not know the main reason, nor can we tell the category of the main reason as it could be *personal reasons* or *cultural interest*. As both reasons are, however, part of the consumption motive, we can assign that motive to that participant.) This allows us to increase the sample size as for the final sample we only drop observations if we neither have information about the respondent’s main reason nor are able to impute the motive. It is however possible that there are missings if we look at main reasons or categories as is the case in Fig. 2.

¹⁰ In this representation, numbers do not add up to 100% due to the “No answer” category for the main reasons. See also the preceding footnote.

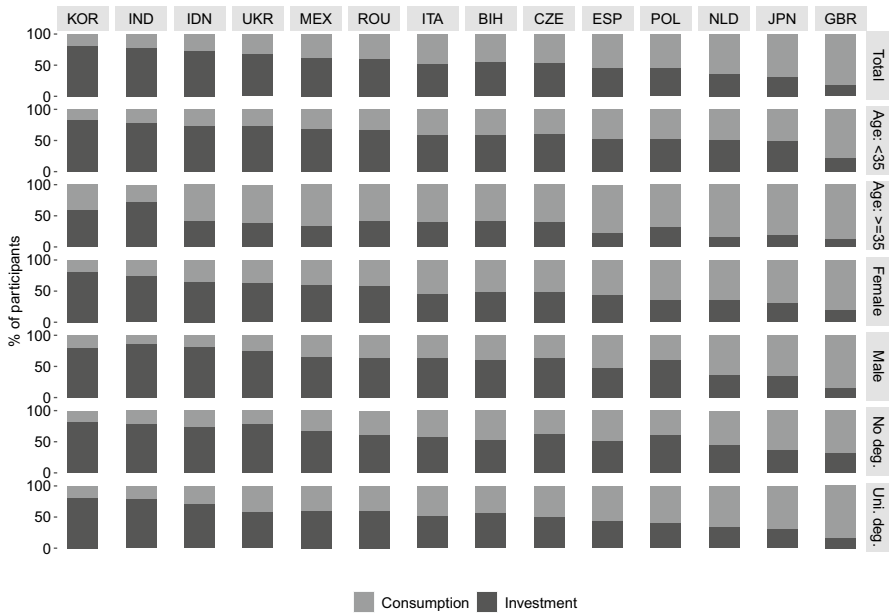


Fig. 3 Human capital and consumption motives by countries for age, gender and education subgroups. Note: The countries are arranged in descending order according to the share of participants indicating a main reason categorized as human capital investment

interest. The same holds for Korea, Indonesia, Ukraine and Mexico with *education*, for Romania with *educational and labour migration*, and for the Netherlands with *family* which translates into the category *personal interest*. There are only three countries, for which this pattern does not hold: Bosnia and Herzegovina, India and Great Britain.

Finally, Fig. 3, upper part, gives the distribution of the investment motive and the consumption motive by countries. The variation across countries is large and the share of human capital investment as main motive ranges from 17.1 to 80.1 % (and vice versa for the consumption motive). We see three groups of countries. First, the investment motive is much more important than the consumption motive in Korea, India, Indonesia, Ukraine, Mexico and Romania. Second, the shares for investment and consumption motives are much more equal with a slight tendency towards investment in Italy, Bosnia and Herzegovina and the Czech Republic. Third, in Spain, Poland, the Netherlands, Japan and Great Britain the consumption motive is more important than the investment motive.

As to differences across age groups, we see that for all countries with the exception of India, the consumption motive becomes more important for individuals older than 35 years of age compared to the full sample. In most countries, it is even more important than the investment motive. The three countries with the highest shares of the consumption motive are Great Britain, Japan and the Netherlands. It is important to note, however, that the share of older people is relatively low in some countries (see Table 12 in the Appendix). When looking at the younger age group, the pattern

is relatively close to the full sample for the three groups of countries reflecting the large share of younger individuals in the sample. The investment motive dominates in most countries except in Great Britain.

Considering gender, there are no strong differences between male and female participants in most countries. If at all, the investment motive seems to be slightly more important for men. There are also no pronounced differences between participants with and without a university degree. If at all, the human capital motive is more probable for those without a university degree (yet) than for those with a university degree. This could be partially driven by participants who are still in education and see language skills as complementary to their acquisition of other human capital.

Overall, we conclude that differences in the gender and education composition cannot explain much of the differences in the relative importance of the human capital motive and the consumption motive across countries. Age, however, seems to play an important role for the two motives behind the decision to learn a foreign language. The different composition of the participants in the different countries as far as their age is concerned translates – at least partially – into the observed cross-country differences of the importance of the two motives. In the empirical analysis in Sect. 5, we will complement the cross-country perspective by an investigation of the within-country variation.

We are also interested in the question whether a human capital motive indeed leads to a professional use of German. According to our data, the probability of using German in the labour market is on average quite high with 3.67 on a scale from 1 (very unlikely) to 5 (very likely) (see Fig. 6 in the Appendix). Around 60 % indicate that they will likely or very likely use German in the labour market. These shares can be expected if one assumes that those with investment motives also indicate a high probability of professional use, and those with consumption motives give a low or medium probability. We take this as the benchmark for our comparison, when investigating the heterogeneity of responses across countries and for our subsamples in the following.

The distribution again varies across countries as presented in Fig. 4, upper part. In the group of countries with a very high share of the human capital investment motive, we also expect a very high share of participants that indicate a high probability of using German in the labour market. In Korea, India and Indonesia, however, this share is much smaller than we expect and also smaller than in Ukraine, Mexico and Romania, where participants indicate the highest probability of a professional use compared to all other countries. A similarly mixed picture emerges for the next group of countries. In Italy, the share is smaller than expected, while it is larger in Bosnia and Herzegovina (75 %) and the Czech Republic (65 %). The last group of countries, when we follow the grouping used before, includes two countries with Spain and Poland, where we expect the share of those with a high probability to be smaller than 50 %, while it is actually around 60 %. The same holds on a somewhat lower level for the Netherlands and Great Britain, while the probability of professional use is smaller in Japan than expected.

There seem to be factors at play that hinder those with a human capital investment motive to think that they will be able to use German in a work-related context, and which vice versa make those with a consumption motive expect a professional

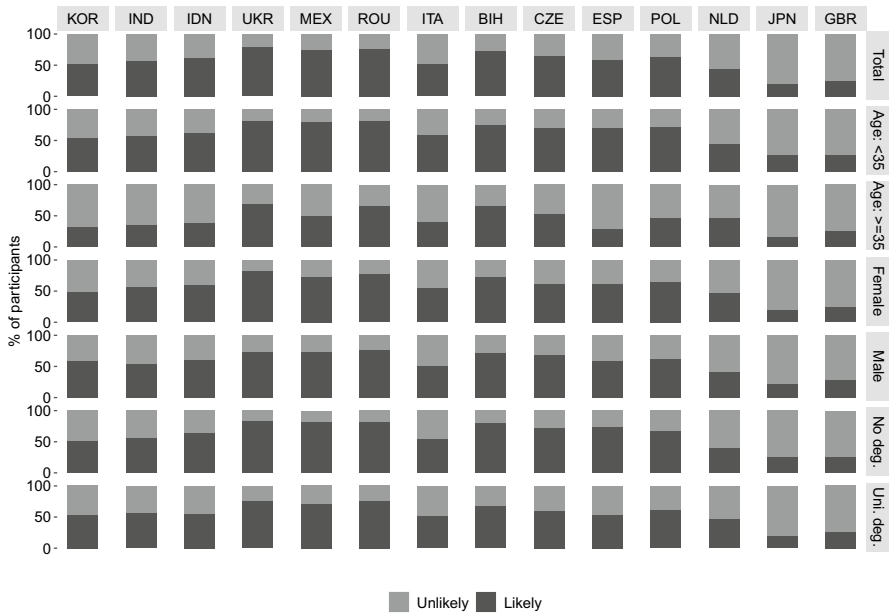


Fig. 4 Use of German on the labour market by countries for age, gender and education subgroups. Note: We aggregate the five point scale to the binary variable “Professional use of German” where values 1 to 3 correspond to “unlikely” and values 4 and 5 to “likely”. As in Fig. 3 the countries are ordered by their share of the human capital investment motive with the country with the highest share, i.e. Korea, at the very left and the country with the lowest share, i.e. Great Britain, at the very right

use of their foreign language skills. Before we examine in detail possible determinants, we again look at the distributions by age, gender and education.

Figure 4 shows that the pattern for the younger age group closely follows the pattern for the full sample. This does not hold for older participants where the likelihood of a professional use is much lower, in general. This emphasizes the role of age for the likelihood of using German on the labour market and in the educational context. The latter is supported by the pattern for those who do not have (yet) a university degree, while those with a degree indicate a somewhat lower probability of a professional use. On the contrary, there seems to be no – or in some countries (Italy, Korea, Netherlands) only a small – relationship between gender and the likelihood of using German in the labour market.

The graphical analyses above provides some evidence that there is no perfect correlation between the human capital investment motive and a high likelihood of using German in the labour market. Nevertheless, as Fig. 7 shows, there is a positive correlation between those two: In all countries, the share of those with a human capital investment motive is larger among the participants with a high likelihood for professional use.

In Sect. 5, we will first investigate the determinants of the human capital investment motive and, second, try to better understand the imperfect relationship between the human capital motive and the professional use of German. To put it differently, we want to see what makes participants with investment

motives to abstain from indicating a high likelihood of using German in the labour market and what creates spillovers from the consumption motive to the labour market.

4 Estimation strategy

We explore individual-specific and country-specific determinants of the human capital investment motive when learning German on the one hand and of a professional use of that language on the other hand. In particular, we are interested if the descriptive evidence found above for age, gender and education continues to hold after controlling for other factors.

We estimate the probability of both of our outcome variables via maximum likelihood method in a binary probit model:

$$Pr(G_i | X_i, C_i) = \alpha + \beta' X_i + \gamma' C_i + \varepsilon_i \quad (1)$$

where G_i takes a value of 1 if respondent i states to have a human capital investment motive and 0 otherwise when considering the determinants of the motive or, alternatively, takes a value of 1 if respondent i states a high probability for professional use of German and 0 otherwise when studying the use of the German language in the labour market. X_i represents a set of individual-specific explanatory variables of respondent i as presented in Table 8 in the Appendix following our theoretical considerations in Sect. 3.2. C_i captures either country-level factors as listed in Table 9 or country-fixed effects to control for country-specific heterogeneities. ε_i is an idiosyncratic error term. Standard errors are heteroscedasticity robust White standard errors.

Country-level factors are control variables that are based on the country characteristics as described in Table 1 (see also Sect. 3.2). They include linguistic differences by distinguishing between Germanic or non-Germanic official languages. They also consider economic differences by distinguishing lower-middle and upper-middle income countries as well as high income countries. Further, we categorize the countries into three groups based on their geographic distance to Germany, but also on the absence or presence of migration barriers: one group consists of EU countries, which are close to Germany and for which migration restrictions are non-existent, the second group comprises non-EU countries. This group is further subdivided into European countries, where the geographic distance to Germany is still rather small, but migration to Germany, Austria and also Switzerland is much more restricted, and non-European countries, where migration restrictions are equally relevant and, in addition, the geographic distance to German-speaking countries is much larger. As geographical distance does not always correspond to cultural distance, we further add two variables capturing this based on Hofstede and Minkov (2013): one variable about the distance in “long-term orientation” (LTO) to Germany and one variable about the distance in “indulgence versus restraint” (IVR).

5 Estimation results

5.1 Determinants of the human capital motive

We present our main results in Table 3. Column 1 includes individual-specific characteristics only. When adding country-specific controls in Column 2 via country-fixed effects and in Column 3 via country-specific characteristics, the goodness-of-fit measured with the McFadden Pseudo R² and the percentage of correctly predicted observations increase.¹¹

5.1.1 Country-level determinants

Before focusing on the individual characteristics, we take a closer look at the country-specific characteristics in Column 3. The probability of the investment motive decreases when the language spoken in the country is a Germanic language (i.e. English in Great Britain and India, and Dutch in the Netherlands) in comparison to a non-Germanic language. The benefits of learning German do not seem to be very large for those with another Germanic language as mother tongue. Given the linguistic closeness of these languages, speakers of Dutch (and German) have relatively low costs of learning English, which is the most spoken foreign language of the world (“lingua franca”). At the same time, Dutch and English allow its speakers a relatively easy access to German. Both might reduce the need for formal learning of German at adult age.¹²

For language learners from non-EU countries outside Europe, the probability of the human capital investment motive is significantly larger than for those from European countries, both in and outside the EU. This can be related to several reasons. First, language skills are often a prerequisite for legal migration to German-speaking countries from non-EU countries particularly outside Europe. This makes it more likely for language course participants from these countries to acquire language skills for investment purposes, as we saw in Fig. 3. Second, geographic proximity, which is given for all European countries, can be related to a larger migrant stock in German-speaking countries due to previous migration. This makes the consumption motive of language learning more likely for participants from these countries. Both reasons might explain the observed differences between European, EU- and non-EU countries on the one hand and non-European countries on the other hand. Considering a more direct measure of cultural proximity shows that a larger distance with respect to long-term orientation (LTO) is related to a smaller probability of the human capital motive and more importance attributed to the consumption motive.

¹¹ In Table 16, we include fixed effects interacted with age and gender. As the model fit is unchanged, we opt for the country-fixed effects without interactions as our main specification. This allows studying the importance of gender and age more explicitly.

¹² In Sect. 5.1.2, we discuss how individual proficiency of the English language relates to the motive of German language learning in order to see whether both languages are substitutes.

Table 3 Human capital investment: basic specifications

	Dependent variable: Human capital investment		
	(1)	(2)	(3)
Age: under 35 years	0.225*** (0.023)	0.157*** (0.023)	0.189*** (0.023)
Gender: female	-0.068*** (0.013)	-0.082*** (0.013)	-0.072*** (0.013)
Children	-0.033 (0.025)	-0.071*** (0.025)	-0.053** (0.025)
Partner (native German)	-0.441*** (0.027)	-0.360*** (0.034)	-0.402*** (0.030)
Partner (other native)	-0.043*** (0.016)	-0.022 (0.016)	-0.032** (0.016)
Occ.: high appl./low comm. skills	0.101*** (0.029)	0.106*** (0.027)	0.110*** (0.028)
Occ.: high appl./high comm. skills	0.114*** (0.020)	0.119*** (0.020)	0.125*** (0.020)
Occ.: in education	0.212*** (0.025)	0.186*** (0.026)	0.213*** (0.025)
University degree	0.023 (0.017)	0.045*** (0.017)	0.051*** (0.017)
Risk attitude	0.011*** (0.003)	0.006** (0.003)	0.007** (0.003)
Patience	0.010*** (0.003)	0.005* (0.003)	0.007** (0.003)
English speaker	-0.042*** (0.014)	-0.013 (0.014)	-0.025* (0.015)
Germanic language			-0.094*** (0.020)
Non-EU (European)			-0.070* (0.038)
Non-EU (Non-European)			0.058*** (0.020)
Upper-middle income			-0.039 (0.029)
High income			-0.115*** (0.024)
Cultural distance: LTO			-0.158*** (0.046)
Cultural distance: IVR			0.057 (0.037)
Country FE	No	Yes	No
McFadden Pseudo R2	0.13	0.17	0.15

Table 3 (continued)

	Dependent variable: Human capital investment		
	(1)	(2)	(3)
Percent. correctly predicted	69.5	71.3	71.0
Observations	4,891	4,891	4,891

Average marginal effects. Reference category for the occupation categories is “occupation with low internationally applicable skills and low or high communication needs” (for other reference categories, see Table 8). Heteroscedasticity robust White standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Finally, the country-wide income level plays an important role: participants from higher income countries are on average less likely to learn German for investment purposes. The higher the average income level the more likely participants have the means to see learning a foreign language as a consumption good and not as a way to reap monetary benefits. Obviously, there is a large overlap between EU countries and high-income countries (see also Table 1) which is reflected in the results here.

5.1.2 Individual-level determinants

On the individual level, there are only few differences between Columns 2 and 3. As there is a higher goodness-of-fit in Column 2 with country-fixed effects, we use that specification for the discussion about the results in the following and for the estimations by age, gender and education subgroups in Table 4.

In line with the human capital theory, the probability of the investment motive is larger for the younger age group. Being under 35 years of age leads to a 15.7 % point increase; this relationship is also robust within gender subsamples, while it only holds for those with a university degree when considering education subsamples.

We find less language learning for investment purposes for women. Having a native German partner also reduces the probability of the investment motive in comparison to singles. The absolute size of the average marginal effect is 36.0 % point in the full sample and thus almost three times as large as the age effect of the younger age group. A native German partner increases the opportunities where the consumption motive of German language skills seems to be more likely, e.g. communication with the partner as well as family members and friends. This relationship is robust within all subsamples. Also the presence of children makes the investment motive less likely. This is however mostly limited to women and those without a university degree.

The probability of the investment motive is larger by 18.6 % point for course participants in education compared to those who are in the labour market (with low internationally applicable skills). Not surprisingly, this relationship does not hold for participants in the older age group. We also observe a large probability of the investment motive for those in the labour market who have highly internationally applicable skills. When considering the subsamples, this pattern holds for those who also have high communication needs and is more mixed when communication is

Table 4 Human capital investment: subsample by age, gender and education

	Dependent variable: Human capital investment					
	(1)	(2)	(3)	(4)	(5)	(6)
	Female	Male	Age: < 35	Age: ≥ 35	No uni. degree	University degree
Age: under 35 years	0.147*** (0.031)	0.172*** (0.037)			-0.057 (0.063)	0.171*** (0.024)
Gender: female			-0.090*** (0.015)	-0.051* (0.027)	-0.085*** (0.021)	-0.082*** (0.016)
Children	-0.083** (0.034)	-0.052 (0.039)	-0.025 (0.044)	-0.055* (0.029)	-0.319*** (0.088)	-0.036 (0.026)
Partner (native German)	-0.354*** (0.038)	-0.358*** (0.069)	-0.405*** (0.045)	-0.239*** (0.033)	-0.298*** (0.107)	-0.351*** (0.031)
Partner (other native)	-0.028 (0.022)	-0.022 (0.025)	-0.015 (0.019)	-0.024 (0.032)	0.018 (0.031)	-0.037** (0.019)
Occ.: high appl./low comm. skills	0.073 (0.045)	0.133*** (0.035)	0.128*** (0.031)	0.070 (0.051)	0.167 (0.082)	0.101*** (0.030)
Occ.: high appl./high comm. skills	0.108*** (0.028)	0.131*** (0.030)	0.120*** (0.023)	0.123*** (0.036)	0.122* (0.063)	0.120*** (0.022)
Occ.: in education	0.190*** (0.034)	0.173*** (0.040)	0.232*** (0.029)	0.320 (0.272)	0.308*** (0.078)	0.194*** (0.028)
University degree	0.043* (0.023)	0.048* (0.027)	0.071*** (0.019)	-0.048 (0.048)		
Risk attitude	0.006 (0.004)	0.006 (0.005)	0.002 (0.004)	0.020*** (0.006)	0.007 (0.005)	0.006 (0.004)
Patience	0.002 (0.004)	0.009** (0.004)	0.004 (0.003)	0.007 (0.006)	-0.0004 (0.005)	0.007** (0.003)
English speaker	-0.019 (0.020)	-0.0004 (0.021)	-0.005 (0.016)	-0.033 (0.029)	0.013 (0.023)	-0.024 (0.018)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
McFadden Pseudo R2	0.15	0.20	0.11	0.15	0.08	0.20
Percent. correctly predicted	68.8	74.9	71.2	72.9	71.7	71.8
Observations	2,810	1,934	3,807	1,084	1,885	3,006

Average marginal effects. Reference category for the occupation categories is “occupation with low internationally applicable skills and low or high communication needs” (for other reference categories, see Table 8). Heteroscedasticity robust White standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

less important. Similarly, having a university degree increases the probability of the investment motive. While English-language proficiency has the expected negative relation with the investment motive as long as we do not control for country-fixed

effects, this relation becomes insignificant in the presence of these fixed effects. Apparently, variation within countries is not so large.

Human capital investment is a risky endeavour as the benefits are not certain, which shows up in the positive and significant coefficient. Looking at the subsamples, we find a positive relation for risk attitude only for the older age group. Patience is an important characteristic as well with regard to human capital investment, where benefits realize much later – if they realize at all. This shows up in the positive correlation between patience and the probability of having an investment motive, which is however more pronounced when country-fixed effects are not included. This seems to be mainly due to male participants and those with a university degree.

Figure 8 graphically displays the results by country-groups (see Table 17 for the estimates). As can be seen, differences are not large.

In order to see how sensitive our results are to our allocation of the *education* category to the investment motive (see the discussion in Sect. 3.1), we run estimations where we exclude all those with *education* as their main reason and where we categorize *education* as part of the consumption motive (see Table 19). The results for the individual characteristics are qualitatively the same as in Table 3. Not too surprisingly given that *education* is particularly important for some country-groups, there are a few changes when it comes to the country characteristics. Given that we mostly focus on specifications with country-fixed effects, this does, however, not affect our regression analyses. Overall, we see that the specifications with a different treatment of the *education* category are inferior to our main specification in terms of model fit (McFadden Pseudo R² and the percentage of correctly predicted observations).

5.2 Determinants of the professional use of German

Apart from the reasons behind their decision to learn German, participants also indicated the probability of using their foreign language skills in the labour market. If opportunities arise to use them in a professional environment, they have a productive value independent from the main reason behind the learning decision. Based on our expected associations and our graphical illustrations (see Sects. 3.2 and 3.3), we want to inquire what makes the professional use of German language skills more likely. For this, we estimate its determinants and try to identify possible spillovers from the consumption motive.

5.2.1 Country- and individual-level determinants

We present our main results in Table 5. Column 1 includes the same individual-specific characteristics and country-fixed effects as in Table 3, Column 2 additionally adds the investment motive dummy and Column 3 includes dummies for the

Table 5 Professional use: basic specifications

	Dependent variable: Professional use of German			
	(1)	(2)	(3)	(4)
Investment		0.211*** (0.015)		0.222*** (0.015)
Domestic labour market			0.052** (0.024)	
Educational and labour migration			0.033 (0.022)	
Personal reasons			-0.147*** (0.024)	
Cultural interest			-0.198*** (0.021)	
Age: under 35 years	0.113*** (0.023)	0.078*** (0.022)	0.069*** (0.022)	0.094*** (0.021)
Gender: female	0.015 (0.014)	0.032** (0.013)	0.033** (0.014)	0.039*** (0.013)
Children	-0.049* (0.026)	-0.034 (0.025)	-0.039 (0.026)	-0.018 (0.024)
Partner (native German)	0.043 (0.031)	0.100*** (0.028)	0.093*** (0.031)	0.074** (0.028)
Partner (other native)	0.032* (0.017)	0.038** (0.016)	0.034** (0.017)	0.034** (0.016)
Occ.: high appl./low comm. skills	0.039 (0.033)	0.014 (0.033)	0.007 (0.034)	0.011 (0.033)
Occ.: high appl./high comm. skills	0.113*** (0.021)	0.087*** (0.021)	0.078*** (0.022)	0.086*** (0.021)
Occ.: in education	0.169*** (0.025)	0.129*** (0.025)	0.136*** (0.026)	0.135*** (0.025)
University degree	0.007 (0.018)	-0.003 (0.017)	0.002 (0.018)	-0.002 (0.017)
Risk attitude	0.022*** (0.003)	0.020** (0.003)	0.020*** (0.003)	0.019*** (0.003)
Patience	0.004 (0.003)	0.003 (0.003)	0.003 (0.003)	0.004 (0.003)
English speaker	0.011 (0.015)	0.017 (0.015)	0.022 (0.015)	0.017 (0.015)
Germanic lang.				-0.165*** (0.021)
Non-EU (European)				0.018 (0.038)
Non-EU (Non-European)				-0.158*** (0.020)

Table 5 (continued)

	Dependent variable: Professional use of German			
	(1)	(2)	(3)	(4)
Upper-middle income				-0.001 (0.029)
High income				-0.136*** (0.024)
Cultural distance: LTO				0.087* (0.045)
Cultural distance: IVR				0.103*** (0.038)
Country FE	Yes	Yes	Yes	No
McFadden Pseudo R2	0.12	0.15	0.15	0.14
Percent. correctly predicted	68.1	70.1	70.2	70.0
Observations	4,891	4,891	4,654	4,891

Average marginal effects. Reference category for the occupation categories is “occupation with low internationally applicable skills and low or high communication needs” (for other reference categories, see Table 8). Heteroscedasticity robust White standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

categories with *education* as reference category instead of the investment motive dummy. Column 4 re-estimates Column 2 with country characteristics instead of country-fixed effects.¹³

When adding the investment motive in Table 5, Column 2, and the categories in Column 3, the goodness-of-fit measured with the McFadden Pseudo R2 and the percentage of correctly predicted observations improves. Furthermore, some variables turn significant, e.g. being female or having a native German partner. As we expected, an investment motive increases the probability of a professional use of German significantly by 21.1 % points. Categories that belong to the consumption motive are negatively associated compared to the reference category *education*. Within the investment categories, individuals that indicate a reason that refers to the *domestic labour market* are more likely to have a high probability of professional use of German in comparison to those who indicate reasons which are part of the *education* category and also of the *educational and labour migration* category.

The country characteristics in Column 4 are similarly correlated to a professional use of German as in Table 3, where we considered the determinants of the human capital investment motive, as far as linguistic and economic factors are concerned. On the contrary, a larger geographic distance as we have for non-European countries is now negatively associated with a larger probability of a professional use and a larger cultural distance is now positively associated. The latter holds, in particular, for the IVR measure, which was insignificant in the analyses above.

¹³ Note that the sample is slightly smaller in Column 3 due to the imputation of the categories and not of the motives as otherwise done. This leads to some missings as described in Sect. 3.1.

Table 6 Professional use: subsamples by age, gender and education

	Dependent variable: Professional use of German					
	(1)	(2)	(3)	(4)	(5)	(6)
	Female	Male	Age: < 35	Age: ≥ 35	No uni. deg.	Uni. deg.
Investment	0.186*** (0.019)	0.255*** (0.025)	0.162*** (0.017)	0.381*** (0.032)	0.154*** (0.024)	0.248*** (0.019)
Age: under 35 years	0.091*** (0.029)	0.059* (0.036)			0.187*** (0.080)	0.058** (0.023)
Gender: female			0.034** (0.015)	0.002 (0.028)	0.057 (0.022)	0.021*** (0.017)
Children	-0.025 (0.033)	-0.058 (0.040)	-0.007 (0.048)	-0.003 (0.029)	0.123* (0.058)	-0.049* (0.026)
Partner (native German)	0.087** (0.035)	0.133** (0.048)	0.107*** (0.034)	0.071 (0.050)	0.020*** (0.100)	0.111 (0.030)
Partner (other native)	0.015 (0.021)	0.074*** (0.027)	0.074*** (0.019)	-0.062* (0.032)	0.057 (0.032)	0.031* (0.019)
Occ.: high appl./low comm. skills	0.051 (0.047)	-0.026 (0.049)	0.021 (0.041)	-0.016 (0.051)	0.137 (0.119)	0.006 (0.034)
Occ.: high appl./high comm. skills	0.098*** (0.027)	0.066* (0.036)	0.091*** (0.027)	0.092*** (0.036)	0.179*** (0.054)	0.086** (0.023)
Occ.: in education	0.135*** (0.031)	0.110*** (0.042)	0.165*** (0.029)	-0.065 (0.134)	0.179*** (0.075)	0.139** (0.028)
University degree	-0.019 (0.022)	-0.004 (0.028)	0.010 (0.019)	-0.003 (0.042)		
Risk attitude	0.019*** (0.004)	0.022*** (0.005)	0.023*** (0.004)	0.009 (0.006)	0.030*** (0.005)	0.015*** (0.004)
Patience	0.002 (0.004)	0.005 (0.005)	0.003 (0.003)	0.008 (0.006)	0.0003 (0.005)	0.005 (0.003)
English speaker	0.038** (0.019)	-0.008 (0.023)	0.024 (0.017)	0.001 (0.030)	0.018 (0.024)	0.020 (0.018)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
McFadden Pseudo R2	0.16	0.14	0.11	0.22	0.11	0.17
Percent. correctly predicted	70.9	69.8	70.0	74.9	71.0	70.2
Observations	2,810	1,934	3,807	1,084	3,006	1,885

Average marginal effects. Reference category for the occupation categories is “occupation with low internationally applicable skills and low or high communication needs” (for other reference categories, see Table 8). Heteroscedasticity robust White standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

The results for individual characteristics are partially in line with previous results on the human capital investment motive (see also Table 6 for subgroups by age, gender and education), but there are three important differences. First, the relationship between being female and the probability of a professional use

of German now turns positive and significant, but only among the younger individuals and those with a university degree. Second, while having a partner with German as native language is associated negatively with the investment motive, we find a positive and significant relationship for the professional use of German, but more pronounced among younger respondents and those without a university degree. It also holds for both gender subsamples. Children, however, are not significantly related. Third, a university degree and an occupation with highly applicable skills, but no high communication needs are not significantly associated with the probability of a professional use, while both coefficients were significant and positive in the investment motive specification.

As to the different country-groups, we see many similarities, but also some differences (see Fig. 9 and Table 18). First, Germanic countries stand out as individual characteristics there mostly do not play a role for the professional use of German language skills. Second, EU countries and high-income countries are very similar and the same holds for culturally close, upper-middle income and non-Germanic countries. There are also comparable patterns for lower-middle income countries and culturally more distant countries.

In order to assess the sensitivity of our results relative to our allocation of the *education* category to the investment motive (see the discussion in Sect. 3.1), we also run estimations where we exclude all those with *education* as their main reason and where we categorize *education* as part of the consumption motive (see Table 20). Results for both the individual characteristics and the country characteristics are qualitatively very similar to those in Table 5. Overall, we see that the specification with *education* as part of the consumption motive is inferior to the specification where it is part of the investment motive in terms of model fit, while the specification where we drop those with *education* as main reason is comparable to it.

5.2.2 Spill-overs

Three individual characteristics – female gender, partner with German as native language and a younger age – are important determinants for spillovers from a consumption motive to a professional use of German as can be seen in Table 7. In addition, it shows in Column 3 that being still in education (but without a university degree) or in occupation with internationally applicable skills and high communication needs and also a risk proneness lead to a higher probability of a professional use of German, when the main motive of language learning is consumption. We further include in this specification if the respondent indicated a least one reason that we categorize as an investment reason, which is positively related to the likelihood of a professional use of German.

Overall, it seems that while language learning has a larger consumption value for younger women with a native German partner, who might be considered “tied-movers”, a professional use of German language skills is not unlikely. This holds especially if the investment motive plays a role as well (even though not the main one).

Table 7 Professional use: subsamples by consumption/investment

	Dependent variable: Professional use of German			
	(1)	(2)	(3)	(4)
	Consumption good	Hum. capital investment	Consumption good	Hum. capital investment
Age: under 35 years	0.137*** (0.029)	-0.038 (0.033)	0.092*** (0.029)	-0.038 (0.033)
Gender: female	0.046** (0.022)	0.015 (0.017)	0.055*** (0.021)	0.016 (0.017)
Children	-0.017 (0.033)	-0.045 (0.044)	-0.002 (0.032)	-0.045 (0.043)
Partner (native German)	0.121*** (0.036)	0.069 (0.082)	0.138*** (0.035)	0.070 (0.082)
Partner (other native)	0.039 (0.026)	0.037* (0.021)	0.031 (0.026)	0.037* (0.021)
Occ.: high appl./low comm. skills	0.068 (0.046)	-0.031 (0.048)	0.057 (0.045)	-0.031 (0.048)
Occ.: high appl./high comm. skills	0.073** (0.031)	0.094*** (0.030)	0.056* (0.030)	0.095*** (0.030)
Occ.: in education	0.152*** (0.039)	0.132*** (0.034)	0.126*** (0.039)	0.133*** (0.034)
University degree	-0.028 (0.029)	0.017 (0.021)	-0.040 (0.028)	0.017 (0.021)
Risk attitude	0.017*** (0.005)	0.021*** (0.004)	0.014*** (0.005)	0.021*** (0.004)
Patience	0.002 (0.004)	0.003 (0.004)	0.001 (0.004)	0.003 (0.004)
English speaker	0.010 (0.022)	0.026 (0.019)	0.007 (0.022)	0.026 (0.019)
Other investm. reason			0.226*** (0.023)	-0.005 (0.017)
Country FE	Yes	Yes	Yes	Yes
McFadden Pseudo R2	0.08	0.14	0.08	0.18
Percent. correctly predicted	71.8	69.2	72.1	71.6
Observations	2,067	2,824	2,067	2,824

Average marginal effects. Reference category for the occupation categories is “occupation with low internationally applicable skills and low or high communication needs“ (for other reference categories, see Table 8). Heteroscedasticity robust White standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

6 Conclusion

While the productive value of language skills has been shown in previous literature, our contribution is to highlight that it is not enough to focus on the human capital aspect of language learning. For the full picture, we enlarge the human capital framework by

adding the consumption motive of foreign language learning. Based on a unique dataset collected from close to 5000 language course participants in 14 countries worldwide, we analyse language learning in a cross-country perspective.

Our results show that the group of language learners is heterogeneous within and between countries. From the perspective of German-speaking countries, two points of interest emerge: First, the human capital motive is particularly relevant for course participants in the context of education and the labour market, both abroad or at home. Policy measures targeting this group, such as subsidies for foreign language learning or, in general, more language-learning opportunities, should therefore be one focus.¹⁴

Second, there are possible spillovers from the consumption motive to a professional use of German in the labour market, which might be of interest for policy-makers as well. This group mostly comprises younger women with a native German partner, who might be considered “tied-movers”. Even though the consumption motive is the main reason for them to learn a foreign language, a professional use of the language is not unlikely. As a consequence, policy measures aiming at this group not only support their social, but also their economic integration in Germany.

The immigration-related regulations of recent years in Germany can be viewed in the context of our findings: The “A1 requirement” for family reunification, which became effective in Germany in 2007, introduced the requirement that spouses from non-EU countries must have basic knowledge of German at the A1 level before being granted permission to live in Germany with their partners. By establishing a minimum level of language proficiency of migrants, this regulation lies the basis for the spillover effects from language learning for consumption reasons to an application of the acquired skills in the labour market. In contrast to this, the new Skilled Immigration Act effective since 2020 facilitates access of skilled workers from third countries to the German labour market and, by doing so, reduces uncertainty related to the returns of investing in the language of the destination country (see Uebelmesser et al. 2021, for an analysis about the consequences for language learning incentives based on macro data from the GI). Overall, the two policies address the two different motives: the consumption motive in the former case and the investment motive in the latter case. Due to the spillover effects identified above, they foster language learning in the migration context and lead to better integrated individuals.

While language learning related to migration is of importance, we need to keep in mind that there are also reasons for language learning in the absence of any migration intention both for investment or consumption purposes. Only when considering all contexts, do we get the full picture.

Appendix 1: Tables

See Tables 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 and 20.

¹⁴ For macro-level analyses of language-learning opportunities and migration to Germany based on aggregate data from the GI, see Huber and Uebelmesser (2019).

Table 8 Variable description: individual characteristics

Variable name	Type	Description
Age: under 35 years	Binary	Indicates whether participant is under 35 years, based on age group according to the ranges: <i>under 18, 18 to 24, 25 to 34, 35 to 49, 50 to 64, 65 and older</i> . Reference category is 35 years and older.
Gender: female	Categorical	Indicates whether respondent's gender is female. Reference category is male. Includes a third category of respondents indicating "No answer/prefer not to say" or if response is missing that is not reported in the result tables.
Partner: native German	Categorical	Indicates whether respondent has a partner with German as native language. Reference category is single.
Partner: other native	Categorical	Indicates whether respondent has a partner with other native languages than German. Reference category is single.
Children	Binary	Indicates whether respondent has any children. Reference category is no children.
University degree	Binary	Indicates whether respondent has a university degree based on highest educational qualification: <i>no university degree (no degree, school diploma which cannot lead to higher education, school diploma which can lead to higher education), university degree below PhD, PhD</i> . Reference category is no university degree.
Occ.: in labor market	Categorical	Indicates whether respondent's main occupation is in labor market based on the category on main occupation: <i>employee/civil servant with non-highly skilled job, employee/civil servant with highly skilled job, self-employed graduate (lawyer, doctor, ...)/freelance, other self-employed</i> .
Occ.: in education	Categorical	Indicates whether respondent's main occupation is in education based on the category on main occupation: <i>pupil, student, student apprentice/unpaid trainee or apprentice</i> .
Occ.: other/no answer	Categorical	Indicates whether respondent's main occupation is not answered or other based on the category on main occupation: <i>unemployed, housewife/househusband, retiree or other</i> . Not reported in the result tables.
International applicability of skills	Binary	Measures respondents' self-evaluated international applicability of skills (" <i>If you work abroad, you may be able to use only some of your acquired skills there. How much of your education or professional skills do you think you can use abroad?</i> ") on a 5-point scale from 1 for "none" to 5 for "all". Respondents indicating a value under 4 are categorized as "low applicability" and respondents indicating a value of 4 and above as "high applicability".
Importance of communication skills	Binary	Measures respondents' self-evaluated importance of communication skills in professional life (" <i>How important are communication skills in your professional life?</i> ") on a 6-point scale from 1 for "not at all important" to 6 for "very important". Respondents indicating a value under 5 are categorized as "low communication skills" and respondents indicating a value of 5 and above as "high communication skills".

Table 8 (continued)

Variable name	Type	Description
Risk attitude	Numerical (0–10)	Measures respondents' self-reported willingness to take risks (“Would you describe yourself as someone who tries to avoid risks (<i>risk-averse</i>) or as someone who is willing to take risks (<i>risk-prone</i>)?”) on a 11-point scale from 0 for “ <i>risk-averse</i> ” to 10 for “ <i>risk-prone</i> ”.
Patience	Numerical (0–10)	Measures respondents' self-reported patience (“Would you describe yourself as an impatient or a patient person in general?”) on a 11-point scale from 0 for “ <i>very impatient</i> ” to 10 for “ <i>very patient</i> ”.
English speaker	Binary	Indicates whether respondent is a native speaker of the English language or has command of the English language comparable to a native speaker, indicated by the highest language level (5) on a 5-point-scale of 1 for “ <i>basic knowledge</i> ” to 5 for “ <i>very high language level</i> ”. Reference category is no English speaker.

Table 9 Variable description: country-characteristics

Variable name	Type	Description
Germanic language	Binary	Indicates whether the language spoken in the country in which the survey took place is a Germanic language (i.e. English in Great Britain and India, and Dutch in the Netherlands). Reference category is non-Germanic.
European / EU	Categorical	Indicates whether the country in which the survey took place is a European country that is a member of the EU (Netherlands, Great Britain, Spain, Italy, Czech Republic, Poland and Romania), a European country that is not a member of the EU (Ukraine and Bosnia and Herzegovina), or a non-European country (Japan, Korea, Mexico, India and Indonesia). Reference category is EU-country.
Income	Categorical	Indicates income levels of the country in which the survey took place: High income (Czech Republic, Great Britain, Italy, Japan, Netherlands, Poland, South Korea and Spain), upper-middle income (Bosnia and Herzegovina, Mexico and Romania), and lower-middle income (India, Indonesia and Ukraine). <i>World Development Indicator in 2018</i> (World Bank 2021).
Cultural Distance to Germany: LTO	Rate (0–1)	Distance in long-term orientation index to Germany by country, normalized between 0 (lowest distance) to 1 (largest distance), as of 2013 or latest year available. (Hofstede and Minkov 2013)
Cultural Distance to Germany: IVR	Rate (0–1)	Distance in indulgence-restraint index to Germany by country, normalized between 0 (lowest distance) to 1 (largest distance), as of 2013 or latest year available. (Hofstede and Minkov 2013)

Table 10 Sample shrinkage

Step	Action	EU	non-EU		Total
			European	Non-European	
1	Completed questionnaires	2645	1052	2967	6664
2	Remove missings: Age	2631	1047	2958	6636
3	Remove missings: Children and partner	2616	1030	2928	6574
4	Remove missings: University degree	2593	1016	2902	6511
5	Remove missings: Risk attitude	2582	1014	2876	6472
6	Remove missings: Patience	2579	1010	2869	6458
7	Remove missings: English speaker	2579	1009	2856	6444
8	Remove missings: Applic. and comm. skills	2375	971	2821	6167
9	Remove missings: Reasons for lang. learn.	2306	892	2765	5963
10	Remove missings: Main reason for lang. learn.	2040	754	2097	4891

No missings for gender and occupations (see Table 8 for further explanations)

Table 11 Descriptive statistics: means of individual characteristics by country-groups (sample before droppings)

Variable	EU	Non-EU		Total
		European	Non-European	
	<i>n</i> = 2645	<i>n</i> = 1052	<i>n</i> = 2967	<i>n</i> = 6664
Age: under 35 years (28)	0.63	0.79	0.88	0.77
Gender (0)				
Gender: male	0.39	0.39	0.39	0.39
Gender: female	0.58	0.58	0.57	0.58
Gender: n/a	0.03	0.03	0.04	0.04
Children (2)	0.20	0.19	0.06	0.14
Partner (70)				
No partner	0.44	0.57	0.79	0.62
Partner (native German)	0.09	0.03	0.03	0.05
Partner (other native)	0.48	0.40	0.18	0.33
Occupation (290)*				
Occ.: low appl.	0.17	0.11	0.10	0.13
Occ.: high appl./low comm. skills	0.08	0.05	0.04	0.05
Occ.: high appl./high comm. skills	0.31	0.20	0.16	0.22
Occ.: in education	0.28	0.48	0.59	0.45
Occ.: other occ./no answer	0.17	0.16	0.12	0.14
University degree (76)	0.75	0.56	0.52	0.62
Risk attitude (61)	6.10	6.81	6.52	6.40
Patience (52)	5.83	6.60	6.30	6.16
English speaker (14)	0.47	0.34	0.20	0.33

Number of missing observations per variable in parentheses. Missing observations excluded from means.

*Missings due to variables “International applicability of skills” and “Importance of communication skills”

Table 12 Descriptive statistics: means of individual characteristics by countries

Variable	BIH n = 143	CZE n = 362	ESP n = 521	GBR n = 380	IDN n = 705	IND n = 491	ITA n = 272	JPN n = 225	KOR n = 305	MEX n = 367	NLD n = 86	POL n = 154	ROU n = 265	UKR n = 611	Total n = 4891
Age: under 35 years	0.75	0.70	0.75	0.50	0.96	0.97	0.65	0.40	0.93	0.83	0.60	0.68	0.72	0.83	0.78
Gender: male	0.48	0.35	0.41	0.48	0.45	0.37	0.39	0.32	0.33	0.45	0.44	0.34	0.33	0.36	0.40
Gender: female	0.51	0.63	0.57	0.48	0.48	0.61	0.58	0.67	0.66	0.51	0.53	0.61	0.65	0.61	0.57
Gender: n/a	0.01	0.02	0.02	0.04	0.07	0.02	0.03	0.02	0.01	0.03	0.02	0.05	0.02	0.03	0.03
Children	0.20	0.24	0.12	0.22	0.04	0.03	0.16	0.25	0.04	0.09	0.21	0.18	0.22	0.17	0.13
No partner	0.52	0.36	0.62	0.32	0.81	0.94	0.41	0.37	0.91	0.63	0.38	0.42	0.46	0.60	0.61
Partner (native German)	0.02	0.05	0.05	0.27	0.03	0.00	0.05	0.11	0.00	0.04	0.17	0.03	0.02	0.03	0.05
Partner (other native)	0.45	0.58	0.33	0.41	0.16	0.05	0.54	0.52	0.09	0.34	0.44	0.55	0.52	0.37	0.34
Occ.: low appl.	0.06	0.21	0.08	0.20	0.08	0.04	0.15	0.44	0.08	0.10	0.19	0.25	0.17	0.12	0.13
Occ.: high appl./low comm. skills	0.13	0.07	0.08	0.09	0.04	0.00	0.08	0.07	0.03	0.08	0.09	0.08	0.07	0.04	0.06
Occ.: high appl./high comm. skills	0.22	0.31	0.25	0.45	0.13	0.13	0.22	0.12	0.07	0.40	0.30	0.29	0.35	0.18	0.23
Occ.: in education	0.41	0.31	0.42	0.06	0.68	0.77	0.34	0.14	0.67	0.32	0.23	0.29	0.31	0.54	0.45

Table 12 (continued)

Variable	BIH n = 143	CZE n = 362	ESP n = 521	GBR n = 380	IDN n = 705	IND n = 491	ITA n = 272	JPN n = 225	KOR n = 305	MEX n = 367	NLD n = 86	POL n = 154	ROU n = 265	UKR n = 611	Total n = 4891
Occ.: other occ./no answer	0.19	0.10	0.17	0.19	0.08	0.06	0.21	0.24	0.15	0.11	0.19	0.09	0.10	0.12	0.13
University degree	0.59	0.66	0.70	0.95	0.28	0.57	0.71	0.86	0.44	0.71	0.71	0.73	0.70	0.56	0.61
Risk atti- tude	7.01	5.64	6.36	5.52	6.96	6.59	6.89	5.28	5.62	6.91	6.34	5.82	6.21	6.72	6.36
Patience	6.75	5.55	5.87	5.34	6.07	6.57	6.09	6.24	6.18	6.22	5.74	5.75	6.38	6.55	6.11
English speaker	0.41	0.49	0.57	0.26	0.32	0.02	0.43	0.16	0.07	0.47	0.51	0.58	0.57	0.33	0.35

Table 13 Questions on (main)reasons for learning German

(i) Why are you learning German? (Multiple choices)	Labels (in the paper)
<p><input type="checkbox"/> Study/education/training/PhD^b</p> <p><input type="checkbox"/> Internal company communication</p> <p><input type="checkbox"/> German-speaking trading partners</p> <p><input type="checkbox"/> German-speaking customers</p> <p><input type="checkbox"/> Higher income in the [country]</p> <p><input type="checkbox"/> Requirement/support of the employer</p> <p><input type="checkbox"/> Other considerations regarding career/labour market in [country]</p> <p><input type="checkbox"/> Partner or family</p> <p><input type="checkbox"/> Social environment/friends</p> <p><input type="checkbox"/> (Possible) move to a German-speaking country for professional reasons^a</p> <p><input type="checkbox"/> (Possible) move to a German-speaking country for other reasons^a</p> <p><input type="checkbox"/> Requirement for visa^a</p>	<p>Education</p> <p>Company communication</p> <p>Trading partners</p> <p>Customers</p> <p>Income</p> <p>Employer</p> <p>Labour market</p> <p>Family</p> <p>Friends</p> <p>Educational and labour migr.</p> <p>Other migration</p> <p>Holiday</p> <p>Culture (film, literature,...)</p> <p>Culture</p> <p>Interest in languages</p> <p>Interest</p> <p>Other: [free-text]</p>
(ii) Look at your answers and circle the main reason why you are learning German.	

Notes:

^a Note that in Japan, Bosnia and Herzegovina, Great Britain and Poland we have not distinguished between “(Possible) move to a German-speaking country for professional reasons” and “(Possible) move to a German-speaking country for other reasons”. In these cases, we have imputed the reasons by making use of the main reason for a potential move to a German-speaking country, which the respondents answered in the survey as well. The same method was applied for the category “Requirement for visa” in all surveys.

^b In Indonesia, we split the category “Study/education/training/PhD” into “Study/education/training/PhD” and “(Vocational) Training”, but re-merged it for our analysis. Other reasons were categorized according to the free-text field if possible

Table 14 Tetrachoric correlation coefficients

	Interest	Culture	Holiday	Friends	Family	Other migrat.	Educ., labour migration	Educat.	Labour market	Income	Employer	Customer	Trade partner
Interest													
Culture	0.52***												
Holiday	0.30***	0.52***											
Friends	0.27***	0.33***	0.41***										
Family	-0.14***	-0.09***	0.06***	0.24***									
Other migration	0.10***	0.09***	0.17***	0.23***	0.37***								
Educ., lab. mig.	0.04***	0.07***	-0.04**	0.07***	-0.24***	0.03**							
Education	-0.13***	-0.06***	-0.17***	-0.08***	-0.38***	-0.09***	0.27***						
Labour market	0.10***	0.06***	0.09***	0.05***	-0.16***	0.00	0.10***	-0.04***					
Income	0.03**	0.03*	0.04***	0.08***	-0.22***	-0.01	0.21***	0.11***	0.31***				
Employer	0.00	0.05***	0.05***	0.10***	-0.08***	-0.08***	-0.02	-0.12***	0.17***	0.25***			
Customer	-0.02	0.00	0.04***	0.04***	-0.10***	-0.09***	0.05***	-0.19***	0.15***	0.36***	0.45***		
Trade partner	-0.12***	-0.04**	0.08***	0.04***	-0.13***	-0.04***	0.03**	-0.13***	0.11***	0.30***	0.42***	0.70***	
Company comm.	-0.07***	-0.03*	-0.04***	0.06***	-0.09***	-0.07***	0.07***	-0.10***	0.13***	0.28***	0.51***	0.61***	0.57***

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 15 Tetrachoric correlation coefficients for categories

	Cultural interest	Education	Domestic labour market	Education, labour migration
Cultural interest				
Education	- 0.17***			
Domestic labour market	0.04***	- 0.06***		
Education, labour migr.	0.04***	0.27***	0.15***	
Personal reasons	0.16***	- 0.19***	- 0.11***	- 0.01

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 16 Human capital investment: basic specifications with country FE interacted with gender and age

	Dependent variable: Human capital investment		
	(1)	(2)	(3)
Age: under 35 years	0.157*** (0.023)	0.025 (0.085)	0.019 (0.085)
Gender: female	− 0.082*** (0.013)	− 0.083*** (0.013)	− 0.091 (0.068)
Children	− 0.071*** (0.025)	− 0.074*** (0.026)	− 0.074*** (0.026)
Partner (native German)	− 0.360*** (0.034)	− 0.356*** (0.034)	− 0.352*** (0.034)
Partner (other native)	− 0.022 (0.016)	− 0.020 (0.016)	− 0.020 (0.016)
Occ.: high appl./low comm. skills	0.106*** (0.027)	0.105*** (0.028)	0.103*** (0.028)
Occ.: high appl./high comm. skills	0.119*** (0.020)	0.120*** (0.020)	0.117*** (0.020)
Occ.: in education	0.186*** (0.026)	0.189*** (0.026)	0.187*** (0.026)
University degree	0.045*** (0.017)	0.047*** (0.017)	0.048*** (0.017)
Risk attitude	0.006** (0.003)	0.007** (0.003)	0.007** (0.003)
Patience	0.005* (0.003)	0.005* (0.003)	0.005* (0.003)
English speaker	− 0.013 (0.014)	− 0.011 (0.014)	− 0.009 (0.014)
Country FE	Yes	Yes	Yes
Country FE × Age	No	Yes	Yes
Country FE × Gender	No	No	Yes
McFadden Pseudo R2	0.17	0.17	0.17
Percent. correctly predicted	71.3	71.3	71.3
Observations	4,891	4,891	4,891

Average marginal effects. Heteroscedasticity robust White standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 17 Human capital investment: subsamples by country characteristics

		Dependent variable: Human capital investment											
	(1)	Language			Income			Geography				Culture	
		(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Culture		
											Germanic	non-Germanic	Lower-middle
Age: under 35 years	0.067 (0.047)	0.163*** (0.026)	0.204*** (0.055)	0.197*** (0.058)	0.195*** (0.027)	0.131*** (0.029)	0.145* (0.066)	0.115*** (0.043)	0.213*** (0.033)	0.139*** (0.032)			
Gender: female	-0.056** (0.025)	-0.086*** (0.015)	-0.118* (0.021)	-0.067*** (0.035)	-0.103*** (0.019)	-0.059*** (0.021)	-0.068*** (0.033)	-0.103*** (0.019)	-0.085*** (0.018)	-0.098*** (0.019)			
Children	-0.079 (0.056)	-0.076*** (0.029)	0.003 (0.047)	-0.043 (0.064)	-0.020*** (0.032)	-0.094*** (0.035)	-0.091 (0.061)	-0.007 (0.045)	-0.069** (0.034)	-0.079* (0.038)			
Partner (native German)	-0.348*** (0.061)	-0.332*** (0.046)	-0.501*** (0.080)	-0.497*** (0.076)	-0.511*** (0.035)	-0.276*** (0.036)	-0.286*** (0.099)	-0.486*** (0.069)	-0.463*** (0.051)	-0.369*** (0.043)			
Partner (other native)	-0.002 (0.038)	-0.020 (0.019)	-0.035** (0.030)	-0.087 (0.040)	-0.054 (0.022)	0.008 (0.023)	0.005 (0.038)	-0.061 (0.029)	-0.031 (0.022)	-0.019 (0.024)			
Occ.: high appl./low comm. skills	0.045 (0.071)	0.111*** (0.032)	0.173 (0.038)	0.051*** (0.075)	0.128** (0.039)	0.090 (0.043)	0.059*** (0.055)	0.190*** (0.041)	0.121*** (0.035)	0.135 (0.044)			
Occ.: high appl./high comm. skills	0.142*** (0.039)	0.122*** (0.024)	0.132* (0.030)	0.097*** (0.056)	0.126*** (0.028)	0.118*** (0.031)	0.086** (0.050)	0.121*** (0.027)	0.154*** (0.027)	0.122*** (0.031)			
Occ.: in education	0.249*** (0.065)	0.187*** (0.028)	0.239** (0.045)	0.147*** (0.066)	0.206*** (0.036)	0.182*** (0.040)	0.111*** (0.062)	0.309*** (0.038)	0.238*** (0.034)	0.204*** (0.040)			

Table 17 (continued)

Dependent variable: Human capital investment										
Language	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Income			Geography			Culture			
	Germanic			EU			Non-EU		Far	
	Germanic	non-Germanic	Lower-middle	Upper-middle	High	EU	European	Non-Euro-pean	Close	Far
University degree	0.024 (0.033)	0.036* (0.019)	0.041** (0.027)	0.121 (0.049)	0.066 (0.025)	0.032 (0.029)	0.012 (0.039)	0.014*** (0.025)	0.094* (0.023)	0.045* (0.026)
Risk attitude	0.006 (0.006)	0.006 (0.004)	0.011 (0.005)	-0.0002** (0.008)	0.007 (0.004)	0.004 (0.005)	0.002 (0.008)	0.0004** (0.005)	0.011 (0.004)	0.006 (0.005)
Patience	0.007 (0.006)	0.006* (0.003)	0.006 (0.005)	0.009 (0.007)	0.007 (0.004)	0.002 (0.004)	0.0003 (0.007)	0.006** (0.004)	0.009 (0.004)	0.004 (0.004)
English speaker	0.049 (0.038)	-0.019 (0.017)	-0.011 (0.025)	-0.003 (0.034)	-0.009 (0.021)	-0.016 (0.021)	-0.012 (0.035)	0.007 (0.024)	-0.017 (0.019)	-0.026 (0.021)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	957	3,720	775	1,807	2,309	2,040	754	2,097	2,627	2,264

Average marginal effects. Heteroscedasticity robust White standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$ For the groupings, see Table 9; culturally close countries to Germany are Bosnia and Herzegovina, Czech Republic, Indonesia, Italy, Japan, South Korea and Ukraine and culturally more distant countries are Great Britain, India, Mexico, Netherlands, Poland, Romania and Spain based on the sum of distances of the two culture distance measures (LTO and IVR)

Table 18 Professional use: subsamples by country characteristics

	(1) Language		(2) non-Germanic		(3) Income		(4) Upper-middle		(5) High		(6) Geography		(7) European		(8) non-European		(9) Culture		(10) Far	
											EU		non-EU							
Investment	0.215*** (0.042)	0.225*** (0.017)	0.152*** (0.026)	0.199*** (0.035)	0.257*** (0.022)	0.234*** (0.023)	0.172*** (0.035)	0.196*** (0.026)	0.198*** (0.021)	0.227*** (0.023)										
Age: under 35 years	-0.071 (0.049)	0.100*** (0.025)	0.101*** (0.054)	0.117* (0.052)	0.050* (0.028)	0.049* (0.028)	0.041 (0.053)	0.148*** (0.047)	0.062* (0.032)	0.073** (0.031)										
Gender: female	-0.0002 (0.032)	0.040** (0.016)	0.061 (0.023)	0.022*** (0.032)	0.013 (0.020)	0.025 (0.021)	0.110*** (0.031)	0.010 (0.022)	0.043** (0.019)	0.017 (0.020)										
Children	-0.044 (0.062)	-0.029 (0.028)	-0.003 (0.053)	0.024 (0.051)	-0.062* (0.034)	-0.065* (0.048)	0.013 (0.048)	-0.012 (0.055)	-0.024 (0.035)	-0.059 (0.038)										
Partner (native German)	0.068 (0.058)	0.134*** (0.036)	0.237 (0.047)	0.023*** (0.106)	0.092** (0.036)	0.082** (0.037)	0.143* (0.054)	0.189*** (0.059)	0.200*** (0.036)	0.050 (0.042)										
Partner (other native)	0.024 (0.049)	0.028 (0.018)	0.066 (0.032)	0.033** (0.036)	0.027 (0.023)	0.034 (0.023)	0.002 (0.035)	0.064** (0.031)	0.046** (0.023)	0.027 (0.025)										
Occ.: high appl./low comm. skills	0.126 (0.086)	-0.013 (0.038)	0.100 (0.061)	-0.050 (0.073)	-0.008 (0.045)	-0.061 (0.046)	0.031 (0.062)	0.111* (0.063)	0.052 (0.044)	-0.042 (0.050)										
Occ.: high appl./high comm. skills	-0.005 (0.056)	0.085*** (0.024)	0.170 (0.034)	0.048*** (0.048)	0.045 (0.031)	0.050 (0.031)	0.154*** (0.034)	0.068 (0.042)	0.138*** (0.028)	0.026 (0.033)										
Occ.: in education	0.069 (0.064)	0.126*** (0.027)	0.263* (0.042)	0.102*** (0.057)	0.070** (0.036)	0.066* (0.038)	0.154*** (0.052)	0.200*** (0.043)	0.146*** (0.033)	0.121*** (0.039)										
University degree	0.049 (0.039)	0.011 (0.019)	0.022 (0.029)	-0.018 (0.042)	-0.005 (0.026)	-0.021 (0.028)	-0.020 (0.032)	0.039 (0.030)	-0.013 (0.024)	0.016 (0.026)										

Table 18 (continued)

	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)					
	Language				Income				Upper-middle				High				Geography				Culture			
	Germanic	non-Germanic	Lower-middle	Upper-middle	High	EU	non-EU	European	non-European	Close	Far	European	non-European	Close	Far	European	non-European	Close	Far	European	non-European	Close	Far	
Risk attitude	-0.003 (0.008)	0.020*** (0.004)	0.022*** (0.006)	0.020*** (0.007)	0.020*** (0.005)	0.018*** (0.005)	0.030*** (0.007)	0.020*** (0.005)	0.020*** (0.005)	0.027*** (0.005)	0.012*** (0.005)	0.030*** (0.007)	0.020*** (0.005)	0.027*** (0.005)	0.012*** (0.005)	0.030*** (0.007)	0.020*** (0.005)	0.027*** (0.005)	0.012*** (0.005)	0.030*** (0.007)	0.020*** (0.005)	0.027*** (0.005)	0.012*** (0.005)	
Patience	-0.004 (0.007)	0.005 (0.003)	-0.005 (0.005)	0.003 (0.006)	0.008* (0.004)	0.005 (0.004)	0.004 (0.006)	0.004 (0.005)	0.004 (0.005)	0.006 (0.004)	-0.0003 (0.004)	0.004 (0.006)	0.004 (0.005)	0.006 (0.004)	0.006 (0.004)	0.004 (0.006)	0.004 (0.005)	0.006 (0.004)	0.006 (0.004)	0.004 (0.006)	0.004 (0.005)	0.006 (0.004)	-0.0003 (0.004)	
English speaker	0.074 (0.049)	0.015 (0.017)	-0.003 (0.027)	0.022 (0.031)	0.031 (0.022)	0.022 (0.021)	-0.009 (0.031)	0.032 (0.027)	0.032 (0.027)	0.012 (0.020)	0.019 (0.022)	-0.009 (0.031)	0.032 (0.027)	0.012 (0.020)	0.019 (0.022)	-0.009 (0.031)	0.032 (0.027)	0.012 (0.020)	0.019 (0.022)	-0.009 (0.031)	0.032 (0.027)	0.012 (0.020)	0.019 (0.022)	
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Observations	892	3,525	735	1,715	2,204	1,984	726	1,944	2,153	2,501	2,153	726	1,944	2,501	2,153	726	1,944	2,501	2,153	726	1,944	2,501	2,153	

Average marginal effects. Heteroscedasticity robust White standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$ For the groupings, see Table 9; culturally close countries to Germany are Bosnia and Herzegovina, Czech Republic, Indonesia, Italy, Japan, South Korea and Ukraine and culturally more distant countries are Great Britain, India, Mexico, Netherlands, Poland, Romania, and Spain based on the sum of distances of the two culture distance measures (LTO and IVR)

Table 19 Human capital investment: basic specifications without *education* as motive and with *education* as consumption motive

	Dependent variable: Human capital investment			
	without <i>education</i> as motive		<i>education</i> as consumption motive	
	(1)	(2)	(3)	(4)
Age: under 35 years	0.125*** (0.023)	0.156*** (0.023)	0.084*** (0.020)	0.109*** (0.020)
Gender: female	- 0.095*** (0.016)	- 0.079*** (0.016)	- 0.059*** (0.013)	- 0.039*** (0.013)
Children	- 0.066** (0.027)	- 0.046* (0.028)	- 0.046* (0.024)	- 0.027 (0.025)
Partner (native German)	- 0.309*** (0.029)	- 0.357*** (0.024)	- 0.234*** (0.022)	- 0.278*** (0.017)
Partner (other native)	- 0.013 (0.019)	- 0.019 (0.019)	0.003 (0.016)	- 0.0001 (0.017)
Occ.: high appl./low comm. skills	0.114*** (0.034)	0.110** (0.035)	0.092*** (0.033)	0.081** (0.035)
Occ.: high appl./high comm. skills	0.130*** (0.025)	0.131** (0.025)	0.104*** (0.024)	0.096*** (0.025)
Occ.: in education	0.104*** (0.029)	0.147*** (0.029)	- 0.008 (0.025)	0.025 (0.025)
University degree	0.058*** (0.021)	0.083*** (0.021)	0.055*** (0.017)	0.084*** (0.018)
Risk attitude	0.009** (0.004)	0.009** (0.004)	0.008*** (0.003)	0.008*** (0.003)
Patience	0.007** (0.003)	0.010*** (0.003)	0.005* (0.003)	0.009*** (0.003)
English speaker	- 0.025 (0.017)	- 0.042** (0.018)	- 0.024* (0.014)	- 0.041*** (0.014)
Germanic lang.		- 0.023 (0.023)		0.077*** (0.021)
European (Non-EU)		- 0.051 (0.043)		0.005 (0.038)
Non-European		- 0.003 (0.024)		-0.062*** (0.021)
Upper-middle income		0.013 (0.034)		0.096*** (0.030)
High income		- 0.095*** (0.030)		- 0.005 (0.023)

Table 19 (continued)

	Dependent variable: Human capital investment			
	without <i>education</i> as motive		<i>education</i> as consumption motive	
	(1)	(2)	(3)	(4)
Cultural distance: LTO		– 0.121** (0.053)		– 0.035 (0.048)
Cultural distance: IVR		0.031 (0.044)		– 0.012 (0.037)
Country FE	Yes	No	Yes	No
McFadden Pseudo R2	0.14	0.10	0.09	0.05
Percent. correctly predicted	67.2	63.5	69.7	68.0
Observations	3,651	3,651	4,891	4,891

Average marginal effects. Heteroscedasticity robust White standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 20 Professional use: basic specifications without *education* as motive and with *education* as consumption motive

	Dependent variable: Professional use of German			
	Without <i>education</i> as motive		<i>Education</i> as consumption motive	
	(1)	(2)	(3)	(4)
Investment	0.226*** (0.017)	0.239*** (0.016)	0.157*** (0.015)	0.168*** (0.014)
Age: under 35 years	0.081*** (0.023)	0.096*** (0.022)	0.099*** (0.022)	0.116*** (0.022)
Gender: female	0.036** (0.016)	0.041*** (0.016)	0.024* (0.014)	0.030** (0.014)
Children	− 0.027 (0.026)	− 0.013 (0.025)	− 0.042* (0.025)	− 0.025 (0.024)
Partner (native German)	0.098*** (0.030)	0.081*** (0.030)	0.074** (0.029)	0.049 (0.029)
Partner (other native)	0.033* (0.019)	0.031 (0.019)	0.032* (0.017)	0.026 (0.017)
Occ.: high appl./low comm. skills	0.021 (0.035)	0.020 (0.035)	0.024 (0.033)	0.023 (0.032)
Occ.: high appl./high comm. skills	0.080*** (0.023)	0.080*** (0.023)	0.098*** (0.021)	0.099*** (0.021)
Occ.: in education	0.110*** (0.028)	0.115*** (0.028)	0.170*** (0.025)	0.177*** (0.024)
University degree	− 0.023 (0.020)	− 0.027 (0.020)	− 0.002 (0.018)	− 0.005 (0.018)
Risk attitude	0.019*** (0.004)	0.018*** (0.004)	0.021*** (0.003)	0.020*** (0.003)
Patience	0.003 (0.003)	0.004 (0.003)	0.003 (0.003)	0.004 (0.003)
English speaker	0.013 (0.017)	0.015 (0.017)	0.016 (0.015)	0.017 (0.015)
Germanic lang.		− 0.175*** (0.023)		− 0.199*** (0.021)
Non-EU (European)		− 0.035 (0.043)		0.0002 (0.038)
Non-EU (Non-European)		− 0.172*** (0.022)		− 0.135*** (0.020)
Upper-middle income		− 0.030 (0.033)		− 0.023 (0.029)

Table 20 (continued)

	Dependent variable: Professional use of German			
	Without <i>education</i> as motive		<i>Education</i> as consumption motive	
	(1)	(2)	(3)	(4)
High income		- 0.193*** (0.029)		- 0.159*** (0.024)
Cultural distance: LTO		0.039 (0.049)		0.051 (0.045)
Cultural distance: IVR		0.096** (0.042)		0.121*** (0.038)
Country FE	Yes	No	Yes	No
McFadden Pseudo R2	0.16	0.15	0.13	0.13
Percent. correctly predicted	70.8	70.4	69.7	69.4
Observations	3,651	3,651	4,891	4,891

Average marginal effects. Heteroscedasticity robust White standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Appendix 2: Figures

See Figs. 5, 6, 7, 8 and 9.

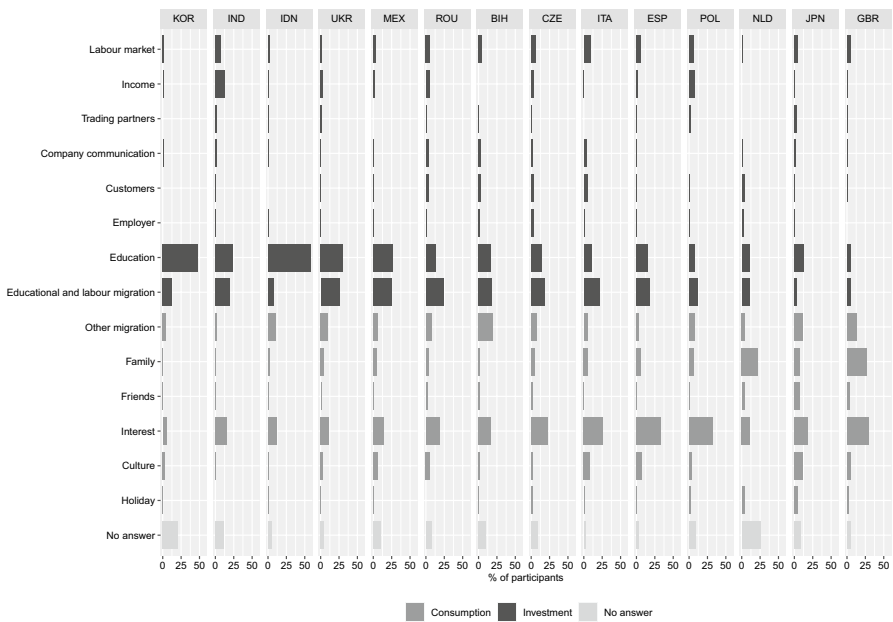


Fig. 5 Main reasons for learning German by countries (n = 4891)

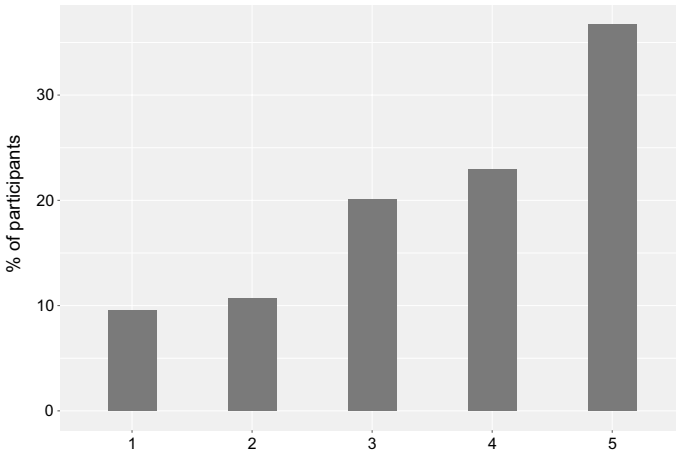


Fig. 6 Professional use of German: 1 = very unlikely, 5 = very likely (n = 4891)

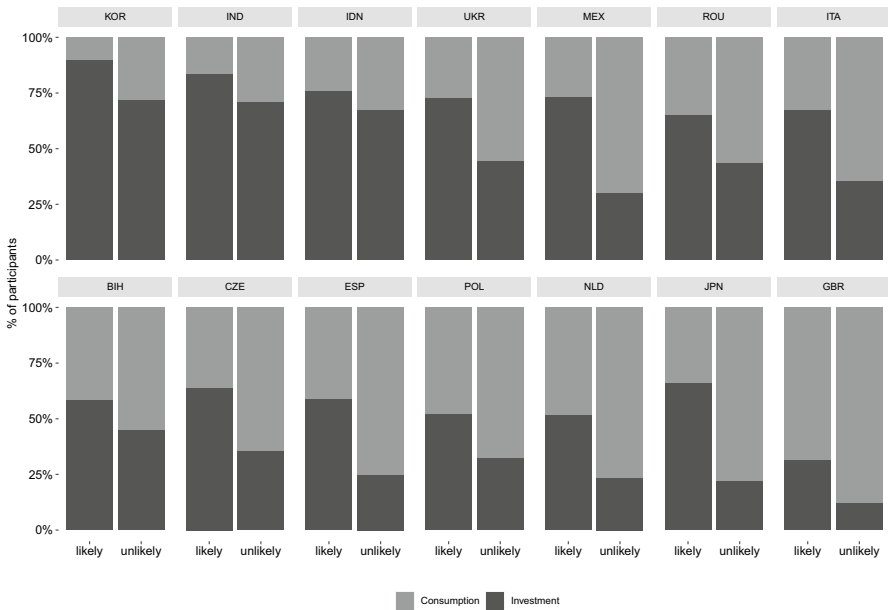


Fig. 7 Professional use of German: subsamples by countries and economic motive. Note: We aggregate the five point scale to the binary variable “Professional use of German” where values 1 to 3 correspond to “unlikely” and values 4 and 5 to “likely”

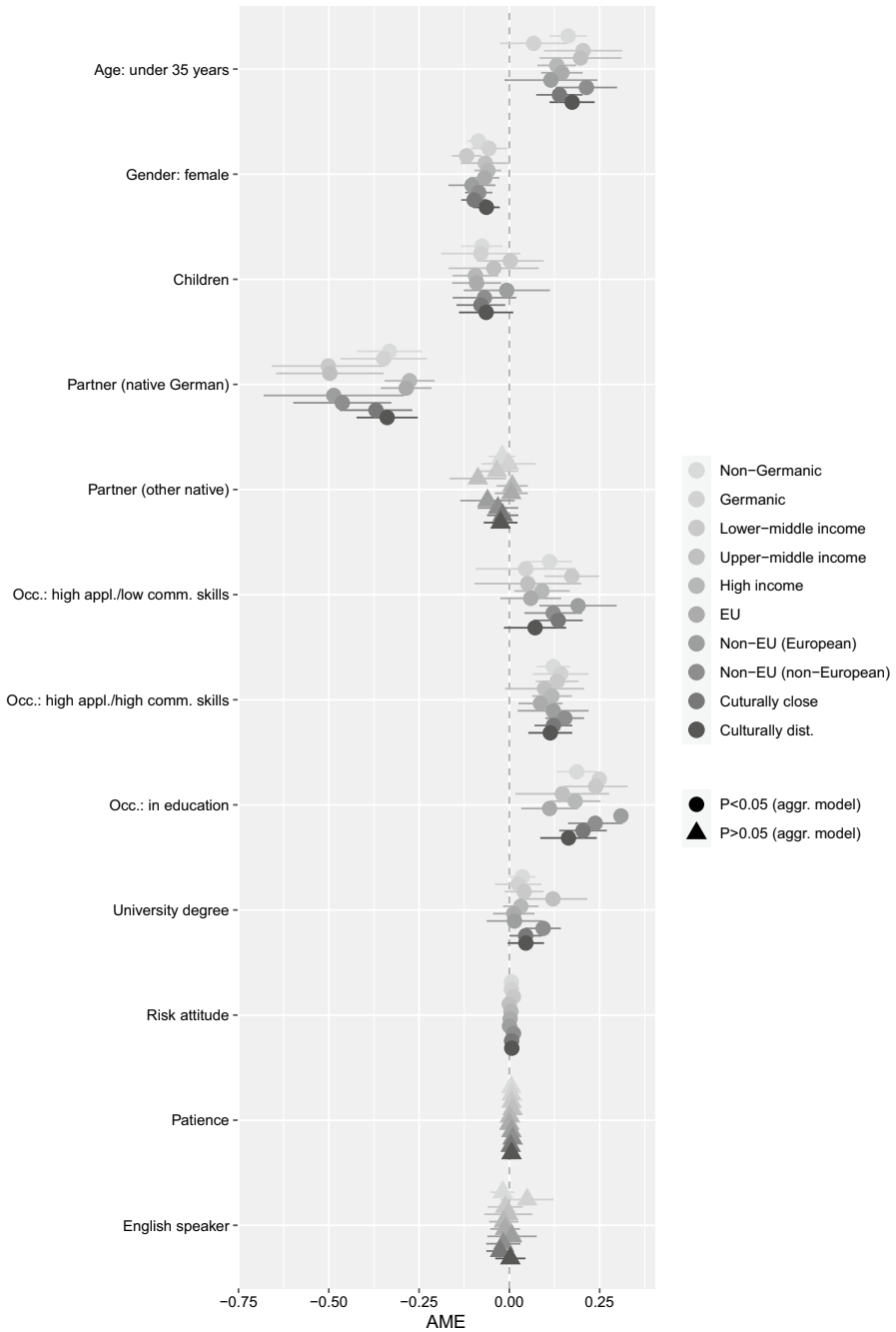


Fig. 8 Human capital investment: subsamples by country-groups according to country characteristics. Note: Average marginal effects and 95 % CI. Shapes according to the p-values of the aggregated model, see Table 3 Column 2. For detailed estimation results, see Table 17 in the Appendix

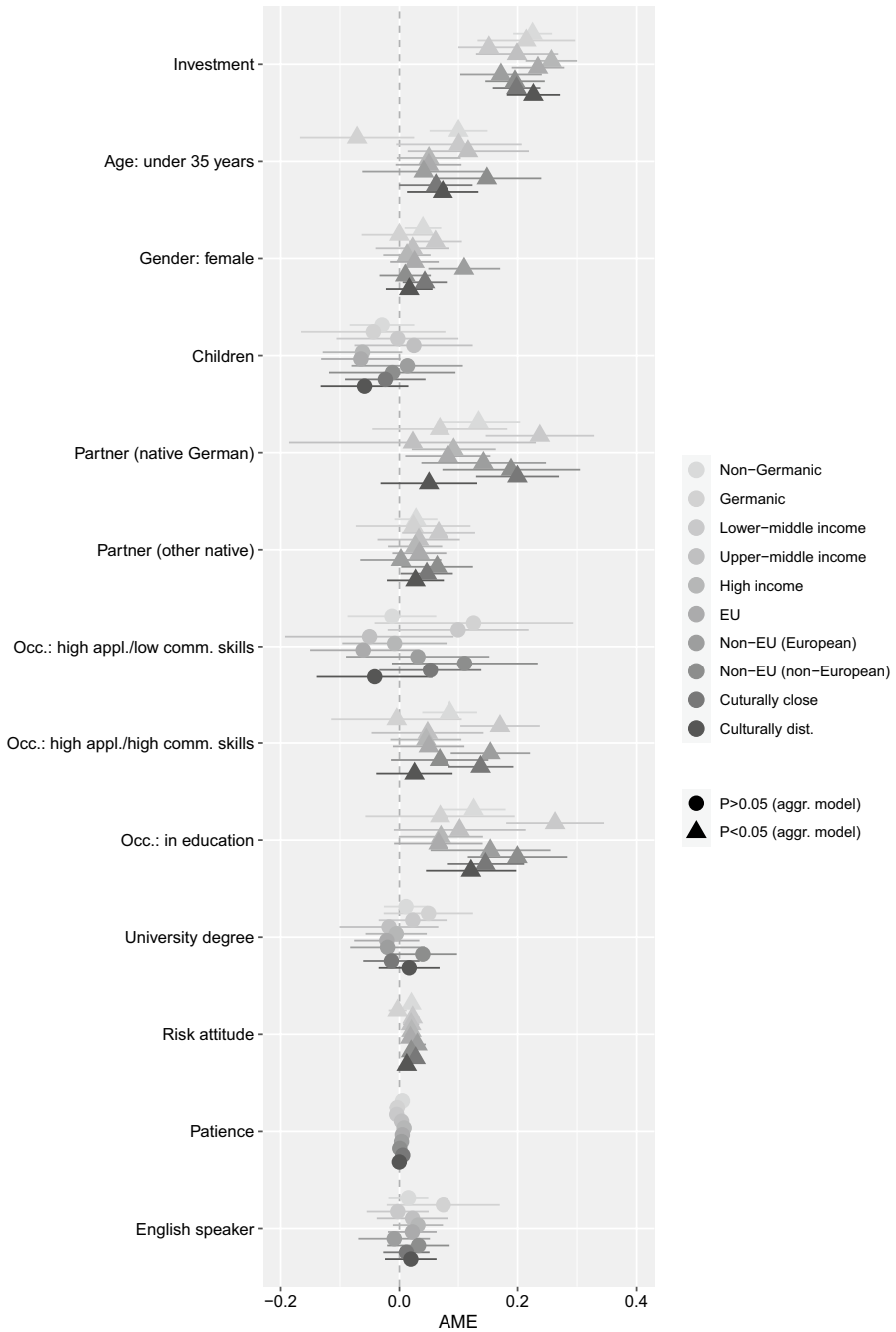


Fig. 9 Professional use: subsamples by country-groups according to country characteristics. Note: Average marginal effects and 95 % CI. Shapes according to the p-values of the aggregated model, see Table 5 Column 3. For detailed estimation results see Table 18 in the Appendix

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10663-022-09548-7>.

Funding Open Access funding enabled and organized by Projekt DEAL. Funding was provided by Deutsche Forschungsgemeinschaft; grant no. UE 124/2-2–270886786.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aldashev A, Gernandt J, Thomsen SL (2009) Language usage, participation, employment and earnings: evidence for foreigners in West Germany with multiple sources of selection. *Labour Econ* 16(3):330–341
- Alstadsaeter A (2011) Measuring the consumption value of higher education. *CESifo Econ Stud* 57(3):458–479
- Arcidiacono P (2004) Ability sorting and the returns to college major. *J Econom* 121(1–2):343–375
- Becker GS (1964) Human capital: a theoretical and empirical analysis, with special reference to education. National Bureau of Economic Research, New York
- Budría S, Colino A, de Ibarreta CM (2019) The impact of host language proficiency on employment outcomes among immigrants in Spain. *Empirica* 46(4):625–652
- Canton E, de Jong F (2005) The demand for higher education in The Netherlands, 1950–1999. *Econ Educ Rev* 24(6):651–663
- Chiswick BR, Miller PW (1995) The endogeneity between language and earnings: international analyses. *J Law Econ* 13(2):246–288
- Dustmann C, Fabbri F (2003) Language proficiency and labour market performance of immigrants in the UK. *Econ J* 113(489):695–717
- Dustmann C, van Soest A (2001) Language fluency and earnings: estimation with misclassified language indicators. *Rev Econ Stat* 83(4):663–674
- Geist C, McManus PA (2011) Different reasons, different results: implications of migration by gender and family status. *Demography* 49(1):197–217
- Ginsburgh VA, Prieto-Rodríguez J (2011) Returns to foreign languages of native workers in the European Union. *ILR Rev* 64(3):599–618
- Isphording IE (2013) Returns to foreign language skills of immigrants in Spain. *Labour* 27(4):443–461
- Kodde DA, Ritzén JMM (1984) Integrating consumption and investment motives in a neoclassical model of demand for education. *Kyklos* 37(4):598–608
- Lazear E (1977) Education: Consumption or production? *J Polit Econ* 85(3):569–597
- Mincer J (1978) Family migration decisions. *J Polit Econ* 86(5):749–773
- Oosterbeek H, van Ophem H (2000) Schooling choices: preferences, discount rates, and rates of return. *Empir Econ* 25(1):15–34
- Oosterbeek H, Webbink D (1995) Enrolment in higher education in The Netherlands. *De Econ* 143(3):367–380
- Saiz A, Zoido E (2005) Listening to what the world says: bilingualism and earnings in the United States. *Rev Econ Stat* 87(3):523–538
- Schaafsma J (1976) The consumption and investment aspects of the demand for education. *J Hum Resour* 11(2):233
- Schultz TW (1960) Capital formation by education. *J Polit Econ* 68(6):571–583

- Sjaastad LA (1962) The costs and returns of human migration. *J Polit Econ* 70(5):80–93
- Stöhr T (2015) The returns to occupational foreign language use: evidence from Germany. *Labour Econ* 32:86–98
- Toomet O (2011) Learn English, not the local language! Ethnic Russians in the Baltic States. *Am Econ Rev* 101(3):526–31
- Uebelmesser S, Huber M, Weingarten S (2018) The German language worldwide: a new data set on language learning. *CESifo Econ Stud* 64(1):103–121
- Auswärtiges Amt and Goethe-Institut (2004) *Rahmenvertrag*. Munich and Berlin
- Chiswick BR, Miller PW (2015) Chapter 5 - International Migration and the Economics of Language. In: *Handbook of the Economics of International Migration*. In: Chiswick BR, Miller PW (eds) *Handbook of the Economics of International Migration*, vol 1. North-Holland, pp 211–269
- Goethe-Institut (2014) *Jahrbuch 2013/14*. Munich
- Goethe-Institut (2021) *Jahrbuch 2020/21*. Munich
- Hofstede G, Minkov M (2013) *Values survey module*
- Huber M, Uebelmesser S (2019) Presence of language-learning opportunities and migration. *CESifo Working Paper* 7569
- Uebelmesser S, Sommerfeld AM, Weingarten S (2022) A macro-level analysis of language learning and migration. *German Economic Review* 23(2): 181–232
- World Bank (2021) *World development indicator*

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.