Thema

Empirical Studies on Language-Learning, Migration and
Trade

Dissertation
zur Erlangung des akademischen Grades
doctor rerum politicarum
(Dr. rer. pol.)

vorgelegt dem
Rat der Wirtschaftswissenschaftlichen Fakultät
der Friedrich–Schiller–Universität Jena

am 15.12.2021

von: M.Sc. Matthias Huber

geboren am: 15. August 1987   in: Passau

Gutachter

1. Prof. Dr. Silke Übelmesser,
    Friedrich-Schiller-Universität Jena,
    Lehrstuhl für AVWL/Finanzwissenschaft

2. Prof. Dr. Bernd Suessmuth, Professor für
    Volkswirtschaftslehre / Ökonometrie, Universität
    Leipzig

Datum der Verteidigung: 18.05.2022

Contents

List of Figures

vi

List of Tables

Acknowledgements

Co-Authorship and Statement of Contribution

This thesis is a publication-based – cumulative – dissertation. Chapters 2 to 6 rest on five individual empirical research papers.

The research project 'Language Learning and Migration' supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) has resulted in three papers for this dissertation:

- Chapter 2 is based on a paper titled 'The German Language Worldwide: A New Dataset on Language Learning', which is published in *CESifo Economic Studies* 64(1), 103-121. It is co-authored by Silke Uebelmesser (CESifo, University of Jena) and Severin Weingarten (University of Jena). The research was initiated by Silke Uebelmesser. While Severin Weingarten and I collected the data and prepared the datasets for publication, I implemented the empirical analysis. The writing of the paper was done by Silke Uebelmesser and me to about equal parts.

- Chapter 3 builds on 'Presence of language-learning opportunities and migration' (CESifo Working Papers No. 2019-7569) which is joint work with Silke Übelmesser. While Silke Uebelmesser mainly developed the research question and contributed to a collaborative development of the paper, I focused on the preparation of the dataset and conducted the empirical analysis. My contribution to the drafting of the manuscript and final revisions was slightly larger than that of Silke Übelmesser.

- Chapter 4 is based on a paper titled 'Learn German, Buy German? Language-learning opportunities abroad and exports' (Jena Economic Research Papers (JERP) No. 2021-008). It is now in the revision and resubmission process at *The World Economy*. This paper is joint work together with Omar Martin Fieles-Ahmad (Otto von Guericke University Magdeburg). I have developed the research idea and initiated the research

myself. My contribution to the data preparation and implementation of the empirical analysis was larger than that of Omar Martin Fieles-Ahmad. The responsibility for writing the paper was shared equally.

Within the framework of the follow-up project 'Language-Skill Investments and Migration Decisions' which was also supported by the DFG, two surveys were conducted: one among language course participants at Goethe-Institutes (GI) and one among university students, both in various countries all over the world. While the questionnaire was developed jointly by Silke Uebelmesser, Panu Poutvaara, Till Nikolka and me, my main responsibility was the conduction and data preparation of the survey at the GI. In this context, Chapters 5 and 6 were prepared:

- Chapter 5 rests on the paper titled 'Language Learning: Human Capital Investment or Consumption?'(Jena Economic Research Papers (JERP) No. 2021-019) which is co-authored by Silke Uebelmesser. It is now in the revision and resubmission process at *Empirica*. While the contribution to the development of the research idea and the draft of the manuscript was about equal, my main responsibility was the data collection and preparation (survey at GI) as well as the implementation the empirical analysis.

- Chapter 6 builds on the yet unpublished paper 'Migration aspirations and intentions' which is joint work with Silke Übelmesser, Ann-Marie Sommerfeld (Goettingen University), Till Nikolka (German Youth Institute) and Panu Poutvaara (University of Munich, ifo Institute, CESifo, CReAM and IZA). Within the framework of the research project, the main responsibility for this paper was by Silke Uebelmesser and me. We both mainly developed the research question. While I focussed on the collection and preparation of the GI institute data, the data for survey among university students was provided by Till Nikolka and Panu Poutvaara. Ann-Marie Sommerfeld implemented the empirical analysis and drafted the largest part of manuscript with my and Silke Uebelmesser's support. All co-authors contributed to collaborative development of the paper.

Signature:

German Summary

Kenntnisse in einer Sprache sind eine fundamentale Fähigkeit, die in ganz Europa an Schulen gelehrt wird und dort meist obligatorisch sind. Im Gegensatz zu Schülerinnen und Schülern entscheiden sich Erwachsene aktiv für den Erwerb von Sprachkenntnissen. Was die Vorteile von Sprachkenntnissen sind und was sie determiniert, kann aus zwei Blickwinkeln betrachtet werden: aus der aggregierten Makro-Perspektive und aus der individuellen Mikro-Perspektive. Die Kapitel 2 bis 4 dieser Dissertationschrift beschäftigen sich mit diesem Zusammenhang aus der Makro-Perspektive, Kapitel 5 aus der Mikro-Perspektive. Während sich die Kapitel 2 bis 5 hauptsächlich mit dem Spracherwerb beschäftigen, der in den Herkunftsländern stattfindet bevor die Migration erfolgt, befasst sich Kapitel 6 hingegen nun allgemein mit potenziellen Migrantinnen und Migranten in ihren Herkunftsländern und mit ihren Migrationsaspirationen und -absichten, losgelöst vom Spracherwerb.

Wenn man Sprachkenntnisse aus der Makro-Perspektive betrachtet, konzentriert sich die Literatur dazu vor allem auf Migrations- und Handelsströme. Dies gilt sowohl für die Determinanten als auch für die Vorteile. Bei der Betrachtung der Vorteile von Sprachkenntnissen im Bezug auf Migration und Handel werden hauptsächlich linguistische Eigenschaften der Amts- oder Verkehrssprachen der beteiligten Länder verwendet. Nur wenige Ausnahmen zeigen, dass auch erworbene Sprachkenntnisse Handels- und Migrationsströme erhöhen können. Der Grund für das weitestgehende Fehlen von Studien zu Sprachkenntnissen aus der Makro-Perspektive ist die schwere Messbarkeit von Fremdsprachenkenntnissen und einem damit einhergehenden Mangel an Daten, die eine quantitative Analyse erlauben. Spracherwerb bzw. Sprachkenntnisse sind in vielerlei Hinsicht heterogen, angefangen beim Sprachniveau, den Motiven für den Spracherwerb (siehe dazu Kapitel 5) und dem Kontext, in dem die Sprachkenntnisse erworben werden - insbesondere ob innerhalb oder außerhalb des Schulsystems.

Kapitel 2 trägt dazu bei, den Mangel an Daten zu überwinden, indem es Daten zum Erwerb der deutschen Sprache außerhalb des Schulsystems für den Zeitraum von 1965 bis 2014 präsentiert. Die Datensätze sind die ersten, die es ermöglichen, die Entwicklung des Erwerbs der deutschen Sprache weltweit und im Laufe der Zeit nachzuvollziehen. Anhand einer deskriptiven Analyse werden die Entwicklungen der einzelnen Weltregionen in den drei Datensätzen dargestellt.

Die drei Datensätze stammen aus den Jahresberichten des Goethe-Instituts (GI), einem wichtigen Akteur der deutschen Auswärtigen Kulturpolitik, dessen Aufgabe die Förderung der deutschen Kultur und Sprache weltweit ist. Dies wird erfüllt, indem das GI Sprachdienstleistungen anbietet und über die deutsche Kultur und Gesellschaft informiert. Der erste Datensatz enthält Informationen über die Angebotsseite des Spracherwerbs und stellt diese anhand der Präsenz von GI auf Stadtebene für den Zeitraum von 1965 bis 2014 dar. Die beiden anderen Datensätze konzentrieren sich auf die Nachfrage nach Sprachdienstleistungen. Der eine Datensatz enthält Informationen über die Teilnahme an Sprachkursen (Anmeldungen von 1990 bis 2014 sowie verkaufte Kurseinheiten von 1972 bis 1989 und 1997 bis 2014) und über Prüfungen an ausländischen Instituten (1986 bis 2014). Der letzte Datensatz beinhaltet detaillierte Informationen über die jährliche Zahl der Anmeldungen für Sprachkurse in Deutschland, aufgegliedert nach Staatsangehörigkeit für die Jahre 1966 bis 2014.

In Kapitel 3 wird nun der Datensatz zur Präsenz der Sprachinstitute verwendet, um den Zusammenhang zwischen Sprachlernangeboten und Migration nach Deutschland zu untersuchen. Potenzielle Migrantinnen und Migranten könnten die Vorteile von Sprachkenntnissen bei der Integration in den Arbeitsmarkt und die Gesellschaft in Betracht ziehen, wenn sie Migrationsentscheidungen treffen und sich für ein Zielland entscheiden. Während Aparicio Fenoll und Kuehn (2016) einen positiven Zusammenhang zwischen dem obligatorischen Spracherwerb an Schulen und Migrationsströmen innerhalb der Europäischen Union feststellen, gibt es unseres Wissens keine Belege für die Auswirkungen des Spracherwerbs von Erwachsenen auf Migration. Daher trägt Kapitel 3 dazu bei, diese Lücke in der Literatur zu schließen, indem es die Auswirkungen von Sprachlernangeboten für Erwachsene auf Migration nach Deutschland untersucht. Es wird dazu ein Panel-Datensatz mit 69 Ländern für die Jahre 1977 bis 2014 verwendet.

Mit Hilfe von Fixed-Effects(FE)-Schätzungen wird gezeigt, dass die Anzahl der Sprachinstitute in einem Land positiv mit der Migration aus diesem Land

nach Deutschland korreliert ist. Durch die Unterscheidung von Instituten mit und ohne Sprachlernangeboten ist es möglich zu zeigen, dass diese Korrelation durch Sprachkurse getrieben ist.

Um die Ergebnisse kausal interpretieren zu können, befassen wir uns mit der Frage der umgekehrten Kausalität. Betrachtet man die Strategien und Ziele des GI in unserem Beobachtungszeitraum, so wird deutlich, dass Zuwanderung nach Deutschland keine besondere Rolle bei der Eröffnung der Institute spielte. Trotzdem schätzen wir auch den Effekt von Sprachinstituten auf die Migrationsströme in die Schweiz, da Migration in die Schweiz als exogen zur Anzahl der Institute angesehen werden kann. Wir finden ähnliche Ergebnisse für Migration in die deutschsprachige Schweiz und interpretieren die Ergebnisse daher als kausal. Darüber hinaus zeigen Schätzungen mit dem Common-Correlated-Effects(CCE)-Schätzer (Pesaran, 2006), dass die Hälfte des Gesamteffekts ein direkter Effekt auf die Migration von Personen nach Deutschland ist, die sonst nicht zugewandert wären. Die andere Hälfte ist ein indirekter Effekt von umgeleiteten Migrationsströmen ins Zielland Deutschland. Mit dem CCE-Schätzer kann für die Multilateral Resistance to Migration (Bertoli und Fernández-Huertas Moraga, 2013) kontrolliert werden, d.h. dass Migrationsströme nicht mehr nur, wie in FE-Schätzungen, von den Merkmalen des Herkunfts- und Ziellandes abhängen, sondern auch von der Attraktivität alternativer Ziele.

In Kapitel 4 wird ebenso unter Verwendung der Daten zur Präsenz der Sprachinstitute untersucht, ob sich Sprachlernangebote auch auf deutsche Exporte auswirken. Der Paneldatensatz enthält jährliche Beobachtungen zu 134 Ländern für den Zeitraum von 1978 bis 2014. Sprachkenntnisse senken die Transaktionskosten der Handelspartner. Diese Ansicht wird durch Studien über die Beziehung zwischen den linguistischen Eigenschaften der Sprachen der Handelspartner bzw. den erworbenen Sprachkenntnissen und Handelsströmen gestützt. Mit unseren Ergebnissen liefern wir weitere Belege für die Bedeutung von erworbenen Sprachkenntnissen für internationale Handelsströme und zeigen ein politisches Instrument zur Exportförderung auf.

Mit Hilfe von FE-Schätzungen stellen wir fest, dass die Anzahl der Sprachinstitute in einem Land positiv mit deutschen Exporten in dieses Land korreliert. Dieser Zusammenhang ist nicht-linear in der Anzahl der Institute, d.h. es gibt einen positiven, aber marginal abnehmenden Einfluss der Institute. Darüber hinaus befassen wir uns mit dem Problem der umgekehrten Kausalität, indem wir den Effekt der deutschen Institute auf die Schweizer Exporte schätzen. Die Idee und Umsetzung ist die gleiche wie in Kapitel 3: Die Entscheidung, ein

Institut zu eröffnen, ist exogen zu Schweizer Exporten. Eine steigende Anzahl an Instituten erhöht Exporte aus der deutschsprachigen Schweiz in die Länder, die diese Institute beherbergen.

Unsere Ergebnisse bestätigen Lien und Lo (2017) und Lien et al. (2019), die eine positive Beziehung zwischen der Eröffnung verschiedener Kulturinstitute (z.B. GI für Deutschland und Cervantes-Institut für Spanien) und Handel zeigen. Während sie aber Kulturinstitute als Mittel zur Steigerung der Soft Power (Nye, 2004) ansehen, fokussieren wir uns auf den Sprachkanal, der Transaktionskosten senkt. Transaktionskosten sind jedoch bei weitem nicht für alle Waren und Handelspartner gleich. Besonders hoch sind sie beim Handel mit Gütern, die eine umfangreiche schriftliche und mündliche Kommunikation erfordern. Mit den verwendeten Daten, die eine Unterscheidung zwischen Instituten erlauben, die Sprachdienstleistungen anbieten, und solchen, die dies nicht tun, kann gezeigt werden, dass nur Sprachinstitute signifikant mit Exporten korrelieren. Darüber hinaus verwenden wir Rauchs (1999) Produktkategorien, um Exportströme nach dem unterschiedlichen Kommunikationsbedarf der gehandelten Güter und damit nach ihren Transaktionskosten zu unterscheiden. Einerseits stellen wir fest, dass es keinen signifikanten Zusammenhang zwischen der Anzahl der Institute und dem Export von homogenen Gütern gibt, für die nur wenig Kommunikation notwendig ist. Andererseits steigt der Export von differenzierten Gütern, die ein hohes Maß an Kommunikation beim Handel erfordern, mit der Anzahl der Institute. Daraus folgern wir, dass Sprache der relevante Kanal ist.

Aus individueller Perspektive werden die Vorteile von Sprachkenntnissen in zwei Kontexten untersucht: Sprachkenntnisse von Migrantinnen und Migranten im Zielland und deren positiver Einfluss auf die Integration in Arbeitsmarkt und Gesellschaft, sowie Sprachkenntnisse auf dem heimischen Arbeitsmarkt. Aufgrund der Vorteile für den ausländischen und inländischen Arbeitsmarkt können Fremdsprachenkenntnisse als Humankapital betrachtet werden. Die Humankapitaltheorie könnte möglicherweise eine Erklärung bieten, warum sich eine Person für den Erwerb von Fremdsprachenkenntnissen entscheidet. Es gibt allerdings nur wenig Forschung über die Gründe des Spracherwerbs. Am Nähesten zu dieser Frage sind Studien, die sich auf die Determinanten der Fremdsprachenkenntnisse von Migrantinnen und Migranten konzentrieren. Diese Gründe können unter den drei E's zusammengefasst werden: Exposition (exposure), Effizienz (efficency) und wirtschaftliche Anreize (economic incentives) (Chiswick und Miller, 2015). Letzteres bezieht sich auf das Motiv der Humankapitalinvestition, andere Gründe werden weitgehend ignoriert.

Kapitel 5 schließt an dieser Lücke an und befasst sich mit der Entscheidung zum Spracherwerb und den verschieden dahinterstehenden Motiven. Die Motive basieren auf dem Lernen als Humankapitalinvestition oder als Konsum. Um der Frage nachzugehen, warum sich Menschen für den Erwerb von Fremdsprachenkenntnissen entscheiden, verwenden wir Umfragedaten, die wir in 14 Ländern weltweit von Sprachkursteilnehmenden des Goethe-Instituts erhoben haben.

Fremdsprachenkenntnisse haben auf dem Arbeitsmarkt einen produktiven Wert und können somit aus der Perspektive der Humankapitaltheorie (Becker, 1964) betrachtet werden. Diese Theorie besagt, dass Individuen das Humankapital wählen, das ihren erwarteten Nettobarwert des Einkommens nach einer Kosten-Nutzen-Abwägung maximiert. Sjaastad (1962) bezieht Migrationsentscheidungen in die Humankapitaltheorie durch einen Vergleich der erwarteten Erträge auf dem ausländischen und heimischen Arbeitsmarkt mit ein. Empirisch lassen sich jedoch Bildungsentscheidungen von Individuen nicht vollständig durch die Humankapitaltheorie bzw. die erwarteten Vorteile auf dem Arbeitsmarkt erklären. Eine mögliche Erklärung für dieses Verhalten ist, dass Bildung einen Konsumwert hat und einen direkten Nutzen erzeugt (Lazear, 1977; Schaafsma, 1976). Individuen entscheiden sich dann für eine Menge bzw. Art von Bildung, die nicht zu optimalen monetären Erträgen führt (Alstadsaeter, 2011; Arcidiacono, 2004).

Wir präsentieren Evidenz für Konsum- und Investitionsmotive. Diese Motive bauen auf den von den Teilnehmenden angegebenen Hauptgründen für ihren Spracherwerb auf. In Bezug auf Länder und Altersgruppen gibt es starke Heterogenitäten der Motive. Mit binären Probit-Schätzungen werden individuelle und länderspezifische Determinanten des Investitionsmotivs ermittelt. Der Lebensmittelpunkt in einem europäischen Land, in einem Land mit einer germanischen Amtssprache oder in einem Land mit einem höheren Durchschnittseinkommen ist mit einer geringeren Wahrscheinlichkeit des Investitionsmotivs verbunden. Auf individueller Ebene gibt es starke Heterogenitäten der Determinanten in Bezug auf Ländergruppen, Alter und Geschlecht.

Mit der erwarteten Wahrscheinlichkeit der Verwendung der Fremdsprachenkenntnisse in einem beruflichen Umfeld können wir zeigen, dass ein bestimmtes Motiv für Humankapitalinvestitionen nicht unbedingt mit einer hohen Wahrscheinlichkeit der beruflichen Verwendung der Sprache einhergeht. Bei binären Probit-Schätzungen zeigen außereuropäische Länder und Länder mit niedrigerem Einkommen nun eine negative Korrelation, während sie für das Investiti-

onsmotiv positiv war, d.h. Befragte aus diesen beiden Ländergruppen geben im Durchschnitt eher ein Investitionsmotiv an, schätzen aber gleichzeitig die Wahrscheinlichkeit einer beruflichen Nutzung der deutschen Sprache geringer ein. Auf individueller Ebene lassen sich wiederum Heterogenitäten in einer Analyse nach Alter, Geschlecht und Ländergruppen feststellen. Bei der Betrachtung von möglichen Spillover-Effekten von einem bestimmten Konsummotiv auf eine hohe Wahrscheinlichkeit einer berufliche Nutzung stellen wir fest, dass diese vor allem in einem "tied-mover"-Kontext bei jüngeren Frauen mit einem deutschen Partner auftreten.

Diese Dissertationschrift ist in den Kontext potenzieller Migrantinnen und Migranten in ihrem Herkunftsland und dem Sprachenlernen an Goethe-Instituten weltweit eingebettet. Während sich die Kapitel 2 bis 5 auf den Spracherwerb in den Herkunftsländern vor der Migration konzentrieren, fokussiert Kapitel 6 sich auf potenzielle Migrantinnen und Migranten in ihrem Herkunftsland, losgelöst vom Spracherwerb.

Ein großer Teil der ökonomischen Migrationsliteratur legt den Schwerpunkt auf Migrantinnen und Migranten in den Zielländern. Dies hat jedoch den Nachteil, dass es sich bei den Migrantinnen und Migranten um eine Selektion derjenigen handelt, die tatsächlich migriert sind. Wenn man sich hingegen auf Umfragedaten unter Personen in ihrem Herkunftsland stützt, bevor die tatsächliche Migration stattfindet, kann dieser Nachteil überwunden werden, da potenzielle Migrantinnen und Migranten identifiziert werden können. Dafür werden häufig Daten des Gallup World Poll (GWP) verwendet. Ziel des Kapitels 6 ist es, Einschränkungen, die dieser Datensatz mit sich bringt, zu beheben und eine umfassende Analyse von Migrationsaspirationen und -absichten zu liefern. Der besondere Augenmerk wird auf diejenigen potenziellen zukünftigen Migrantinnen und Migranten gelegt, die durch die Art der Fragenstellung im GWP oft übersehen werden.

Potenzielle Migration wird einerseits durch Migrationsaspirationen gemessen, die den Wunsch ausdrücken, unter idealen Umständen zu migrieren, anderseits durch Migrationsabsichten, die sich auf Überlegungen und Pläne basierend auf der tatsächlichen Situation beziehen. Der GWP benutzt beide Konzepte. Die Daten sind jedoch mit zwei wichtigen Einschränkungen verbunden: Erstens filtert die Frage zu den Migrationsaspirationen die folgenden Fragen zu Migrationsabsichten, so dass diese nur beantwortet werden, wenn die Befragten Migrationsaspirationen angeben. Dieser Filter erzeugt eine Konditionalität und nimmt fälschlicherweise an, dass Migrationsaspirationen für Migrations-

absichten notwendig sind. Zweitens sind die Fragen zu den Aspirationen und Absichten auf einen permanten Zeithorizont beschränkt und vernachlässigen daher potenzielle Migrantinnen und Migranten, die nur eine vorübergehende Migration anstreben oder beabsichtigen.

Kapitel 6 verwendet nun zwei multinationale Umfragen, eine unter Sprachkursteilnehmenden in 14 Ländern (siehe auch Kapitel 5) und eine unter Universitätsstudierenden in sechs Ländern, mit dem Ziel, die beiden oben erwähnten übersehenen Gruppen beobachten zu können. Da wir sowohl die Aspirationen als auch die Absichten aller Befragten beobachten, können wir diese beiden Größen vergleichen und gemeinsam analysieren. Darüber hinaus sind wir in der Lage, zwischen temporären und permanenten Migrationsabsichten zu unterscheiden und mögliche Unterschiede bei den Determinanten zu untersuchen. Wir tragen zu einem vollständigeren Bild potenzieller künftiger Migrantinnen und Migranten und zu einem besseren Verständnis der individuellen Faktoren bei. Im Vergleich zu Studien, die sich auf die GWP-Daten stützen, sind wir die ersten, die sich mit diesen Gruppen potenzieller Migrantinnen und Migranten befassen.

Wir schätzen die Determinanten auf individueller Ebene für die Aspirationen und Absichten, vorübergehend oder dauerhaft auszuwandern, mittels multinomialer Probit-Regression und stellen fest, dass die Aspirationen eine wichtige Determinante für die Migrationsabsichten sind. Aber selbst nachdem für Aspirationen kontrolliert wird, bleiben individuelle Merkmale für die Erklärung von Migrationsabsichten wichtig. Die Unterschiede zwischen Aspirationen und Absichten, dauerhaft auszuwandern, werden in erster Linie durch familiäre Bindungen im Herkunftsland erklärt. Familiäre Bindungen im Ausland hingegen können dazu führen, dass trotz fehlender Aspirationen die Absicht besteht, dauerhaft auszuwandern. Die Absicht, vorübergehend auszuwandern, wird in erster Linie durch Bildungsvorhaben erklärt. Die Relevanz der Determinanten ist jedoch je nach Geschlecht unterschiedlich. Der positive Zusammenhang zwischen familiären Bindungen im Ausland und der Absicht, dauerhaft auszuwandern, wird in beiden Stichproben hauptsächlich von Frauen hergestellt. Der positive Zusammenhang zwischen einem Hochschulabschluss und der Absicht, vorübergehend auszuwandern, ist nur für Männer in beiden Stichproben stabil.

Chapter 1

General Introduction

1.1 Introductory Remarks

Language learning is a fundamental skill taught in schools all across the World. In the European Union[1], e.g. almost $93\,\%$ of the pupils in 2019 learn a foreign language already in primary education (Eurostat, 2021). Language learning in schools is usually compulsory, it is likely to come with few or no costs and the acquired skills might affect later decisions in live, e.g. migration decisions. On the contrary, many adults decide to acquire foreign language skills voluntarily. In that case the direction of causality can be the opposite to that with language learning during childhood or adolescence. The determinants of foreign language skills and their benefits can be approached from two perspectives, the aggregated macro-perspective and the individual micro-perspective. While Part I focusses on adult language learning on the macro-level, Part II uses a micro perspective on language learning. Part III remains on the individual perspective, but concentrates more generally on migration intentions and aspirations, very important aspects of adult language learning.

On the macro-level, determinants and benefits focus mainly on migration and trade flows. Ginsburgh et al. (2017) show that the scope of learning depends negatively on the size of the world population of the native language, while literacy increases the scope of foreign language skills. The selection of a particular language is influenced by the costs of learning the foreign language measured by the linguistic distance between the native and the foreign language, and by

---

[1]   EU-27 without the United Kingdom.

trade with countries speaking the foreign language. Uebelmesser et al. (2021) show for the German language, that migration is positively associated with learning in the respective home country for EU member countries, while for third-country nationals the relationship holds for language learning in the destination country Germany. When looking at benefits, in particular the positive relationship between languages of the trading partners and trade flows, it is mainly based on common official or spoken languages of the trading partners (Egger and Lassmann, 2012; Head and Mayer, 2014) or their respective linguistic proximity (Isphording and Otten, 2013b; Melitz and Toubal, 2014). Only a few exceptions (Egger and Toubal, 2016; Fidrmuc and Fidrmuc, 2016) also show that acquired language skills increase trade flows. The same holds for the relationship between migration flows and language skills: While Adserà and Pytliková (2015) and Belot and Ederveen (2012) show that there is a negative relationship between the linguistic distance of origin and host country language and migration flows, there is only one study on the positive effects of foreign language skills measured by compulsory language learning in schools on migration flows (Aparicio Fenoll and Kuehn, 2016). One of the reasons for the lack of studies is the problem of measuring foreign language skills on the aggregate level which results in poor data availability.

This dissertation contributes to these strands of literature by focussing on the German language, migration flows to Germany and German exports. Germany is one of the world's largest exporting countries. At the same it is under pressure to counteract demographic change and the associated consequences for the labor market and society. Migration is the main solution for this problem. With the cultural institute Goethe-Institut (GI), Germany has established an important supplier of language courses worldwide that allows to analyse language learning for a large amount of countries and many decades. Chapter 2 presents the datasets on German language learning worldwide at the GI and shows developments of German language learning since 1965. Chapter 3 provides evidence on the effects of acquired language skills on migration flows by using a dataset on the presence of language-learning opportunities of the GI presented in Chapter 2, and Chapter 4 investigates the relationship of language-learning opportunities and German exports similar to Chapter 3.

On the micro-level, the benefits of foreign language skills are usually investigated in two contexts: host country language skills of immigrants and foreign language skills in a domestic labour market. First, proficiency in the host country's language improves labour market and integration outcomes. While it increases earnings (see e.g. Chiswick and Miller, 1995; Dustmann and Van

Soest, 2001) and the employment probability (Dustmann and Fabbri, 2003) on the labour market, it also improves social integration e.g. by increasing the probability of intermarriage and decreasing the likelihood of living in an ethnic enclave decreases (Bleakley and Chin, 2010). Second, effects of language skills foreign to the labour market's main language is not that clear. For the US, Saiz and Zoido (2005) have found no or only very small returns to foreign language skills. In the labour market of some European countries, however, there are high returns of foreign language skills for immigrants (Isphording and Otten, 2013b; Toomet, 2011) and natives (Ginsburgh and Prieto Rodriguez, 2011), in particular where the skills are relatively scarce (Ginsburgh and Prieto Rodriguez, 2011). Because of all these benefits on the foreign and domestic labour market, foreign language skills can be seen as human capital. The human capital theory might potentially explain why individuals decide to acquire foreign language skills, but literature on the reasons for learning a foreign language is missing. Closest to this question is the literature that focuses on determinants of foreign language skills of immigrants which are subsumed under the three E's: exposure, efficiency, and economic incentives (Chiswick and Miller, 2015). While the latter refers to the human capital investment motive of language learning, other reasons are widely ignored. With Chapter 5, this dissertation contributes to this question by distinguishing between motives that have human capital investment in mind and motives that refer to language skills as a consumption good. Education in general and language learning, in particular, might have a consumption value that generates direct utility (Kodde and Ritzen, 1984; Lazear, 1977; Schaafsma, 1976). Chapter 5 gives an overview of motives among language course participants and its determinants by using survey data from GIs worldwide.

This thesis is embedded in the context of potential migrants in their origin country and of language learning at the GI worldwide. Chapters 2 to 5 focus mainly on language learning that takes place in the countries of origin before an actual migration has taken place. When researching migrants already in their host country (e.g. Abramitzky et al., 2012; Beine et al., 2011; Borjas, 1987; McKenzie and Rapoport, 2010) a selection already took place between those who actually migrated and those who did not. Using survey data in countries of origin helps to overcome this drawback and allows to identify potential migrants without the selection. In this case, however, a new question arises how to measure this potential migration: literature is divided between using migration aspirations that express a desire (e.g. Bertoli and Ruyssen, 2018; Bertoli et al., 2020; Docquier et al., 2020; Ruyssen and Salomone, 2018)

3

and intentions that include concrete actions (e.g. De Jong, 2000; Friebel et al., 2013; Uebelmesser, 2006; van Dalen and Henkens, 2012). In this context, many studies use the Gallup World Poll (GWP) for an analysis of representative populations in around 160 countries. These data have two important limitations: First, it creates a conditionality by asking questions on intentions only if the respondents indicate aspirations to migrate. Second, it focuses on permanent migration and ignores potential migrants who only aspire or intend to migrate for a predefined time-span. Chapter 6 analyses the determinants of migration aspirations and intentions jointly with a particular focus on the two groups overlooked by the limitations in the GWP by using two surveys, one among language course participants at the GI and one among university students worldwide.

The remainder of Chapter 1 gives a summary of the results and their contribution to the literature. The rest of the thesis is structured in three parts. Part I consists of three chapters: In Chapter 2, we first give an overview of how language learning at the GI institute has evolved since the 1960s based on three datasets that measure the availability of language-learning opportunities with the presence of the GI in origin countries, the quantity of language learning in the institutes abroad and institutes in Germany. In Chapter 3, we use the dataset on the presence of language institutes to investigate the relationship between language-learning opportunities and migration flows to Germany. In Chapter 4, we use a similar approach for language-learning opportunities and German exports. Part II consists of Chapter 5, where we change the perspective to the micro-level and use a survey conducted among language course participants at the GI worldwide to answer what explains the motives behind the decisions to learn a foreign language. Part III consists of Chapter 6, where we use the same survey among language course participants at the GI worldwide in addition to a survey among university students and analyse migration aspirations and intentions. Chapter 7 concludes with giving an overview on the policy implications of our finding, showing limitations of the presented research and presenting possible ideas for future research.

1.2 Contributions to the literature and summary of results

1.2.1 German language skills worldwide

New datasets on language learning

While there are many studies on the benefits of language proficiency on the individual level (see e.g. Bleakley and Chin, 2010; Dustmann and Fabbri, 2003; Dustmann and Van Soest, 2001), evidence on the aggregate level is scarce with only a few exceptions. One of the reasons is the problem of measuring foreign language skills on the aggregate level as language learning and foreign language skills are heterogeneous in many aspects ranging from the level of proficiency, the motives behind learning a language (see Chapter 5) and the context where language skills are acquired – mainly within or outside the school system. Egger and Toubal (2016) and Melitz and Toubal (2014) consider the effect of acquired language skills on bilateral trade flows, but the data on foreign language skills do not vary across time. Aparicio Fenoll and Kuehn (2016) use policy reforms on compulsory language learning at school as a measure for language skills in the context of migration flows within the EU. While their data vary across time, they only focus on the context of language learning at schools.

Chapter 2 contributes to overcoming the lack of data by presenting data on German language acquisition outside the school system for five decades. Data is collected from the annual reports of the GI. The GI is an important actor in Germany's foreign cultural policy. Its main duty is the promotion of German culture and language worldwide by offering language courses and information on German culture and society (Auswärtiges Amt and Goethe-Institut, 2004). The datasets are the first that allow understanding the development of German language acquisition worldwide for many decades.

The three datasets presented in Chapter 2 are based on the annual reports of the GI. The GI reports for each institute data about language courses and exam participation. The first dataset contains information about the supply side of language-learning opportunities, the presence of the GI on the city level for the period from 1965 to 2014. For each city-year combination, the dataset provides data about the types of institutes, their offer of language services as well as their opening and closing years. These data are used in Chapter 3 and 4 and have been applied by Jaschke and Keita (2021) to show that the presence of GI

affects positively language skills and relevant labour-market characteristics of immigrants at arrival. The other two datasets focus on the demand for language services. The second dataset contains information about language courses and exam participation at institutes abroad. From 1986 to 2014, the GI reported for each institute the number of participants in standardized exams. Furthermore, there is information on course participation by two variables, the number of registrations from 1990 to 2014 and sold course units from 1972 to 1989 and 1997 to 2014. The third dataset provides information about the number of registrations for language courses in Germany by nationality from 1966 to 2014. Uebelmesser et al. (2021) used the data on exam participation abroad and registrations in Germany to identify determinants of language acquisition.

We find patterns in the developments of German language learning worldwide that highlight a heterogeneity across time, across regions, and even across demand-side and supply-side. When looking e.g. at the supply of language-learning opportunities, the observation period from 1965 to 2014 can be divided into three phases. Before the 1980s, the GI was mostly present in America and the "old Europe". After the end of the Cold war, the GI expanded to former socialist countries while decreasing the number of existing institutes. Since the early 2000s, the GI expanded to new countries in Africa while holding the already existing presence quite constant. These data on the presence of institutes is used in the following to first analyse their effect on migration and then on their effect on German exports.

Language learning opportunities and migration

Potential migrants might consider the benefits of language proficiency for integration into the labour market and society (see e.g. Bleakley and Chin, 2010; Dustmann and Fabbri, 2003; Dustmann and Van Soest, 2001) when forming migration decisions and location choices. And this is indeed, what many studies show based on linguistic distance between the respective languages of host and origin countries (Adserà and Pytliková, 2015). While linguistic distance is a proxy for the cost of language acquisition, it neglects the actual language acquisition of potential migrants before migration.

Theoretically, individuals decide on staying in their country of origin or migrating to another destination based on their maximized expected utility by comparing their relative attractiveness based on a cost-benefit analysis. The costs on the country level can be geographical and linguistic distance, or visa regula-

tions. On an individual level, learning the language of the potential destination country increases expected earnings there, but also raises the costs of migration. Hence, individuals will decide to acquire language skills if their utility is maximized by doing so. On an aggregated level, increasing language-learning opportunities might decrease costs of language acquisition and therefore have a positive effect on migration flows.

Language learning can be distinguished between language learning as a child and as an adult. This is different in terms of costs, but also in the direction of causality in the context of migration. During childhood or adolescence, the decision to acquire language skills is more likely determined by factors outside the learners' direct control like the school system. Often, it comes with few or no costs but might still affect later migration decisions. On the contrary, language learning during adulthood is more likely based on the individual motive of the learners themselves where migration decisions might result in language learning. The direction of causality can therefore be opposite to the direction of language learning during childhood or adolescence.

While Aparicio Fenoll and Kuehn (2016) find a positive relationship between compulsory language learning at school and migration flows within the European Union, to the best of our knowledge there is no evidence on the effect of adult language learning on migration. Hence, Chapter 3 contributes to closing this gap in the literature and studies the effect of the presence of language-learning opportunities for adults on migration. We use the data on the presence of the GI as presented in Chapter 2 and form a unique panel dataset for 69 countries for the period 1977 to 2014. Starting the empirical analysis with fixed-effects (FE) estimations, we find that the number of language institutes in a country is positively correlated with migration from that country to Germany. By identifying institutes that offer language services, we can show that correlation is driven by language courses.

In order to interpret the results causally, we need to address the issue of reverse causality. When looking at the strategies and overall objectives of the GI during our observation period, it gets clear that migration to Germany does not play any particular role in the location decisions of institutes. Nevertheless, we also estimate the effect of language institutes on migration flows to Switzerland since migration to Switzerland can be considered exogenous to the number of institutes. We find similar results for migration to German-speaking Switzerland and therefore interpret the results as causal. Moreover, estimates using the Common-Correlated-Effects (CCE) estimator (Pesaran, 2006) show

that half of the total effect is a direct effect on migration to Germany by individuals who would not have migrated otherwise. The other half is an indirect effect of redirected migration flows to Germany. The CCE estimator can be used to control for multilateral resistance to migration (Bertoli and Fernández-Huertas Moraga, 2013), i.e., migration flows no longer depend only on origin and destination country characteristics, as in FE estimates, but also on the attractiveness of alternative destinations.

While language-learning opportunities show a positive effect on migration flows, we use a similar approach in the following to analyse their effect on exports.

Language learning opportunities and exports

Language skills reduce transaction costs but are far from uniform across goods and trading partners. They are particularly high for economic activities and exchanges that require extensive written and verbal communication. This view is supported by studies on the relationship of linguistic properties of the trading parters' languages (Egger and Lassmann, 2012; Head and Mayer, 2014) and acquired language skills (Egger and Toubal, 2016; Fidrmuc and Fidrmuc, 2016) on trade flows. Hence they show that language and language acquisition matter for trade flows. Policies that increase domestic language skills abroad hence offer a potential tool for export promotion. In Chapter 4, we test if language-learning opportunities have an effect on exports by using data on the presence of GI as presented in Chapter 2.

Starting the empirical analysis with fixed-effects (FE) estimations, we find that the number of language institutes in a country is positively correlated with German exports to that country. This correlation is non-linear in the number of institutes with a positive but marginally decreasing impact of the institutes. These results confirm Ghosh et al. (2017), Lien and Lo (2017), and Lien et al. (2012, 2019) who show a positive relationship between the establishment of different cultural institutes (e.g. GI for Germany, Cervantes Institute for Spain) and trade flows. While they see cultural institutes as a way of increasing soft power (Nye, 2004), we investigate the language channel which reduces transaction cost. The data allow distinguishing between institutes that offer language services, and those that do not and show that only language institutes are significantly correlated with exports. Additionally, we use Rauch's (1999) product categories to distinguish export flows according to their different need for communication and therefore different search costs when trading. On the

one hand, we find that there is no significant association between the number of institutes and the export of homogeneous goods for which only little communication is needed. On the other hand, the export of differentiated goods which requires more communication when traded increases with the number of institutes. Hence, we conclude that language is the relevant channel.

Furthermore, we address the issue of reverse causality and deal with this problem by estimating the effect of German institutes on Swiss exports. The idea and implementation are the same as in Chapter 3: The German government decides jointly with the GI where to establish institutes, but not jointly with any Swiss institution. Therefore, the decision to open an institute is exogenous to Swiss export flows. We provide evidence that institutes increase exports from the German-speaking part of Switzerland, but not from the non-German-speaking part of Switzerland to countries hosting these institutes.

1.2.2 Language learning: human capital investment or consumption

The effects of language-learning opportunities on migration flows and exports outlined in Part I have their micro-foundation in the productive value of foreign language skills in two different contexts. First, one can think of the language skills of natives and immigrants, which are foreign relative to the main language of the country of residence (see e.g. Ginsburgh and Prieto Rodriguez, 2011; Stöhr, 2015). Second, skills in the foreign language of the host county have a positive effect on the integration into the labour market of the host country by increasing earning (see e.g Dustmann and Van Soest, 2001) and employment probabilities (see e.g Dustmann and Fabbri, 2003) and by improving occupational choices (Aldashev et al., 2009). These pieces of evidence can be viewed from a human capital theory (Becker, 1964; Schultz, 1960) perspective which claims that individuals choose the human capital that maximizes their expected net present value of income. As the human capital acquisition is costly it is expected to lead to monetary returns via increased wages or increased employment probabilities. It does so by increasing individual productivity. Sjaastad (1962) include migration into the human capital framework as expected returns which can be realized either in the domestic or the foreign labour market; in that context, it can be further extended to comprise foreign language skills as a specific type of human capital.

Empirically, individuals' choices of education cannot be fully explained by the human capital theory, as they often choose a higher amount of education or other types of education than would be optimal according to the human capital theory (Canton and Jong, 2005; Oosterbeek and Ophem, 2000). Ignoring irrational choices, expected benefits on the labour market seem not to be the only determinant. One possible explanation for this behavior is that learning has a consumption value and generates direct utility (Kodde and Ritzen, 1984; Lazear, 1977; Schaafsma, 1976). This consumption value can be defined as "the private, intended, non-pecuniary return to education" (Alstadsaeter, 2011). Individuals may then choose a different quantity or type of education which does not lead to optimal monetary returns (Arcidiacono, 2004).

Chapter 5 focusses on the decision to acquire a foreign language and the different motives behind that choice. The motives are based on learning as human capital investment or as consumption. We concentrate on language learning outside the school system in which participants actively decide to acquire language skills. To address the question, why people decide to acquire language skills, we use survey data collected among language course participants at the GI in 14 countries worldwide. To the best of our knowledge, we are the first to apply the human capital framework enlarged by consumption motives to language learning in order to contribute to a better understanding of language acquisition from a cross-country perspective.

We proceed in three steps: First, we present evidence on a measure based on consumption and investment motives. This measure builds upon the main reason to study German indicated by the participants which we then aggregate into the human capital investment motive and the consumption motive. Descriptive evidence identifies strong heterogeneities in motives with respect to countries and age groups. Binary probit estimations identify the individual and country determinants of the human capital investment motive. On the country level, we find that living in a European country, in a country with a Germanic official language, or in an at least upper-middle-income country is associated with a lower probability of having an investment motive of learning German. On the individual level, we identify strong heterogeneities in determinants among country groups, age, and gender.

Second, we present evidence on a second measure – the expected probability of using the foreign language skills in a professional environment –. We can show that a human capital investment motive does not necessarily match with a high probability of professional use. We apply again binary probit estimations to

study individual and country-level determinants of professional use. On the country level, non-European and lower-middle-income countries show now a negative correlation, while it was positive in the specification for the human capital motive. While on average, respondents from these two country groups are more likely to indicate an investment motive, they, at the same time, assess the probability of professional use to be lower than respondents from the respective reference groups. On the individual level, we identify again heterogeneities in a subgroup analysis based on age, gender, and country groups.

Third, by looking at language learning by motives, we are interested in possible spill-overs from a given consumption motive to a professional use and find that it mostly emerges in a "tied-mover" context. This group mostly comprises younger women with a native German partner who indicates migration for reasons that are not non-educational or labour market-related reasons.

1.2.3 Migration aspirations and intentions

As outlined above, potential migration is a very important aspect in the decision to learn a foreign language in the country of origin. Much of the economic migration literature, however, focuses on migrants already in their destination countries (see e.g. Abramitzky et al., 2012; Borjas, 1987; McKenzie and Rapoport, 2010). The shortcoming of this literature is that migrants are a selection of those who actually migrated. When relying on survey data to study individuals in their origin country before actual migration takes place, this drawback can be overcome as potential future migrants can be identified. Potential migration is either measured by migration aspirations, which express the desire to migrate under ideal circumstances, or by migration intentions, which refer to considerations and plans based on the actual situation.

The data of the Gallup World Poll (GWP), a repeated cross-sectional and representative survey covering around 160 countries, are often used in this context. The GWP data come, however, with two important limitations: First, the question on migration aspirations filters questions on migration considerations and plans, such that these questions are only answered if the respondents indicate aspirations. This filter creates a conditionality in responses and assumes that aspirations to migrate are necessary for migration intentions. But intentions to migrate do not necessarily imply aspirations and these cases are overlooked by the GWP. Second, questions on aspirations and intentions are limited to

a permanent time horizon and therefore neglect potential migrants who only aspire or intend to migrate temporarily.

The aim of Chapter 6 is to address these two limitations and to provide a comprehensive analysis of migration aspirations and intentions, with a particular focus on those potential future migrants who have often been overlooked, those with migration intentions without aspirations, and those potential migrants who have a temporary time-frame in mind. We use two multinational surveys: one among language course participants in 14 countries that is used in Chapter 5 as well and one among university students in six countries. These surveys observe migration aspirations intentions for all respondents, such that we can observe the two above-mentioned overlooked groups.

By observing both aspirations and intentions for all respondents, we can compare these two measures and analyse them jointly. Furthermore, we are able to distinguish between intentions to migrate temporarily and permanently and analyse possible differences in their determinants. By responding to both questions on migration intentions and aspirations, we further observe those individuals who do not state aspirations to migrate permanently under ideal circumstances but have intentions to do so. We contribute to a more complete picture of potential future migrants and to a better understanding of the individual-level factors which potentially lead to intentions to move abroad despite having no aspirations to do so. In comparison to studies relying on the GWP data, we are the first to give attention to these groups of potential migrants.

We estimate individual-level determinants of aspirations and intentions to migrate, temporarily or permanently, via multinomial probit regression. We run our estimations for both datasets separately and find similar results. We find that the aspirations are an important determinant of intentions, but even after controlling for aspirations, individual-level characteristics remain important in explaining intentions to migrate, both temporarily and permanently. Differences between aspirations and intentions to migrate permanently are primarily explained by family ties at the origin which can prevent aspirations to migrate permanently from being translated into intentions. Family ties abroad, on the contrary, can lead to intentions to migrate permanently despite having no aspirations. Intentions to migrate temporarily are primarily explained by educational plans. However, the relevance of determinants differs by gender. The positive linkage between family ties abroad and intentions to migrate permanently is driven largely by women. The positive linkage between a university degree and intentions to migrate temporarily is robust only for men.

Part I

German language skills worldwide

Chapter 2

New datasets on language learning

Abstract.[2] This chapter presents a comprehensive overview of German language learning for more than 100 countries (including Germany) over a period of 50 years. We provide new and unique data from the Goethe Institut, a German cultural institute, which offers language courses and standardised exams. These data contain information about the supply of language-learning opportunities, i.e. the number and the geographic distribution of institutes, and the demand in the form of course and exam registrations. These data do not only show the development of language learning for the German language over time, they also underline common trends and heterogeneities across regions.

14

2.1 Introduction

Proficiency in a language different from one's mother tongue can be due to different motives and related to different benefits. Language learning can be an investment good (with language learning enhancing individual human capital) or a consumption good.[3] The latter can be related to personal motives, which comprise cultural interests, like the interest in the respective culture or a general interest in language learning, and motives related to the social environment, like learning the partner's or a friend's mother tongue. Language learning as an investment in human capital can refer to the domestic labour market if the foreign language is important vis-a-vis trading partners or customers, eases the communication within the firm or more generally leads to a wage premium. It can also be related to better perspectives on a foreign labour market and thus be linked to migration.

While empirical evidence on the importance of the consumption motive is mostly missing, there is a literature about the effects on labour market outcomes. Empirical evidence related to the domestic labour market, however, is mixed. While in the US, foreign language proficiency does not contribute to higher wages (Fry and Lowell, 2003), especially English proficiency in non-English speaking countries generates wage premia (Ginsburgh and Prieto Rodriguez, 2011; Lang and Siniver, 2009; Toomet, 2011). On the contrary, Stöhr (2015) finds evidence that a return for other foreign languages exists in Germany, but only in occupations related to trade. Isphording and Otten (2013a) show significant wage premia for English, French and German for immigrants to Spain and relate it to the shortage of language proficiency in the Spanish work force.[4]

Contrary to the mixed evidence for the domestic market, there is evidence of positive effects of migrants' proficiency in the language of the destination country. Proficiency helps migrants to overcome barriers in social and economic integration and hence reduces costs of migration. Language proficiency improves labour market outcomes of migrants by increasing earnings (see e.g. Dustmann and Van Soest, 2001, 2002) and employment probability (Dustmann and Fabbri, 2003). Furthermore, proficient migrants are more successful in social integration into destination countries, as the probability of intermarriage

---

[3] See Becker (1964) and Lazear (1977).

[4] To the best of our knowledge, only Ginsburgh et al. (2017) provide evidence on bilateral trade flows as determinant of aggregated language learning.

increases, and they are less likely to live in an ethnic enclave which complements their economic integration (Bleakley and Chin, 2010).

These studies about the economic benefits of language proficiency have mostly relied on self-reported measures and have linked them to other individual-level socio-economic or work-related data. The process of language learning has so far been largely ignored. This chapter aims at closing this gap for the German language by presenting data about language learning for the time period 1965-2014. Data are from the annual reports of the Goethe-Institut (GI),[5] a German association which promotes the study of German language and culture all over the world.[6] The GI has institutes which, besides cultural events, offer language courses and widely recognised standardised exams.

Exceptions of papers with a focus on language learning are Egger and Toubal (2016) and Melitz and Toubal (2014) who consider the effect of common spoken languages and acquired language skills on bilateral trade flows. They use language data for one point in time, however, which does not allow studying variation across time. Furthermore, Aparicio Fenoll and Kuehn (2016) use compulsory language learning at school as a measure for language skills and relate it to migration flows within the European Union. While their data vary across time, acquired language skills likely go beyond compulsory language learning.

The small number of studies on language learning is mostly due to the lack of data. For the German language, for example, there are reports about language learning in schools, universities and at the GI which have been published every 5 years since the early 1990s by a group of institutions consisting of the German Academic Exchange Service (DAAD), the GI, the Foreign Federal Office and the Central Agency for German Schools Abroad (ZfA) (Auswärtiges Amt, 2015; Netzwerk Deutsch, 2010; Ständige Arbeitsgruppe Deutsch als Fremdsprache, 2003, 2006). While they provide qualitative insights, they are not suited for a quantitative analysis as data are available only for selective periods, often based on estimates from experts and subject to frequent revisions.

Instead of focusing on the process of language learning, the relation between language skills and outcomes has often been captured by measures of linguistic distance between individuals' mother tongues and the languages spoken in the

---

[5] From 2015 onward, the data in the annual reports are no longer presented in the same disaggregated way as before.

[6] In this chapter we stick to the following convention: when referring to the association of the Goethe-Institut we use the abbreviation "GI". When talking about a specific branch of the GI in Germany or abroad, we refer to it as "institute".

countries with which they want to interact. Related to migration as a potential motive, some studies control for common language as determinant of migration flows (Grogger and Hanson, 2011; Mayda, 2010; Ortega and Peri, 2013; Pedersen et al., 2008), while others include more accurate measures of linguistic distance. Belot and Ederveen (2012) for example use for measures based on the linguistic family tree, while Belot and Ederveen (2011) and Adserà and Pytliková (2015) estimate the effect of linguistic distance on international migration flows with two, more fine-grained indices developed by linguists: the Levenshtein Index based on phonetic similarity and the Dyen Index based on similarity of sample words in Indo-European languages. Chiswick and Miller (1998) measure linguistic distance by an approach based on language scores of native English speakers studying foreign languages which indicate difficulties of studying a particular language.

Similar considerations as for migration flows are applicable to international trade flows. The same measures for linguistic distance or a common language as in the context of migration have been used in empirical applications (Hutchinson, 2005; Isphording and Otten, 2013b; Lohmann, 2011; Melitz, 2008).

All these distance-based indicators reflect two aspects of languages. First, they serve as proxies for the costs of acquiring language skills: languages linguistically more distant from the mother tongue are more difficult to learn. Second, linguistic distance is in part an explanation of cultural differences (Belot and Ederveen, 2012) and is related to long-term cultural and historical connections between countries. Not too surprisingly, a negative relationship between linguistic distance and migration flows as well as trade has been established.

Linguistic distance is however not something invariant as often assumed, at least not on an individual level. By acquiring a foreign language, individuals can reduce the barriers resulting from that distance. The learning costs are likely related to linguistic distance, but so might also be the benefits.

The data presented here provide a detailed overview of language learning worldwide over the last five decades for the German language. Beside the number of institutes and their geographic distribution across countries, i.e. the supply of language-learning opportunities, the demand as documented by numbers of course and exam participation in the institutes abroad is presented. In addition, data related to the institutes located in Germany complement the picture. These data allow understanding the development of language learning across time with a focus on similarities and differences across regions.

The rest of the chapter is structured as follows: Chapter 2.2 describes the history and the goals of the GI in more detail. Chapter 2.3 provides information about the three constructed datasets. Chapter 2.4 presents a detailed picture about language learning worldwide over the last five decades with a special focus on regional differences. Chapter 2.5 adds the development of learning the German language in institutes in Germany. Chapter 2.6 concludes.

2.2 The Goethe-Institut

The GI is a German association that promotes German language and culture. It is one of the main actors of the foreign cultural policy and therefore closely related to the German government, in particular to the Federal Foreign Office (FFO). The GI was founded in 1951 as a successor of the "Deutsche Akademie" (German academy) to rebuild and renew the infrastructure for foreign cultural policy after the Second World War.

In 1965, the GI published an annual report for the first time. At that time, the number of institutes grew fast, which was also attributable to the integration of other cultural institutes of the Federal Foreign Office (FFO) into the GI. While in the beginning the GI concentrated on language teaching, in 1961 cultural promotion became a second goal. At the end of the 1960s, the GI got more closely connected to the FFO in order to carry out its foreign cultural policy. During that period, foreign cultural policy evolved as a third component of foreign policy beside diplomacy and foreign-trade policy (Singer, 2003).

In 1970, a new concept of foreign cultural policy was introduced which emphasised cultural cooperation and exchange, instead of cultural export. Initiating and supporting interactions between cultures became the main objective of foreign cultural policy. Within this concept the term "culture" was defined more broadly including science, education and media. The FFO acted as coordinator of different organisations which carried out different parts of foreign cultural policy, among them the GI (Auswärtiges Amt, 1970; Schneider and Schiller, 2000; Singer, 2003).

A framework treaty between the GI and the FFO, signed in 1976, states the relationship between the two institutions. The treaty ensures the content-related autonomy of the GI. The current version of the treaty from 2004 clarifies explicitly the main duties of the GI: First, to promote the German language by providing courses, training for teachers and scholarships for language learning.

Second, to support international cultural cooperation by involving cultural actors of the destination countries in cultural events and exchanges. Third, to inform about social, political and cultural life in Germany by maintaining libraries with German books and media and by organising cultural exchange with Germany (Goethe-Institute and Auswärtiges Amt, 2004). While the GI is mainly funded by the German government, language courses are financed by course fees (Goethe-Institut, 2014).

To fulfil these duties, the GI has institutes all over the world. In 2014, there were 149 institutes in 94 countries out of which 126 institutes offered language services (see Figure 2.1). In addition, there were 12 institutes in Germany. There is no official guideline for location choices of new institutes. According to GI officials, openings and closings of institutes take place in consultation between the FFO and the GI. The process of opening new institutes starts with suggestions for locations mostly by the GI, the FFO or members of the legislature. In a next step, the GI and the FFO discuss the locations on the basis of their general objectives taking into account legal, political and social aspects of the potential host country. Furthermore, aspects related to the overall security in the host country and global developments play a role, as well as considerations about the larger region. Very often location decisions can be seen as reactions to changes in the political, social or economic situation.

Over the last 50 years, the number of countries with at least one institute has continuously increased, while the number of institutes has fluctuated between 130 and 160 for most of the time. Figure 2.2 illustrates these developments for the period 1965 to 2014 and shows that language services are widely offered over the whole period.[7]

There might be self-selection of participants based on the following three characteristics: 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 the costs of other equally suitable learning options. Courses could be more expensive if one considers the GI as a premium provider of language courses, because it is a semi-official German organization with a long tradition and a good reputation. Courses could also be less expensive, because the majority of the GI's funds comes from the German government. Counter-arguments

---

[7] There are three different types of institutes: main institute, subsidiary ("Nebenstelle" or "Außenstelle"), which is linked to a main institute, or liaison office ("Verbindungsbüro"), which is not linked to a main institute. Most of the institutes are main institutes. These main institutes can have one or more subsidiaries, which are directly linked to the institute. Figure 2.2 contains all types of institutes.
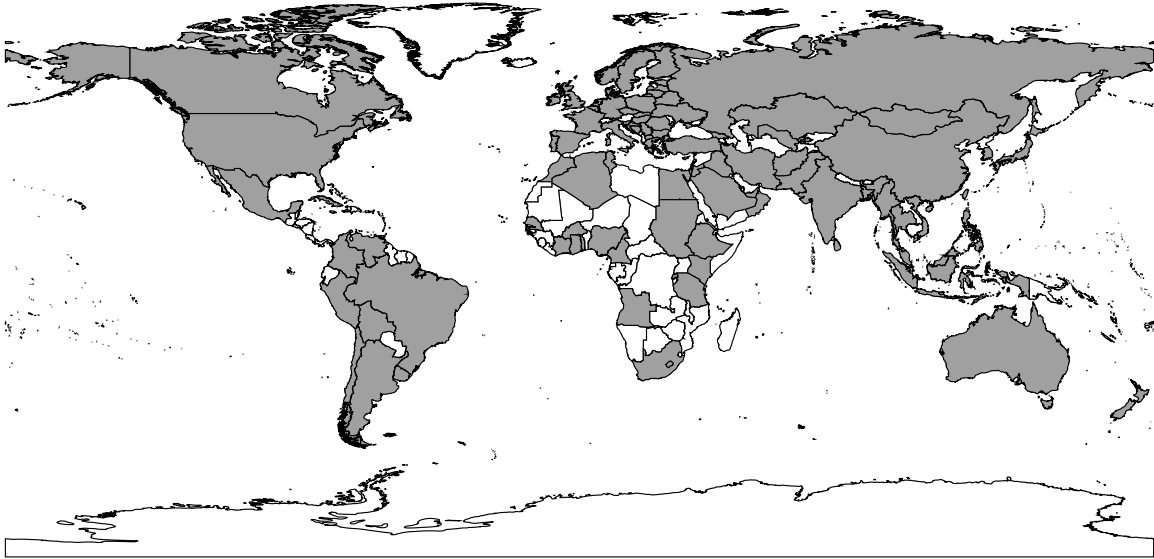
Figure 2.1: The distribution of the GI in 2014.
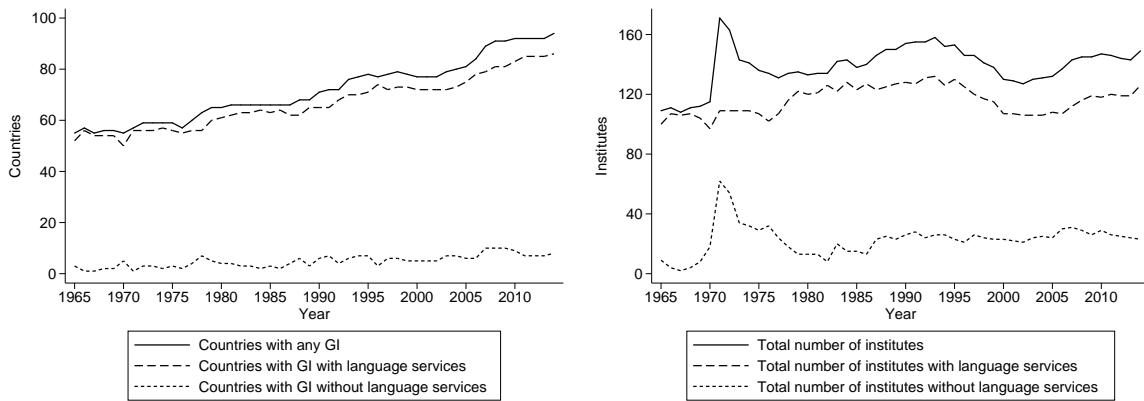
(a) Countries with GI

(b) Institutes

Figure 2.2: Number of countries and institutes.

for both lines of reasoning can, however, be found. Historical price data on language courses are not available. If one looks at current prices instead, no clear pattern emerges. In particular, 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, i.e. not financed by government funding. It can therefore not be expected that the courses of the GI are cheaper than comparable courses offered by other providers.

Institutes are usually located in capitals and other major cities. The lack of institutes in rural areas is likely to lead to an under-representation of language learners from these areas among participants at the GI. However, the bias need not be as large as one would initially expect: the GI offer both extensive and intensive language courses. Extensive courses are based on weekly lessons and last for several months, but intensive courses are taught en-block. Participants of intensive courses do not necessarily have to live in the vicinity of the respective institute. They may also stay there for the duration of the course only. This holds in an analogous way for participation in exams.

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 other language-learning opportunities, including universities, private language schools, and internet platforms. This multitude of alternative learning opportunities might give rise to additional concerns regarding the self-selection of language learners into courses offered by GI and possible changes across time. The language courses taught by the GI are mainly offered as traditional "offline" forms of language learning. Since 2010, this has been complemented however by online and blended-learning courses, which combine traditional and online learning (Goethe-Institut, 2011). At the other end of the spectrum are institutions which only offer online courses. The latter kind of courses may be more attractive to a younger generation of language students, which is more familiar with using the internet in general. While this difference may lead to an over-representation of older participants among the participants in language courses at the GI, the advent of online language learning platforms in the late 2000s falls in the very last years of our data. Given the nature of our macro-level data, we can only speculate about potential changes in the composition of participants over time and regions and, more generally, in the role of the GI.[8] The data allow, however, for a detailed picture of the overall development over a long time period and for a large number of countries.

---

[8] For that reason, micro-level surveys in a selected number of institutes are planned.

2.3 Datasets

Since 1965, the GI has continuously published annual reports in which activities of each institute including data about language course and exam participation are reported. These annual reports are publicly available. We digitised this information in order to construct three datasets.[9]

While the first one focuses on the presence and type of activities of the GI in each country, the two other datasets contain data about the extent of language services abroad and in Germany. The following provides some more details in terms of years and countries as well as content of the three datasets: The first dataset contains information about the presence of the GI on the city level for the period from 1965 to 2014. For each city-year combination the dataset provides data about the types of institutes, their offer of language services as well as their opening and closing years. Over the analysed 50 years, the GI has been present in 272 cities in 109 countries.

The second dataset contains information about language course and exam participation at institutes abroad. From the annual reports of the GI, we construct three variables for different time periods. First, from 1986 to 2014 the GI reported for each institute and year the number of participants in standardised exams ("zentrale Prüfungen") which are widely recognised, e.g. for language requirements in universities. These exams can be categorised in exams for children, adolescents and adults, and for professional use, and they are further differentiated by the level of language skills. The exams can be taken by course participants, but this is not obligatory; they can also be taken by individuals who have not attended a language course at the GI. Only in very few years numbers were reported separately for each type of exam, while for most years only aggregate information was available. Hence, we only use aggregate numbers for exam participation per year and institute.

Furthermore, there is information on course participation by two variables per institute and year: the number of registrations in language courses for the period 1990 to 2014 and an indicator, which we construct, for sold course units for the periods 1972 to 1989 and 1997 to 2014.

As to registrations, courses are organised in course periods, mostly two periods (semesters) or three periods (trimesters) per year. While from 2000 onward

---

[9]   For the data and technical information about the construction of the datasets, see Uebelmesser et al. (2018b,c,d). See Uebelmesser et al. (2021), Chapter 3 and 4 for first applications of the datasets.

the annual number of registrations are reported in the annual reports, in the years 1990 to 1999 only average numbers of students per course period were published. For this period, we construct the number of registrations per year by multiplying the average number of students per course period with the number of course periods.

Concerning sold course units, information was reported in the year 2006 and from 2009 onward. For the years 1972 to 1989 and the remainder of the years between 1997 and 2014, we construct the measure for sold course units as follows: $sold\,course\,units = total\,number\,of\,lecture\,units * average\,course\,size$, where the number of lecture units[10] is the sum of the units all teachers taught at an institute within a year and the average course size is the number of students divided by the number of courses.

Total course and exam participation at institutes worldwide reached their maximum for all three variables in 2014 with 287,630 exams, 229,702 registrations and 17,113,040 sold course units.

The third dataset provides detailed information about the annual number of registrations in language courses in Germany by nationality.[11] Data are available for the period 1966 to 2014. In total, course participants with around 200 different nationalities studied the German language at the GI in Germany over the whole time span.

2.4 Language learning across time – regions abroad

The almost continuous expansion of the presence of the GI over the last five decades as shown in Figure 2.2 hides important heterogeneities across time and across regions. Furthermore, not only the presence of institutes, i.e. the supply of language-learning opportunities, captures the spread of the German language worldwide; also the demand as documented in the numbers of registration and exam participation is an important indicator. In the following, the regional distribution of institutes and data about registrations in these institutes are presented including their evolution over time.

---

[10] A lecture unit has 45 minutes.
[11] E.g. in 2014, there were 12 institutes in Germany: Berlin, Bonn, Bremen, Dresden, Düsseldorf, Frankfurt/Main, Freiburg, Göttingen, Hamburg, Mannheim/Heidelberg, Munich and Schwäbisch Hall (Goethe-Institut, 2015).

2.4.1 The worldwide presence of the GI

The expansion of the GI can be described more precisely if we look at the evolution across regions in Figure 2.3, which shows the number of countries with institutes and the number of institutes separately for different regions for the period 1965 to 2014.[12] The assignment of the countries to regions according to the regional organisation by the GI since 2008 can be found in Table 2.A1 in the Appendix.

Europe (incl. Central Asia) is divided into five regions, which follow three different patterns (see Figures 2.3(a), (b)). First, Southwest and Northwest Europe as the "old Europe" had a relatively high number of institutes before 1990. This number decreased in the 1990s, especially in Northwest Europe. The number in Southwest Europe decreased only slightly before increasing to the old level again. The number of countries in these two regions was fairly stable throughout the whole period. Second, in Southeast Europe the number of institutes was quite stable during the whole period, while the number of countries began to increase in 1990. The latter reflects the expansion of the GI to (former) socialist countries, like Bulgaria, Romania and the successor states of Yugoslavia. At the same time, some institutes in this region were closed, in particular in Greece. Third, Central Eastern Europe, and Eastern Europe and Central Asia reflect even more the expansion of the GI around and after the fall of the iron curtain. While before the end of the 1980s (almost) no institutes were located in these regions, the number of institutes and countries has increased sharply since then.

In Asia (without Central Asia and Middle East), there are three regions (see Figures 2.3(c),(d)): East Asia, Southeast Asia (with New Zealand and Australia) and South Asia. The numbers of institutes and countries in East Asia have slightly increased over the whole period, especially after 2000. In South East Asia, the relatively high number of institutes was reduced slightly in the 1990s, while the number of countries was more or less stable. While at the beginning relatively many institutes were located in South Asia, this number decreased at the end of the 1980s, and so did the number of countries.

---

[12] This number includes all main institutes and subsidiaries or liaison offices, when there are no main institutes in that country (see Footnote 7 for more information on different types of institutes). E.g. Skopje, Macedonia, is first a subsidiary affiliated to Athens, Greece, and later a liaison office. However, there is no other GI in Macedonia. Other cases are Kathmandu in Nepal as a subsidiary of New Delhi in India before 1981 or Havana as a liaison office in Cuba.

In North Africa (including Middle East) and Sub-Saharan Africa (see Figures 2.3(e),(f)), the numbers of institutes and of countries almost coincide, which means, that the GI has no more than one institute in most countries. This number was quite stable until around 2005 when the GI started to expand in both regions.

Finally, America is divided into two regions (see Figures 2.3(g),(h)), North/ Central America, and South America. South America was a core region of the GI until the beginning of the 1990s with an average of at least 2 institutes per country. While there has been no change in the number of countries with institutes after 1966, the number of institutes has changed: it increased until 1980 and stayed constant until the mid of the 1990s. Afterwards, the GI began to close many institutes there. In North America, additionally to Canada, Mexico and the US, the GI has been only present in Costa Rica and Cuba for some years, but not in the remainder of Central America. The number of institutes in North America increased sharply after 1975. At the end of the 1990s, many institutes were closed.

Overall, the years from 1965 to 2014 can be divided roughly into three periods. Until the end of the 1980s, the focus of the GI was mainly on America and the "old Europe", but also on Asia and Africa to a certain extent. With the change of the political landscape in the 1990s the GI expanded to former socialist countries, while simultaneously decreasing its large number of institutes in America and the "old Europe". From the early 2000s onward, the presence of the GI has been quite stable, only in Africa the GI has expanded to new countries.

It is possible to interpret this development in the light of changes to the public funds by the FFO (Schneider and Schiller, 2000). In the beginning, public funds allowed the expansion of the GI. Subsequently, financial cuts and political realignment of foreign cultural policy led to closings (Singer, 2003). This affected the GI mainly in the early 1990s when the GI opened many new institutes in Middle and Eastern Europe and institutes in other regions had to close because of that expansion. At the end of the 1990s, further financial cuts led to more closings of institutes (cf. Figure 2.2).
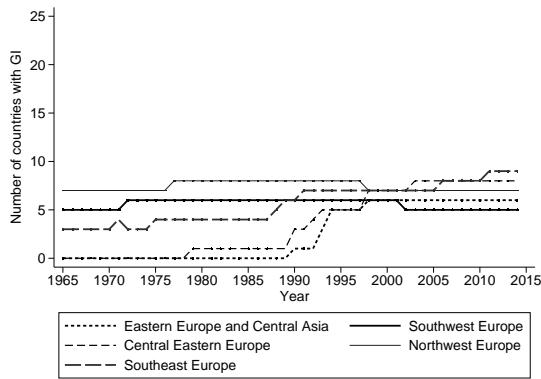
(a) Countries with institutes in Europe
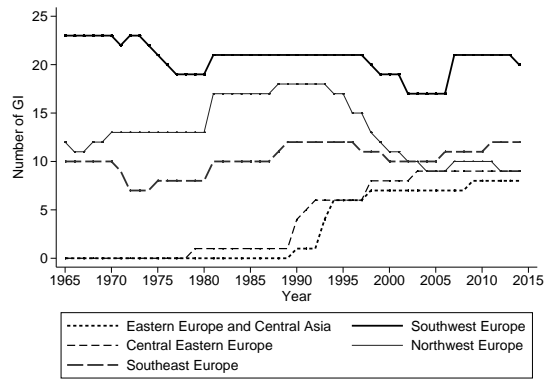
(b) Institutes in Europe
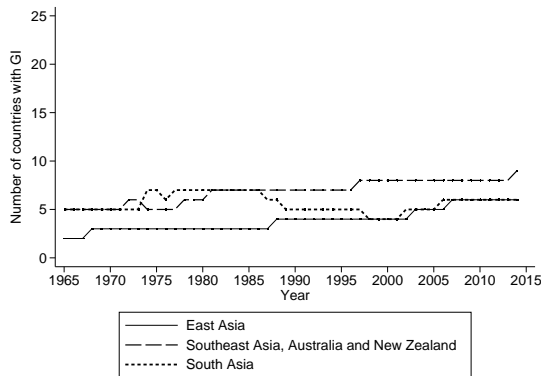
(c) Countries with institutes in Asia (without Central Asia and Middle East)
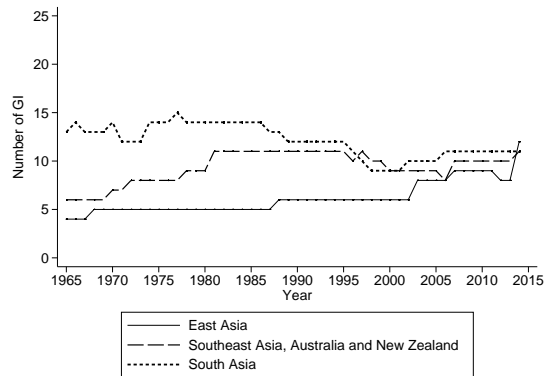
(d) Institutes in Asia (without Central Asia and Middle East)
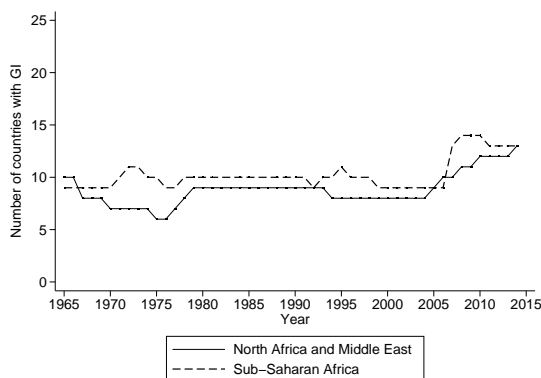
(e) Countries with institutes in Africa and Middle East
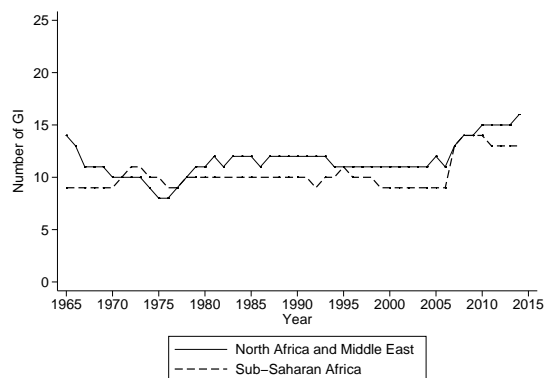
(f) Institutes in Africa and Middle East
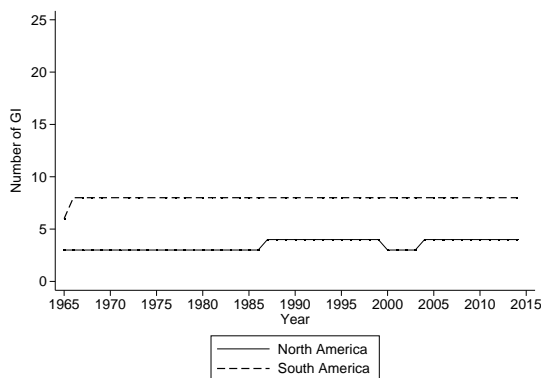
(g) Countries with institutes in America
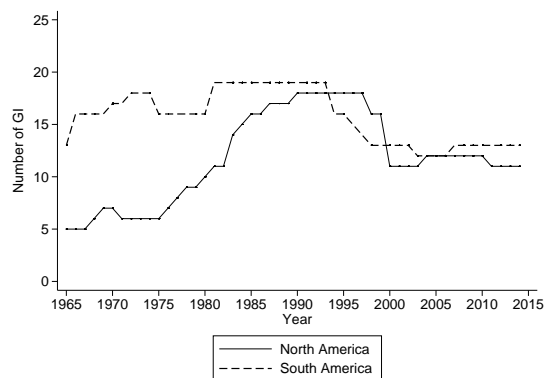
(h) Institutes in America

Figure 2.3: The GI by world regions.

2.4.2 Course and exam participation

While the overall number of institutes with language services has been largely constant, the numbers of course and exam participation show a different pattern with significant differences across regions. Figure 2.4 shows total numbers of exams, registrations and sold course units as well as numbers for Europe with Central Asia and for the rest of the world with some further disaggregation. For all three variables, we can see a similar trend (see panels (a), (b) and (c)): after a long period of quite stable numbers of course and exam participation, numbers have strongly increased since 2005.

In panels (d), (e) and (f), the three variables for the European and Central Asian countries are disaggregated further according to the respective GI regions. Panel (d) shows that the largest part of the increase in exam participation after 2005 took place in Southeast Europe, but also in other regions (except Northwest Europe, and Central Eastern Europe) numbers for exam participation increased. For registrations, we see a sharp increase for Central and Eastern Europe and Central Asia in panel (e), which results from the expansion of the GI in these regions in the 1990s as described in Chapter 2.4.1. Registration numbers for Southeast Europe and Northwest Europe remained quite constant in the whole period from 1990 to 2014. On the contrary, registrations in Southwest Europe decreased until 2008. When the recession set in, numbers began to increase again. Similar observations can be made for the years from 1997 onward in panel (f) which presents sold course units. Before 1990, only three of the regions offered language courses. In particular the number of sold course units was very high in Southwest Europe. The numbers in Southeast Europe increased from 1972 to 1989 continuously and remained constant in Northwest Europe.

The last row of Figure 2.4 disaggregates the variables on the continent level for the rest of the world. Panel (g) shows that the increase in exam participation after 2005 took mainly place in Asia and Africa with Middle East, while in America numbers only increased slightly. For course participation the picture is different. Registrations in America decreased in the 1990s and then stayed constant on a lower level. In Africa the number of registrations increased slightly, while in Asia the numbers increased sharply from 2000 onward. Numbers for sold course units after 1997 again followed the same trends as can be seen in panel (i), while in the period before 1990 the number of sold course units were quite stable in all three continents.
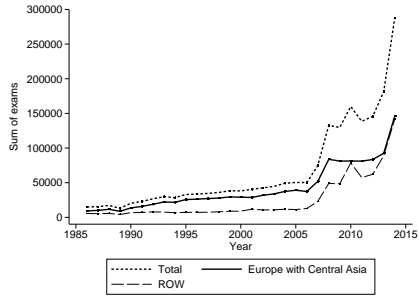
(a) Exams (total)
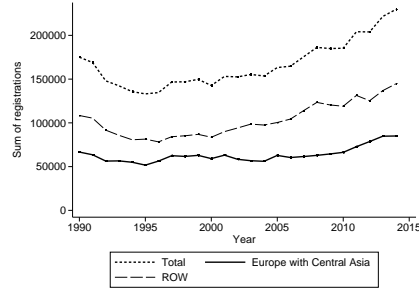
(b) Registrations (total)
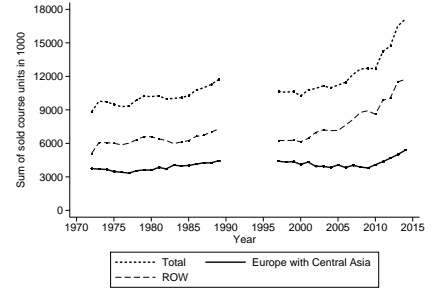
(c) Sold course units (total)
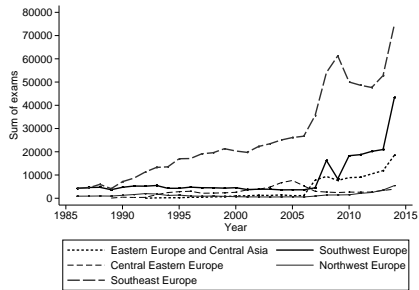
(d) Exams (Europe with Central Asia)
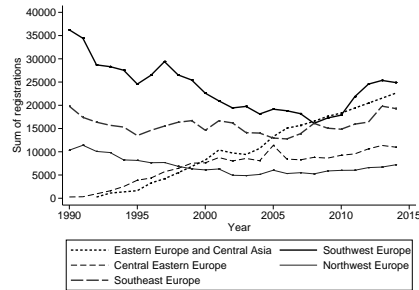
(e) Registrations (Europe with Central Asia)
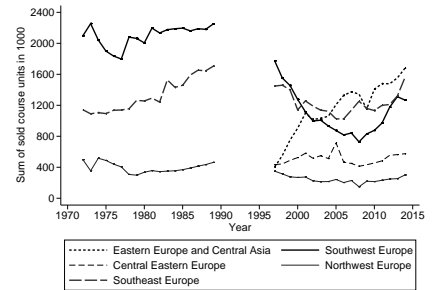
(f) Sold course units (Europe with Central Asia)
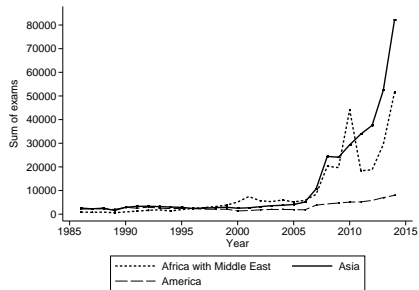
(g) Exams (ROW)
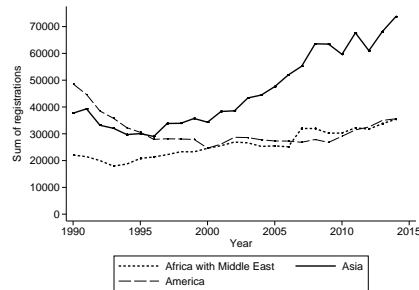
(h) Registrations (ROW)
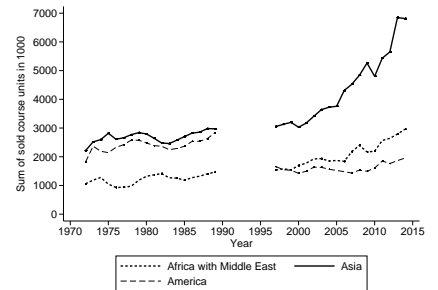
(i) Sold course units (ROW)

Figure 2.4: Participation information abroad.

There has been a general expansion in total course and exam participation over the years, similar to the expansion of the GI to more countries (see Figure 2.2). One reason for the significant increases especially since 2005 might have been the language requirement introduced in 2007 in Germany for non-EU citizens who immigrate to reunite with their family. As the GI offers exams that are recognised for that purpose, the comparatively large increase in the number of exams could be a consequence of this requirement in non-EU countries. The exams can be taken without attending courses at the GI. It is therefore not a surprise that the numbers for course participation, i.e. registrations and sold course units, have increased comparatively less. Nevertheless, a slight increase can be seen for the registrations and sold course units as well, in particular for the rest of the world. As to the most recent increase in the number of exams, this is mainly attributable to a change of the structures of the exams in 2013 (Goethe-Institut, 2014). Before that change, all parts (reading, writing, speaking,...) had to be taken together, while now each part constitutes a separate exam.

Figure 2.5 illustrates the number of registrations in each country in the year 2014.[13] This can be related to the number of institutes in the different countries as well as to their different sizes. In 2014, the average institute had 1,963 registrations and half of all institutes had 1,500 or less registrations. On the other hand, the five largest institutes were spread all over the world: Bangkok (6,800), Rome (6,500), Moscow (6,000), Madrid (5,200) and Mexico (5,200).[14]

2.5 Language learning across time – in Germany

Opportunities to learn German do not only exist abroad. There is also the option to learn the German language in one of the institutes in Germany. Participants can either come explicitly for a course and return to their home country at the end of it; or they are migrants in Germany who want to improve their proficiency of the host country's language.

---

[13] See Figure 2.A1 in the Appendix for a graphical representation based on registrations as a share of the populations in the countries where the institutes are located. Population data are from Feenstra et al. (2015) and United Nations (2017).

[14] Note that numbers for Rome and Madrid are reported jointly with other institutes in their country.
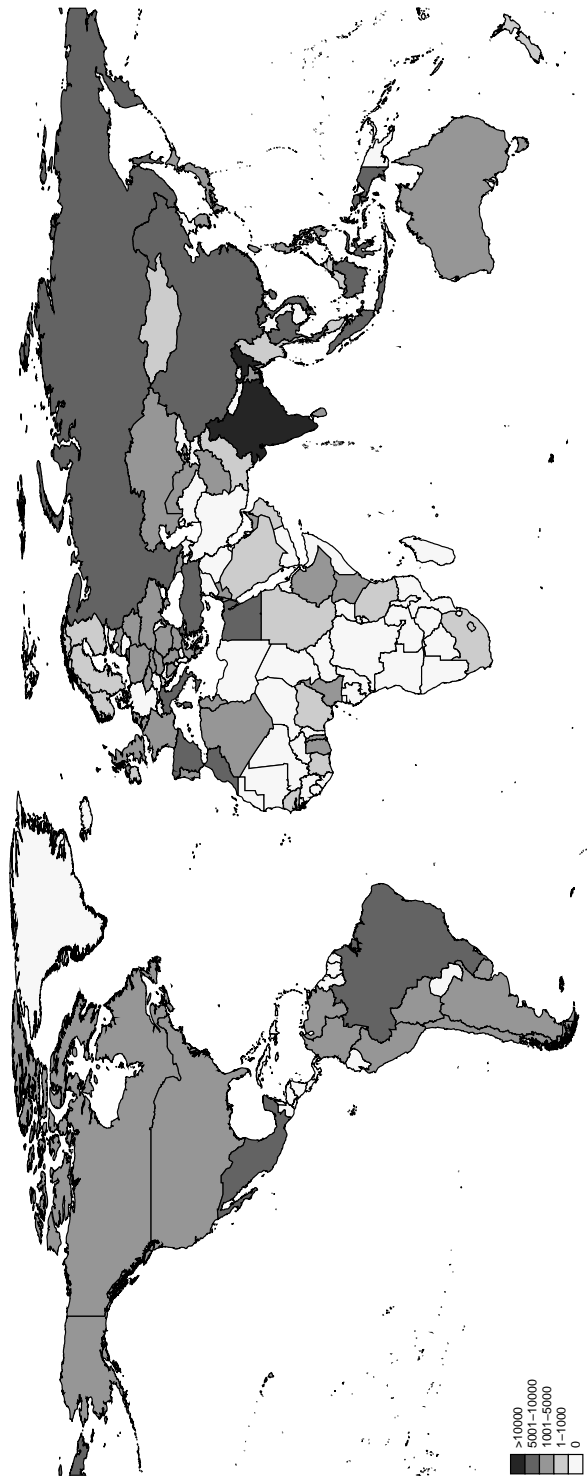
Figure 2.5: Registrations worldwide in 2014 (absolute numbers).

>10000
5001–10000
1001–5000
1–1000
0
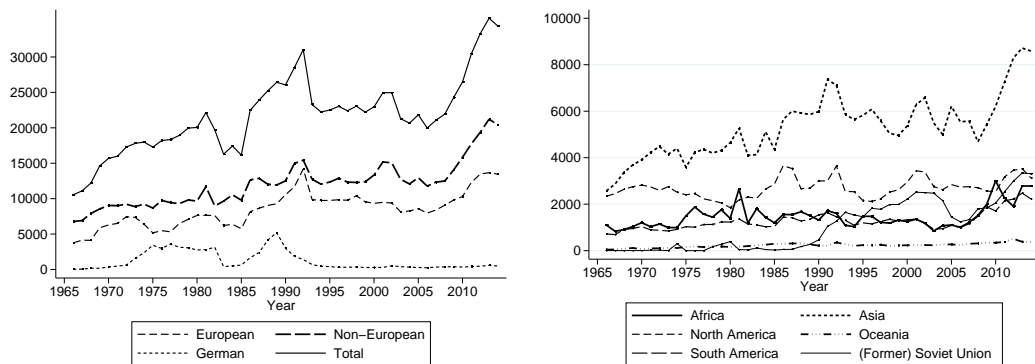
(a) Registrations in Germany from 1966 to 2014

(b) Non-European registrations in Germany from 1966 to 2014

Figure 2.6: Course registrations in Germany.

Figure 2.6 plots the development of the aggregate registrations by continents. Panel (a) distinguishes between European, Non-European[15] and German students, who are mainly ethnic German repatriates[16], and reports the aggregate of all students. The trend was similar for students from European and non-European countries and so was their relative importance over time. In the mid of the 1980s, the number of registrations by European and non-European students began to increase until a peak was reached in 1992, where 31,179 students from all over the world studied at German institutes. After 1992, language course participation decreased and stayed on a quite constant level until 2006. Since 2006, there has been a steep increase in language course participation of European and Non-European students where the total number of registrations reached the maximum in 2013 with 35,468 students. In 2014, registrations by European students amounted to 13,459 and by non-European students to 20,397 from a total of about 200 countries.

---

[15] Due to data limitations in the first years after the dissolution of the Soviet Union, it was not possible to distinguish successor states and assign them to continents. Therefore, the (former) Soviet Union categorises a region itself, while the regions Europe and Asia exclude successor states of the former Soviet Union. Students from the (former) Soviet Union are included as non-European students in Figure 2.6.a

[16] These ethnic German repatriates (Aussiedler) emigrated mostly from Eastern Europe or the former Soviet Union. In the annual reports, in some years they are referred to as students from Germany and sometimes they are categorised as "Aussiedler".

In panel (b) non-European regions are disaggregated by continents and the (former) Soviet Union. Throughout the whole period, students from Asia have constituted the largest group, while the smallest group have been students from Oceania. The number of registrations by Asian students has increased sharply since 2008 which has contributed significantly to the overall increase in course participation numbers of non-European students.

During the existence of the Soviet Union numbers of registrations from that region were very low, but after its dissolution in 1990 the number of students increased sharply up to around 2,000 registrations in 2003. Afterwards, a downward trend began, until course participation started to increase again in line with the common upward trend. At the end of the period, the number of registrations from the former Soviet Union had converged to the number of registrations by students from North America, which was relatively stable over the whole period from 1966 to 2014. The number of registrations by students from Africa and South America were of the same magnitude for most of the time. They also increased, however, following the common trend from 2006 onward. In 2014, 8,587 students came from Asia, 3,311 from successor countries of the Soviet Union, 3,121 from North America, 2,785 from Africa, 2,221 from South America and 372 from Oceania.

The distribution of the number of registrations is clearly right-skewed. In 2014, around 83.5 % of the countries had less than 250 registrations for students studying German at an institute in Germany. A closer look at the left side of the distribution shows, that many countries had even much fewer registrations. While 56.2 % of the countries had less than 50 students at the GI in Germany, the median of the distribution is 29 students and the first quartile is 4 students. On the other side of the distribution, a few countries contribute a large proportion of the total number of registrations at German institutes. While many of them are European[17], only one of the larger countries is South American (Brazil with 1,182 registrations) and African (Libya with 1,318 registrations). Also many students come from the four Asian countries Saudia Arabia (1,509), China (1,319), Japan (1,161) and India (756). With Mexico (608) and the United States (2,000), North America also contributes a large proportion of registrations. Also, another large group of students comes from Russia, as largest successor of the Soviet Union (2,028). Figure 2.7 illustrates the number of registrations by nationalities in the year 2014 and shows the dif-

---

[17] Origin countries in descending order of numbers of registrations in German institutes are: Italy (2,554), Spain (2,359), France (1,569), Turkey (931), Great Britain (886), Switzerland (722) and Poland (661).
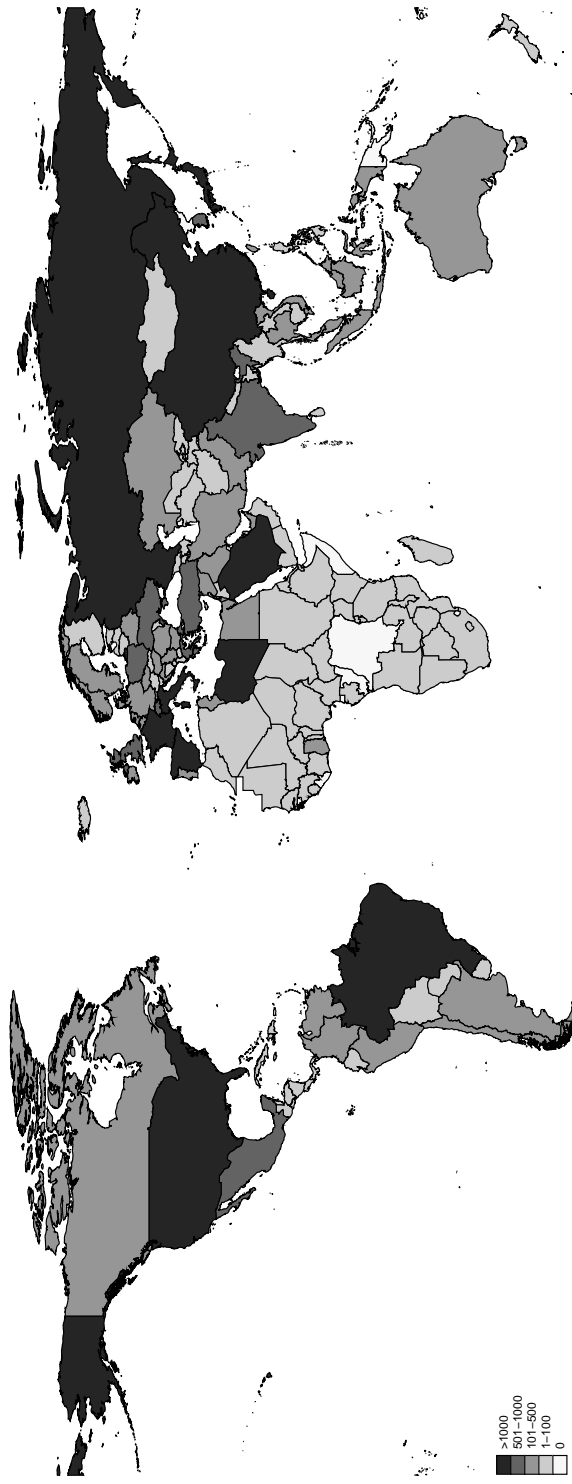
Figure 2.7: Registrations in Germany by nationality in 2014 (absolute numbers).

>1000
501–1000
101–500
1–100
0

ferent importance of different nationalities for the study of German at institutes in Germany.[18]

2.6   Conclusion

This chapter presents a comprehensive overview of German language learning for more than 100 countries (including Germany) over a period of 50 years. Our new and unique data allow providing information about the supply of language-learning opportunities, i.e. the number and the geographic distribution of institutes, and the demand in the form of course and exam registrations. These data do not only show the development of language learning for the German language across time, they also underline common trends and heterogeneities across regions.

With respect to the supply of language-learning opportunities, the years from 1965 to 2014 can be divided into three periods. Before 1980s, the GI was mostly present in America and the "old Europe". In the aftermath of the fall of the iron curtain, the GI expanded to former socialist countries, while simultaneously decreasing its large number of institutes in America and the "old Europe". From the early 2000s onward, the presence of the GI has been quite stable, only in Africa the GI expanded to new countries. Interestingly, while the number of countries with a least one institute continuously increased over the whole period, the overall number of institutes fluctuated for most of the time between 130 and 160 with peaks in the early 1970s, the early 1990s and in the recent past.

As to the demand side, course and exam participation was relatively stable at institutes in Germany and abroad in the first decades. At institutes in Germany, participation numbers began to increase in the mid of the 1980s. Since 1990 or 1995, respectively, there has been a general expansion in total exam and course participation also at institutes outside Germany similar to the expansion of the GI to more countries – even though the number of institutes has not shown the same pattern. For institutes outside Europe, this has been mainly driven by Asia where course participation numbers started to rise around 1995 and numbers for exam participation have much grown since 2005.

---

[18]   See Figure 2.A2 in the Appendix for a graphical representation based on registrations as a share of the populations in the countries where participants are from. Population data are from Feenstra et al. (2015) and United Nations (2017).

A steep increase of registrations by Asians at institutes in Germany followed around 2008. Similarly, registrations by participants from Africa have started to increase after 2005.

In Europe with Central Asia, most changes are driven by three regions: In Southwest Europe, course participation numbers decreased from 1990 onward and only started to increase again around 2010 – about at the same time when numbers for exam participation started to increase as well. The same pattern holds for registrations by Europeans at institutes in Germany. Registrations from Eastern Europe and Central Asia have shown an upward trend since 1990 without, however, a comparable increase in exam numbers. On the contrary, exam numbers for Southeast Europe have been rising from 1990 onward, while numbers of course registrations have not changed much. Numbers of registrations by students from the former Soviet Union at institutes in Germany confirm this rising interest in the German language in particular from 2005 onward.

These different patterns highlight the heterogeneity across time and even more so across regions. In addition, even for the same region, differences can be observed depending on which aspects of language learning (demand-side or supply-side) and on which specific forms of participation (course or exam participation) one is focusing. While we provided some possible explanations for some of the observed patterns, for example, the language requirement introduced in 2007 in Germany for non-EU citizens who immigrate to reunite with their family, more research is needed to fully understand the reasons behind the developments and to which extent they can be attributed to policy changes in Germany or abroad or to other changes of the institutional setting or the individual motives. Also global events, like the financial crisis, might have had an effect. At the same time, the data themselves can contribute to a better understanding of the economic and cultural relations between Germany and countries and regions worldwide and their development over the last five decades.

Appendix to Chapter 2

2.A   Appendix A
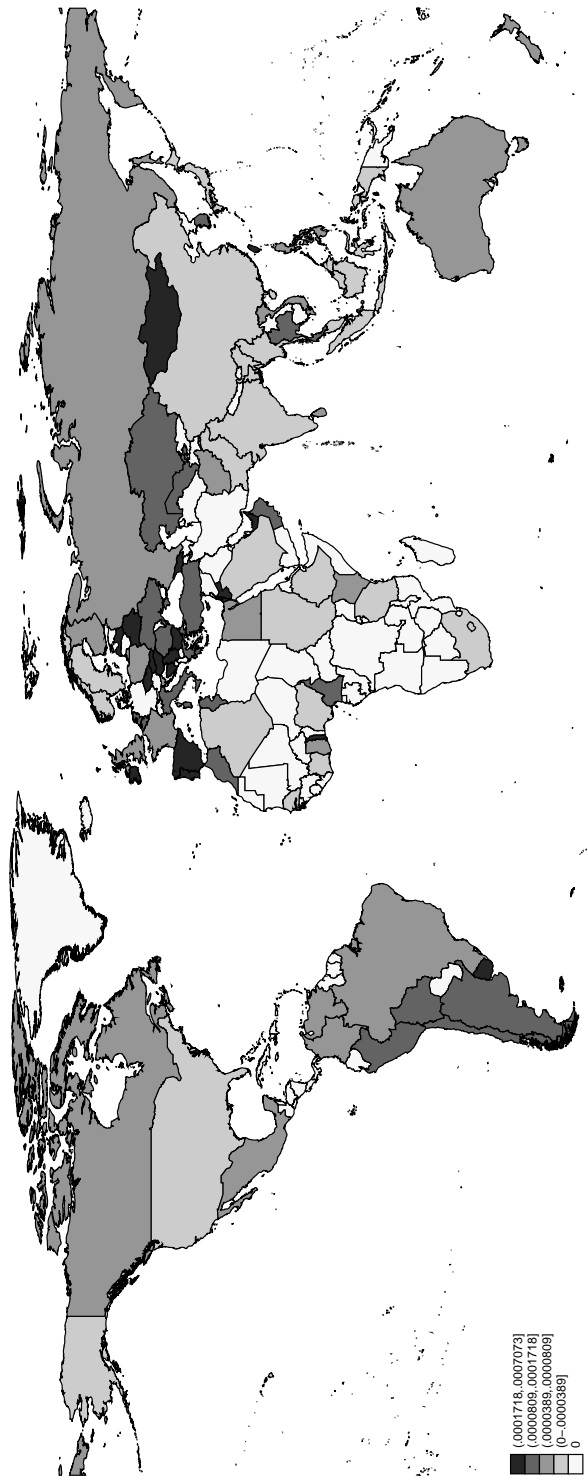
Figure 2.A1: Registrations worldwide in 2014 (as shares of countries' populations).
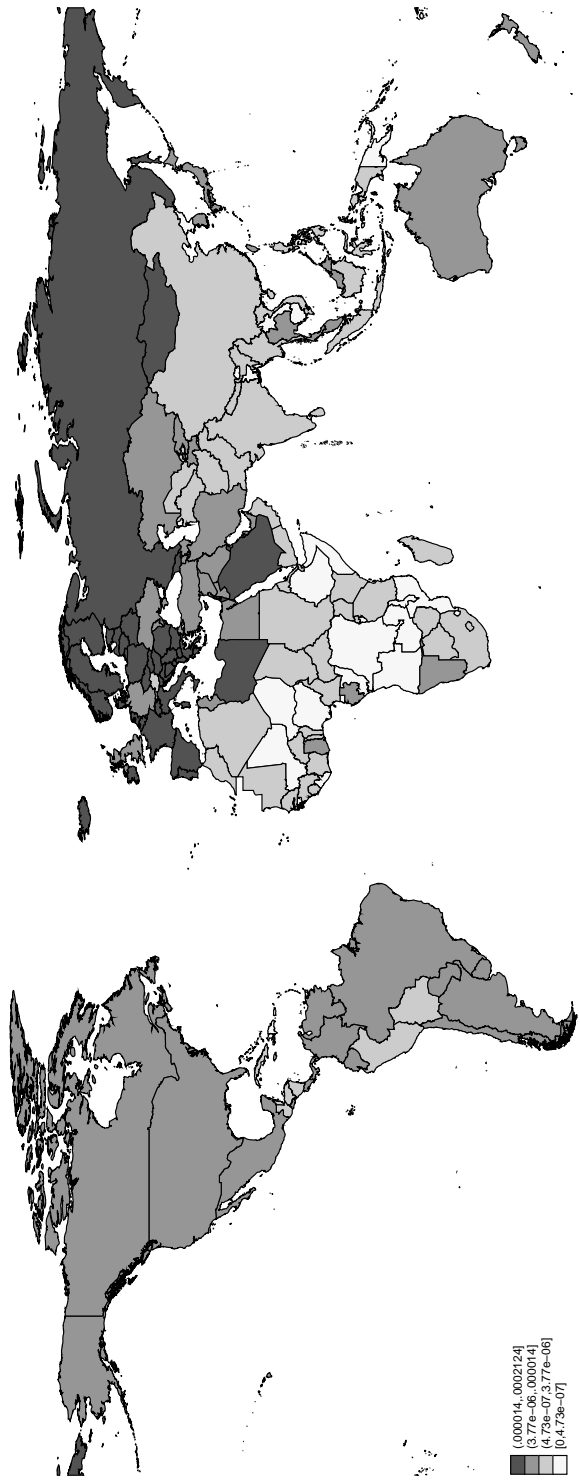
Figure 2.A2: Registrations in Germany by nationality in 2014 (as shares of countries' populations).

[.000014,.0002124]
[3.77e−06,.000014]
[4.73e−07,3.77e−06]
[0,4.73e−07]

Table 2.A1: Regions according to the regional organisation by the GI since 2008

Central Eastern Europe	Northwest Europe	Southwest Europe
Czech Republic	Denmark	Belgium
Estonia	Finland	France
Hungary	Great Britain	Italy
Latvia	Iceland	Luxembourg
Lithuania	Ireland	Portugal
Poland	Netherlands	Spain
Slovakia	Norway	
Slovenia	Sweden	

Southeast Europe	Eastern Europe and Central Asia	South America
Bosnia and Herzegovina	Belarus	Argentina
Bulgaria	Georgia	Bolivia
Croatia	Kazakhstan	Brazil
Cyprus	Russian Federation	Chile
Greece	Ukraine	Colombia
Macedonia	Uzbekistan	Peru
Romania		Uruguay
Serbia		Venezuela
Turkey		

North America	Sub-Saharan Africa	North Africa and Middle East
Canada	Angola	Algeria
Costa Rica	Burkina Faso	Egypt
Cuba	Cameroon	Iraq
Mexico	Congo	Israel
United States of America	Cóte d'Ivoire	Jordan
	Ethiopia	Lebanon
	Ghana	Libya
	Kenya	Morocco
	Madagascar	Oman
	Malawi	Saudi Arabia
	Nigeria	Palestinian Territories
	Rwanda	Sudan
	Senegal	Syrian Arab Republic
	South Africa	Tunisia
	Tanzania	United Arab Emirates
	Togo	
	Uganda	
	Zimbabwe	

Southeast Asia, Australia and New Zealand	South Asia	East Asia
Australia	Afghanistan	China
Indonesia	Bangladesh	Hong Kong
Malaysia	India	Japan
Myanmar	Iran	Mongolia
New Zealand	Nepal	Republic of Korea
Philippines	Pakistan	Taiwan
Singapore	Sri Lanka	
Thailand		
Viet Nam		

Chapter 3

Language-learning opportunities and migration

Abstract.[19] This chapter analyzes the effect of German language-learning opportunities abroad on migration to Germany. We use information about the Goethe-Institut (GI), which promotes the German language worldwide. Our unique dataset covers 69 countries for the period 1977 to 2014. In this multiple-origin and single-destination framework, we estimate fixed-effects models as our basic specification. We find evidence that the number of language institutes of the GI in a country is positively correlated with migration from that country to Germany. To establish causality, we consider Switzerland as an alternative destination, as the number of institutes is exogenous to migration to Switzerland. We find that the number of institutes affects migration to the German-speaking part of Switzerland but not to the French- and Italian-speaking parts. Backed by further extensions, which control for the presence of multilateral resistance, our results point to a causal effect of language-learning opportunities on migration to Germany.

3.1 Introduction

A large part of the migration literature focuses on migrants' proficiency in the language of the destination country. It has been shown that proficiency improves labour market and integration outcomes. Language skills increase earnings (see e.g. Chiswick and Miller, 1995; Dustmann and Van Soest, 2001) and employment (Dustmann and Fabbri, 2003). At the same time, the probability of intermarriage becomes larger and the likelihood of living in an ethnic enclave decreases (Bleakley and Chin, 2010).

Given these benefits of language proficiency, potential migrants can be expected to consider language-related aspects in their migration decision and their location choice. Indeed, many studies show that language is an important determinant of migration flows.[20] To capture the linguistic relationship between the migrants' mother tongue and the language of the destination country, measures of linguistic distance are often used. Adserà and Pytliková (2015) and Belot and Ederveen (2012) find evidence of a negative effect of linguistic distance on international migration flows, based on different measures but, as Bredtmann et al. (2020) show, this effect is reduced by a large ethnic network in the host region which constitutes a substitute for linguistic proximity. However, the concept of linguistic distance neglects language acquisition of potential migrants before migration, which can alleviate or overcome the negative effects of linguistic distance.

The aim of this chapter is to study the effect of the presence of language-learning opportunities on migration. We build a random-utility model. Individuals want to maximize expected utility of migration, of which the expected wage income net of migration costs is an important component. Acquiring language skills of the destination country can increase expected net wage income if the benefits, in terms of higher wages abroad, exceed the costs of learning. Language acquisition can happen at different points in time. We distinguish language learning as a child and as an adult, since this is different in terms of costs and direction of causality in the context of migration. If language skills are acquired during childhood or adolescence, the decision is more likely determined by factors outside the learners' direct control. These factors may be related to parents' preferences and to the school system via compulsory foreign language learning. Language proficiency often comes with few or no costs, while it might affect later migration decisions. Aparicio Fenoll and Kuehn (2016) and Apari-

---

[20] For analyses of the role of languages for trade, see e.g. Melitz (2008).

cio Fenoll and Kuehn (2019) use compulsory language learning at school as a measure for language skills beyond linguistic properties. While Aparicio Fenoll and Kuehn (2016) find a positive relationship with migration flows within the European Union, Aparicio Fenoll and Kuehn (2019) show that migrants with English language skills move to countries where these skills are more scarce.

Language learning during adulthood, on the contrary, is more likely a decision of the learners themselves. This decision can be driven by different motives of personal or economic nature, like better job opportunities at home or abroad. In this context, migration decisions might lead to pre-migration language learning. The direction of causality with language learning of adults can therefore be opposite to the direction with language learning during childhood or adolescence. Uebelmesser et al. (2021) analyse determinants of German language learning of adults by using data from language institutes worldwide and in Germany. They show that general migration and student migration are indeed important determinants of language learning.

For this chapter, we focus on language learning of adults. We use a unique panel dataset for 69 countries for the period 1977 to 2014 with information about the worldwide presence of institutes collected from the annual reports of the Goethe-Institut (GI).[21] The GI is an association which is an important actor of Germany's foreign cultural policy and which promotes German culture and language worldwide (Auswärtiges Amt and Goethe-Institut, 2004). Via its institutes, it offers language courses and standardized language exams as well as information on German culture and society in many different forms, such as cultural events and libraries.

Based on fixed-effects (FE) estimations, we find evidence that the number of language institutes in a country is positively correlated with migration from that country to Germany. By distinguishing between institutes that offer language services and those which do not, we can show that the correlation is indeed driven by language-learning opportunities and not by other factors, such as the provision of information about German culture and society. The relation is stronger for poorer countries and for countries that are linguistically and geographically more distant from Germany.

For a causal interpretation, we address the issue of reverse causality following two approaches: First, we consider the strategies and overall objectives of

---

[21] In this chapter we stick to the following convention: when referring to the association of the Goethe-Institut we use the abbreviation "GI". When talking about specific branches of the GI abroad, we refer to them as "institutes".

the GI. The GI and the Federal Foreign Office (FFO) jointly decide where to open and close institutes by a complex process.[22] Even though there are no indications that migration to Germany plays any particular role in this process, a possible impact cannot be fully ruled out. Therefore, second we also estimate migration flows to the German-speaking part of Switzerland. As migration flows to Switzerland do not have any impact on the decision to open or close institutes, the number of institutes can be considered exogenous to those migration flows. At the same time, it is possible that language-learning opportunities affect migration flows to Switzerland, at least to the German-speaking part. With this specification, we find similar results for the effect of the number of institutes on migration flows to the German-speaking part of Switzerland as to Germany. Furthermore, we apply the common correlated effects estimator (CCE) by Pesaran (2006) to control for multilateral resistance to migration (Bertoli and Fernández-Huertas Moraga, 2013). If migration flows do not only depend on characteristics of the origin and destination countries, but also on the attractiveness of alternative destinations, the FE-estimations are biased. In our case, this bias helps us interpret the relation between the language-learning opportunities and migration flows in more detail, as it allows us to to separate the total effect into a direct and an indirect effect. We find that half of the effect constitutes a direct effect on migration to Germany of individuals who would not have migrated otherwise. The other half is an indirect effect of redirected migration flows to Germany due to a relative decrease in attractiveness of alternative destinations in the aftermath of opening a language institute.

This chapter contributes to the literature in several ways: First, our unique dataset provides new information on language learning for a long time period of 38 years for and a large number of countries all over the world. Second, these data allow us to study the language learning process for adults in the context of migration. So far, the literature mainly looked at linguistic properties. One exception is Aparicio Fenoll and Kuehn (2016), who focus on language learning during childhood. This chapter seeks to answer the question whether the provision of language-learning opportunities, possibly by a potential destination country, has an effect on adults' language learning decisions in the context of migration.This is a topic of large policy relevance in general, and, more specifically, against the background of the new Skilled Immigration Act effective in Germany since March 2020, which aims at migrants from third countries and assigns an active role to the GI (Bundesregierung, 2018; Goethe-Institut, 2018). While Jaschke and Keita (2021) have found that the presence of GI affects the

---

[22] See Chapter 3.2 for more information about the institutional background.

self-selection pattern of immigrants related to their language skills and relevant labour-market characteristics at arrival, we show that the number of institutes not only affects the quality of immigrants but also their quantity.

The remainder of the chapter is structured as follows: Chapter 3.2 describes the history of the GI, its institutional framework and its objectives. Furthermore, it explains the decision process of opening and closing institutes. Chapter 3.3 presents the data and provides descriptive statistics. Chapter 3.4 describes the theoretical micro-foundation for individual migration decisions and presents the estimation strategy. Chapter 3.5 presents our main results. In Chapter 3.6, we provide evidence to support a causal interpretation and deal with multilateral resistance to migration. Chapter 3.7 concludes.

3.2   The Goethe-Institut

The GI is a cultural institute which supports the foreign cultural policy of the German government. As part of this task, the GI is responsible for international cultural cooperation and exchange. Furthermore, it provides information about the social, political and cultural life in Germany. Another important aspect of the GI's work is the promotion of the German language.

To fulfill its duties, the GI has institutes all over the world, which provide language courses and standardized exams as well as scholarships for language learning and training for teachers (Auswärtiges Amt and Goethe-Institut, 2004). Furthermore, the GI maintains libraries and organizes cultural events (Auswärtiges Amt and Goethe-Institut, 2004). In 2014, there were in total 143 institutes in 93 countries (see Figure 3.1) out of which 126 institutes offered language services. The sizes of those institutes in terms of registrations for language courses ranged from 2 registrations in Beirut to 6800 registrations in Bangkok. The average number of registrations amounted to 1963. In total, there were 229,702 course registrations in 2014. While the GI is mainly funded by the German government, language courses are financed by course fees (Goethe-Institut, 2014).

The process of opening and closing institutes is of particular importance for the analysis in this chapter. According to officials of the GI, this is a joint decision by the FFO and the GI on the basis of their general objectives, their regional and strategic focus as well as legal, political and social considerations related to the potential host country. For a better understanding of how the distribution of institutes worldwide evolved and how different factors contributed to it, we

Figure 3.1: The distribution of the GI in 2014.

provide a brief overview of the history of the GI in the context of Germany's foreign cultural policy.

The GI in its current form was founded in 1951 as a successor of the "Deutsche Akademie" and was meant to act as a partner of the German government to support its foreign cultural policy after the Second World War (WW II).[23] For this, the GI and the German government built on existing structures from before WW II (Singer, 2003). While the GI opened its first new international institutes in the 1950s, cultural institutes from other institutions were included stepwise into the GI in the late 1950s and early 1960s (Wittek, 2016), both contributing to the regional expansion of its activities.

In terms of strategy and overall objectives of the GI, several phases have to be distinguished. Until the mid 1960s, the main objective was to convey a positive image of Germany abroad and to support economic relations between Germany and the host countries of the institutes. In that period, it was also important as far as international relations were concerned to be one step ahead of the German Democratic Republic (Kaitinnis, 2018). When the social-democrat Willy Brandt became Federal Minister for Foreign Affairs in 1966, the role of foreign cultural policy changed. In particular, there was a shift from cultural export to cultural exchange with an emphasis on the importance of culture for

---

[23] Other actors were e.g. the German Academic Exchange Service (DAAD), the "Deutsche Welle" (an international broadcaster) and the Alexander von Humboldt Foundation.

international cooperation and peacekeeping (Singer, 2003). In 1969 and 1976, the GI and the FFO signed two General Agreements. The Agreement of 1976 defined the rights and duties of the GI and the FFO and their overall relation for several decades and was only replaced in 2004 by the current agreement. From the Agreements, two points become clear: First, more generally, one can see that the FFO sets the general framework, including the objectives of foreign cultural policy, while the GI is relatively independent within this framework in choosing its activities. Second, more specifically related to the analysis here, until the late 1970s, migration was neither an explicit factor nor an objective linked to any activities of the GI, including the opening or closing of institutes (Schödel, 2007).

With the change to a conservative government in 1982, the focus shifted from cultural work to language promotion and thus more to non-political activities (Kaitinnis, 2018). Language promotion was understood to support economic relations and to foster German exports. In regional terms, Latin America became the main area of expansion. The fall of the iron curtain and the end of the cold war in the early 1990s changed German foreign cultural policy again. The regional focus of the GI shifted to Central and Eastern Europe, where the GI was supposed to support the political and economic transformation process (Kaitinnis, 2018). In 2000, the FFO published its new strategy plan "Konzeption 2000" for its foreign cultural policy. Against the background of a tighter budget, the emphasis was on restructuring and modernizing the organization of foreign cultural policy and included a redefinition of its objectives. The main focus was placed on European cultural policy and again on language promotion, among other goals (Singer, 2003).

During that period, migration did not play a role in determining activities of the GI, nor was migration mentioned in any official document or other material as an objective of foreign cultural policy in general or related to the GI. This has only recently changed in a stepwise way. The first step concerns the change of the immigration law in 2007. Since then, family members who migrate with their spouse to Germany from third countries must possess basic knowledge of the German language (Bundesamtes für Migration und Flüchtlinge, 2008, p. 117). This requires an A1-certificate according to the Common European Framework of Reference for Languages, which the GI issues among others. While this policy change increased the demand for language courses and exams, migration and integration were still not among the objectives of the GI, as evidenced by the first "Target Agreement" for the years 2008-2010 between the FFO and the GI (Goethe-Institut, 2010). In fact, it was not before the

"Target Agreement" for the years 2011-2014 that migration and integration were mentioned among the objectives of the GI (Goethe-Institut, 2013).[24] After our observation period, which goes until 2014, the focus of the GI has shifted more toward fostering migration and integration. More precisely, the GI is supposed to play a major role in the context of the new Skilled Immigration Act (Bundesregierung, 2018; Goethe-Institut, 2018) effective since March 2020.

To summarize, migration was not a relevant topic for the GI for a long time. This changed toward the end of our observation period and has become one of the main topics since then. In our study about the effect of the presence of language-learning opportunities via the GI on migration to Germany, we first follow our reasoning from above and assume that the opening and closing of institutes has been independent from any migration-related considerations. We deal with the issue of potential reverse causality explicitly in Chapter 3.6.

3.3 Data and descriptive statistics

In the following, we describe the data used for the analysis and present descriptive statistics.

3.3.1 Dependent variable: migration rate

For the dependent variable, we use the yearly migration rate, i.e. immigration flows to Germany divided by the population size of the origin country. Migration data are provided in the 'Wanderungsstatistik' by the German Federal Statistical Office (Destatis, 2016). The data document the number of foreign citizens that move to Germany and register their residence in a given year. These immigrants are then categorized according to their citizenship. As this registration is mandatory for all foreign residents staying for more than two months, these data capture legal immigration to Germany in a comprehensive way. Data on population size of the origin countries comes from the Penn World Table (PWT) 9.0 (Feenstra et al., 2015).

---

[24] For robustness, we run our main specifications also for the period 1977-2010. See Chapter 3.5.1 below.

3.3.2 Independent variable: language-learning opportunities

The data on language-learning opportunities are derived from a new dataset comprising information about the presence of the GI and the number of institutes at the country-level (see Chapter 2 and Uebelmesser et al., 2018a). The GI has published annual reports continuously since 1965 in which activities of each institute, including statistics of language course and exam participation, have been reported. The dataset is constructed from these reports and contains information about the presence of GI on the institute-level (see Chapter 2 for a more detailed description of the dataset).

For our analysis, we aggregate the data on the number of institutes in a given country and year. As not every institute offers language services, we construct three different variables: the number of institutes without language services, the number of language institutes, i.e. institutes that offer language services, and the number of all institutes, i.e. the total number of institutes with and without language services. To analyse the effect of language-learning opportunities, for our preferred specification we restrict our attention to institutes with language services.. For robustness checks, we also use information about institutes without language services.

3.3.3 Other control variables

Several additional variables are included as control variables. First, we control for the economic condition in the origin country by including GDP per capita. For this, we use the expenditure-side real GDP ('rgdpe') and data on the population size both from the PWT 9.0 (Feenstra et al., 2015). Second, we construct two dummies which measure the free movement of labour within the European Union (EU). The first one indicates the period of limited access to the German labour market for those countries who became member of the EU 2004 or later (EU12). The second one indicates the period of unlimited access to the German labour market for those EU12 countries as well

as for the countries which joined the EU before (EU15).[25] Third, we include data on bilateral trade flows,[26] i.e. the sum of exports from and imports to Germany, provided by Destatis (2018b) as a control for possible economic relations. Fourth, data on the stock of migrants by origin countries in Germany from the 'Ausländerzentralregister' (Destatis, 2018a) allow controlling for ethnic networks. Fifth, we include a combined measure of political rights and civil liberties from Freedom House (2018). We coded the variable in a way such that higher values indicate more freedom in the origin country. Finally, to capture the effect of violent conflicts in the origin country on migration to Germany, we use the UCDP/PRIO Armed Conflict Dataset Version 4-2016 (Allansson et al., 2017; Gleditsch et al., 2002). The measure takes the value one in case of 25 to 999 battle-related deaths for a given origin country in a given year and the value two in case of more than 999 battle-related deaths. We construct our conflict variable by adding, for a given conflict, the values since the beginning of the conflict up to a given year.[27] This allows us to capture the longer-term burden to the population caused by an ongoing conflict

3.3.4 Sample construction

To construct our dataset, we proceed as follows: First, we include all countries for which we have information on GDP, trade flows and population size for all years between 1977 and 2014. Second, we restrict the sample according to the availability of migration data: We only include countries for which we have migration data in all years but 1990, 2000 and 2001. In these three years, migration data of Destatis included many missing observations due to changes in the data generation process. We interpolate these missing observations linearly

---

[25] E.g. for Poland, which is one of the EU12 countries, the value of both dummies is zero before 2004. However, between 2004 and 2010, the dummy indicating limited free mobility takes the value one, while the dummy indicating unlimited free mobility continues to be zero. From 2011 onward, the dummy indicating unlimited free mobility changes to take the value one, whereas the dummy indicating limited free mobility becomes zero again.

[26] For a comprehensive overview of the interrelations between trade and migration, see Felbermayr et al. (2015).

[27] In case of more conflicts within a country and year, we only consider the conflict with the largest number of battle-related deaths. The dataset is based on the nationality of those participating in the conflict and not on the place of the conflict. As we want to control for violent conflicts in the origin countries, we set the conflict variable to zero when the conflict was not local. This was only the case for conflicts that involved the USA, Great Britain and Australia.

on the basis of the years 1989 and 1991, and 1999 and 2002, respectively. Note that the sample only contains non-zero values for migration flows. Third, we add the data from our GI dataset about the presence of the GI and the number of institutes per country and year, and we assign the value 0 to these variables for countries and years that are not included in the GI dataset.

Furthermore, in some cases, the information about language services in the annual reports are reported jointly for two or more institutes without clarifying which of these institutes actually offered language services. This is mainly the case for institutes that are subsidiaries of main institutes.[28] To avoid any ambiguity, we drop those countries where institutes that offer language services are not clearly distinguishable from those that do not offer language services. As this problem mainly occurs before 1977, we limit the observation period to the years afterwards.

Finally, we end up with a balanced dataset that includes observations for 69 countries in the period from 1977 to 2014 (see Table 3.A1 in the Appendix).

3.3.5 Descriptive statistics

Table 3.A2 in the Appendix provides summary statistics of the variables used in our analysis. Looking more closely at the presence of the GI at the country level in the period 1977 to 2014, we see that there was at least one institute in at least one year in 49 of the 69 countries; in 30 countries, the GI was present in all years of the observation period. The worldwide distribution of the countries in our sample is displayed in Figure 3.A1 where countries are grouped according to the number of years in which the GI was present (all years of the observation period, at least one year and less then all years, and no year). The countries in our sample are spread over all continents and so are those with presence of the GI.[29]

---

[28] For more information on the different types of institutes in the dataset, please refer to Chapter 2.

[29] Note that the (former) Soviet Union and other former socialist countries are not included in our sample. This is due to many newly founded states in the beginning of the 1990s and the lack of GDP and migration data. Furthermore, some Western European countries are missing, as joint reporting has been a widespread phenomenon in these countries. As a robustness check, we use an unbalanced sample which allows the consideration of many more countries, including former socialist countries. We also present two specifications where we include countries with joint reporting (see Chapter 3.5.2).

(a) Countries with GI
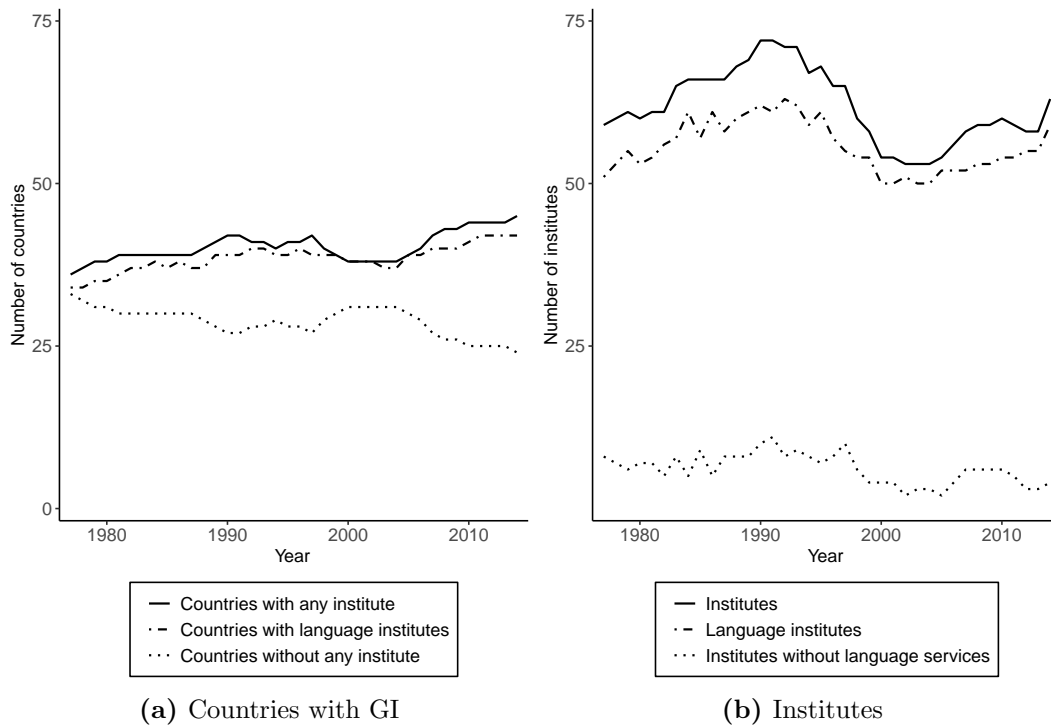
(b) Institutes

Figure 3.2: Number of countries and institutes (based on 69 countries in the sample).
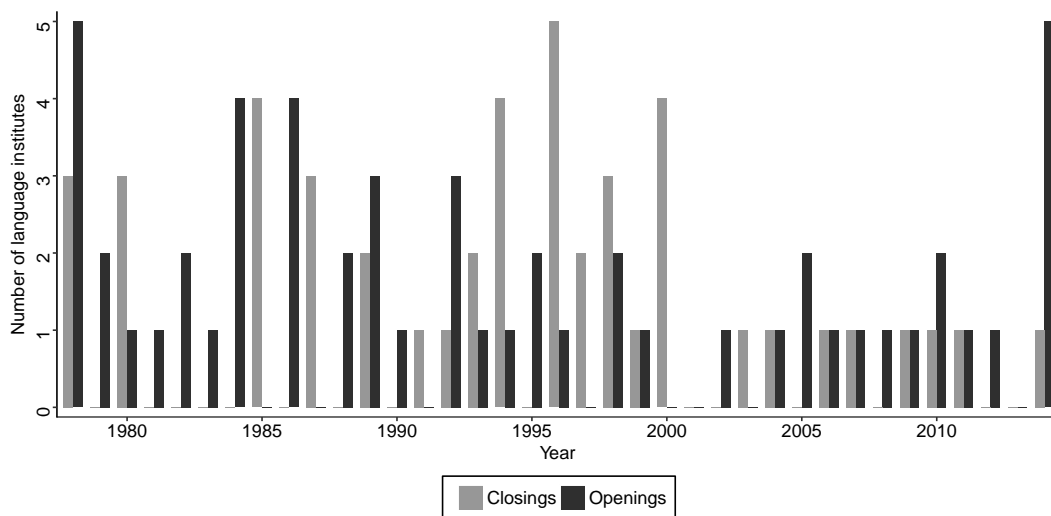
Figure 3.3: Number of openings and closings (based on 69 countries in the sample).

Figure 3.2a presents the number of countries per year in which the GI was present with any type of institute or, respectively, with at least one institute that offered language services. In addition, the number of countries without any institute in that year is included. In most of the years, the GI was present in around 40 out of the 69 countries. Throughout the entire period, the number of countries with institutes which only offered non-language services was negligible.

A different picture emerges when we compare the number of institutes per year with and without language services (Figure 3.2b). While in the entire period 1977 to 2014, the number of institutes which offered language services always clearly exceeded the number of institutes that did not, there were in each year between two (in 2002 and 2005) and 11 (in 1991) institutes without languages services.

The aggregate numbers obscure substantial variation in each year due to the closings and openings of institutes (see also Table 3.A2 for the overall, between and within variation). Figure 3.3 shows this for institutes with language services.[30]

---

[30]   Closing does not necessarily mean that the whole institute was closed, but rather that language services were no longer offered. The same holds for opening.

3.4 Theoretical background and empirical strategy

The micro-foundation of migration choice can be modeled in a random utility model (RUM) where individuals decide to stay in the origin country or to relocate to one out of alternative destination countries. They choose the country which maximizes their utility by comparing the attractiveness of each country and considering expected earnings and costs of migration, such as geographical or linguistic distance or visa regulations. Language learning likely increases expected earnings in the destination country by improving migrants' labour-market outcomes, but it also raises the costs of migration which include the costs of language acquisition. Individuals will opt for language learning if their utility is maximized by doing so. These theoretical foundations with an RUM have been presented in more detail in Grogger and Hanson (2011), Mayda (2010) and others. They adopted the standard multinomial logit model, assuming that the error terms follow an independent and identically distributed extreme value type 1 distribution (McFadden, 1974). In that case, the Independence of Irrelevant Alternatives (IIA) assumption needs to hold, i.e. the ratio of the probabilities of two options does not depend on a third option. In Chapter 3.6.2, we relax the IIA assumption and control for multilateral resistance to migration. This concept was introduced by Bertoli and Fernández-Huertas Moraga (2013) on the basis of the generalized nested logit model (Wen and Koppelman, 2001). It describes a very general form of an RUM which allows for correlation of the error terms across alternative destinations and thus controls for changes of attractiveness in alternative destinations that affect bilateral migration flows.

Based on this theoretical micro-foundation of migration decisions, we can estimate a pseudo-gravity model of migration flows to Germany with the following fixed-effects model specification:

$$y_{jt} = \alpha' GI_{jt} + \boldsymbol{\beta}' \boldsymbol{x_{jt}} + \boldsymbol{\phi'_t} \boldsymbol{d_t} + \boldsymbol{\phi'_j} \boldsymbol{d_j} + \boldsymbol{\phi'_j} \boldsymbol{d_{jT}} + \eta_{jt} \qquad (3.1)$$

where y_{jt} represents the logarithm of the migration flow from origin country j to Germany in year t over the number of people that stay in origin country j. GI represents our main variable of interest, the number of (language) institutes. $\boldsymbol{x}_{jt}$ is the vector of control variables that includes the log of GDP/capita as a measure for economic conditions, a variable that indicates conflicts, dummies

that indicate limited and unlimited movement of labour to Germany, a variable that captures freedom rights, the log of the trade volume (imports and exports) with Germany, and the log of the stock of migrants of country j in Germany in $t-1$ as a measure for previous migration to Germany.[31] Furthermore, we include a set of dummies to control for fixed effects. We add year dummies d_t to control for origin-invariant effects and origin dummies d_j to control for all time-invariant characteristics of the origin country as well as relations between Germany and the origin country j. We further add dummies d_{jT} that vary by origin country j and 10-year time periods T. As our time frame covers 38 years, these fixed effects help to control for slowly changing factors in the relations between Germany and the origin country and hence help to reduce a potential omitted variable bias. η_{jt} is the error term. Standard errors are clustered on the country level. We weight observations by the population of the origin countries to ensure that each potential migrant receives the same weight in the estimations independent of the origin.

3.5 Results for migration flows to Germany

We estimate Equation (3.1) in several specifications. Our preferred specifications in Table 3.1 employ the number of language institutes, i.e. institutes that offer language services. In Table 3.2, robustness checks are presented including lagged specifications. Tables 3.3 and 3.4 and Figure 3.7 show that the sample choice does not influence the results. In Table 3.5, we study potentially heterogeneous effects. As we include origin-country fixed effects and origin-country*10-year fixed effects in all specifications, the estimated effects capture within country, 10-year variations.

3.5.1 Basic specifications

For the basic specifications presented in Table 3.1, we include the control variables in a stepwise way. The coefficients of GDP/capita are negative and significant in all specifications. Better economic conditions in the origin country are thus negatively related with the benefits of migration. Furthermore, both variables that control for the free movement of labour within the EU are signifi-

---

[31]   To assure that the logarithm is also defined if the migrant stock is zero, we always add one migrant to it.

Table 3.1: Estimation results: basic specifications

DV: Migration rate (log)	(1)	(2)	(3)	(4)
Language institutes	0.107*	0.114*	0.100*	0.115**
	(0.060)	(0.060)	(0.052)	(0.054)
GDP/capita (log)		-0.361*	-0.469***	-0.399***
		(0.186)	(0.164)	(0.143)
Free movement (limited)		0.791***	0.820***	0.717***
		(0.218)	(0.222)	(0.177)
Free movement		1.044***	1.119***	0.851***
		(0.385)	(0.379)	(0.294)
Conflict			0.016***	0.016***
			(0.005)	(0.005)
Freedom rights			0.056***	0.043***
			(0.015)	(0.014)
Trade (log)			0.104	0.0905
			(0.082)	(0.099)
Migrant stock/pop. (log), lag=1				0.387***
				(0.135)
Constant	-11.660	-8.849	-10.480	-7.014
	(178,843)	(128,218)	(142,899)	(75,879)
Observations	2,622	2,622	2,622	2,622
Number of countries	69	69	69	69
Years	1977-2014	1977-2014	1977-2014	1977-2014
Adjusted R-squared	0.611	0.623	0.641	0.668
Year FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes	Yes

Cluster-robust standard errors in parentheses. Observations are weighted by population size. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

cantly and positively correlated with migration to Germany. They capture less restrictive immigration regulations which lower migration costs. While violent conflicts and more freedom rights are positively associated with migration to Germany, trade with Germany is not significant in any specification. Finally, the stock of migrants from the origin country already present in Germany is positively and highly significantly associated with migration to Germany, which might capture network effects.

Turning to our main variable of interest, the association between the number of language institutes and the migration rate to Germany is positive and significant in all specifications. Referring to the specification in Column 4, which will be our preferred specification in the following, the coefficient of 0.115 can be interpreted as follows: one more language institute is associated with an average increase of the migration rate to Germany by $e^{0.115} - 1 \approx 12.2\%$.[32]

We perform two back-of-the-envelope calculations based on our preferred specification in order to assess the size of the effect and the economic significance from the perspective of the origin countries and the destination country. First, we take the destination perspective. Figure 3.4 compares the predicted number of migrants to Germany with the actual number of migrants to Germany for our sample and in total. We see that the time trend is similar in all three cases. Furthermore, we find that the predicted numbers and the actual numbers are very close for our sample.

In order to assess the economic significance of additional migration flows related to language institutes, Figure 3.5 compares the predicted number of migrants – again estimated with our preferred specification – to the predicted number of migrants for a hypothetical situation where the number of language institutes is set to zero for all countries and all years, while the other control variables remain unchanged. The lower dashed line shows the difference between the two predicted numbers presented by the upper two lines. We predict that, on average, there were 57,382 more migrants to Germany every year compared to a hypothetical situation without any language institutes. This corresponds to around 16.8 % of the total number of migrants to Germany from the countries in our sample.

---

[32] The results are robust, but slightly larger if observations are not weighted by population size of the origin country (see Table 3.A3 in the Appendix). This also holds, when we limit our period to 1977-2010, i.e. to the time before migration and integration were first mentioned in official documents of the GI (see Table 3.A4 in the Appendix).
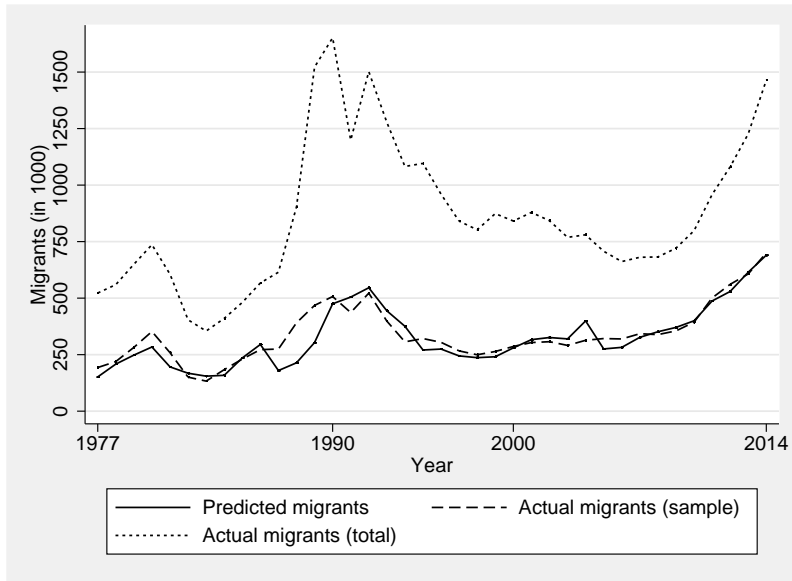
Figure 3.4: Actual and predicted migration flows to Germany, 1977-2014
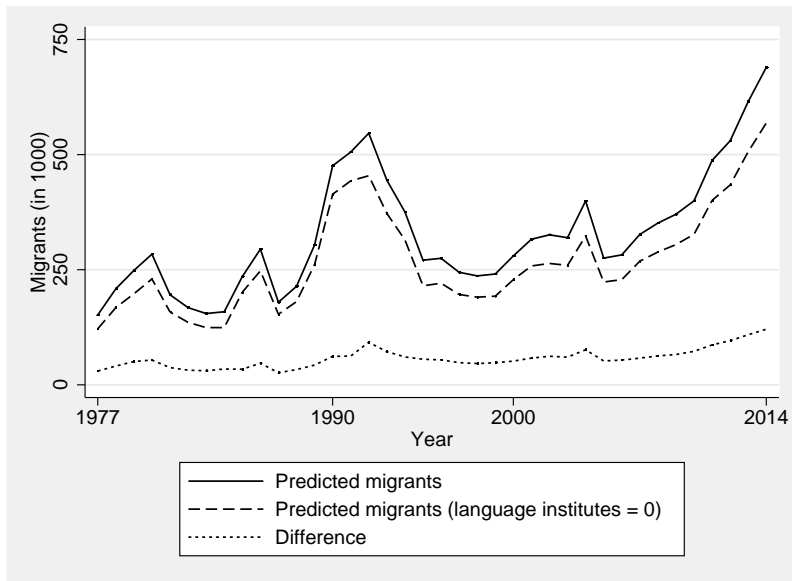
Figure 3.5: Predicted migration flows to Germany, with and without GI, 1977-2014

57
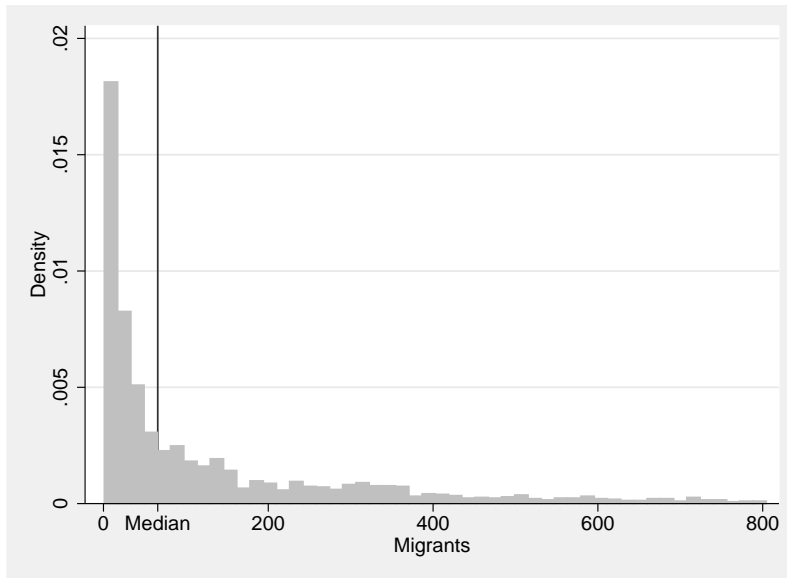
Figure 3.6: Predicted additional migrants, one additional institute (upper 10 % percentiles not shown)

Second, we take the perspective of the origin countries. We want to understand what the increase of the migration flows by 12.2 % related to one additional language institute means in actual numbers. Again, we compare the predicted number of migrants to Germany to a hypothetical situation where we add one language institute to all our country-year observations. Figure 3.6 shows the difference between those two predicted values, i.e. we estimate the size of the additional outflow associated with one additional language institute. The vertical line in the histogram represents the median at 65.86, i.e. in 50 % of the country-year observations an additional language institute is related to around 66 or fewer additional migrants. The distribution of the difference is strongly skewed to the right, with a mean of 590 additional migrants and a maximum of 11,539 additional migrants.

3.5.2 Robustness checks

In Table 3.2, we present robustness checks based on our preferred specification (Column 4 in Table 3.1). First, the positive correlation between the number of language institutes and the migration rate might not measure the effect of language-learning opportunities, but rather other aspects that come with the

Table 3.2: Robustness checks: independent variable choice

DV: Migration rate (log)	(1)	(2)	(3)	(4)	(5)	(6)
All institutes	0.059** (0.027)					
Institutes w/o lang. services		0.003 (0.021)				
Language institutes		0.115** (0.055)	0.082* (0.041)	0.055*** (0.019)		
Language institutes, lag=1			0.076** (0.036)	0.096* (0.051)	0.111** (0.045)	
Language institutes, lag=2				-0.001 (0.022)		
Lang. institutes/pop. (log)						0.049** (0.024)
Other controls	Yes	Yes	Yes	Yes	No	Yes
Other controls lagged	No	No	No	No	Yes	No
Observations	2,622	2,622	2,553	2,484	2,553	2,622
Number of countries	69	69	69	69	69	69
Years	1977-2014	1977-2014	1977-2014	1977-2014	1977-2014	1977-2014
Adjusted R-squared	0.665	0.668	0.661	0.671	0.654	0.664
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes

Cluster-robust standard errors in parentheses. Observations are weighted by population size. Other controls (log(GDP/capita), free movement (limited), conflict, freedom rights, log(trade), first lag of log(migrant stock/population)) and constant are included but not shown. *** p<0.01, ** p<0.05, * p<0.1.

opening of a new institute by the GI. Beyond language services, the institutes provide information about German culture and society. This might reduce uncertainty about life in Germany for potential migrants and therefore increase migration to Germany. The first two columns in Table 3.2 take these other aspects into account. In Column 1, we replace the number of language institutes in a given country and year with the number of all institutes, with and without language services. We find that the coefficient on the migration rate is reduced to 0.0585. In Column 2, we split the total number of institutes into two independent variables – the number of institutes with language services and the number of institutes without language services. We find that only additional language institutes are significantly associated with the migration rate, whereas institutes without language services are not. The coefficient for the number of language institutes is the same as in our preferred specification. From the results of these two robustness checks, we conclude that our variable of interest, when it comes to the empirical relation with migration to Germany, is indeed the variable which measures language-learning opportunities and not other effects of the GI.

Furthermore, we also include the first lag of the number of language institutes (Column 3), and the first and second lags (Column 4). Language learning and migration might not take place in the same period, as the acquisition of language skills requires some time. This is indeed what we find: a new institute is also strongly associated with the migration rate to Germany after one year, as we can see in Column 3, while for higher lags there are no further significant effects (see Column 4). In Column 5, we lag all control variables by one year, as individuals might react to conditions in the previous year. The results do not change compared to our preferred specification.

Additionally, it might be the case that institutes must be sufficiently large to be able to influence the migration rate and therefore drive our results. Unfortunately, we cannot measure the actual size of institutes in terms of course participation since there is no consistent measure available over this long period.[33] In Column 6, we therefore relate the number of institutes in a country to the population size. This controls for the possibility that, for example, one more institute in India and one more institute in Iceland might be differently associated with the respective migration rates. We find that the coefficient for the log of the number of language institutes per 1 million inhabitants remains positive and significant.

---

[33] For available data about language course participation, see Chapter 2

Table 3.3: Robustness checks: joint reporting of institutes and unbalanced sample

DV: Migration rate (log)	(1)	(2)	(3)
Language institutes	0.103**	0.090**	0.094*
	(0.043)	(0.041)	(0.052)
Other Controls	Yes	Yes	Yes
Observations	3,268	3,268	4,374
Number of countries	86	86	152
Years	1977-2014	1977-2014	1977-2014
Adjusted R-squared	0.655	0.654	0.645
Year FE	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes

Cluster-robust standard errors in parentheses. Observations are weighted by population size. Other controls (log(GDP/capita), free movement, free movement (limited), conflict, freedom rights, log(trade), first lag of log(migrant stock/population)) and constant are included but not shown. In Column 1, it is assumed that with joint reporting of language services activities for two or more institutes only the first-named institutes offered language services, and in Column 2, it is assumed, that in that case all institutes offered language services, first-named institutes and not-first-named institutes. Column 3 includes all countries with at least ten consecutive observations. *** p<0.01, ** p<0.05, * p<0.1.

3.5.2.1 Sample choice

Table 3.3 shows that the results do not depend on choices regarding the sample. In Columns 1 and 2, we include countries for which it is not always entirely clear for some institutes if they have offered language services or not, because of joint reporting in the annual reports. In those cases, annual reports show numbers on language course and exam participation jointly for a first-named institute (mostly main institute) followed by one or more institutes in parentheses (mainly subsidiaries). The independent variable in Column 1 assumes that only the first-named institutes offered language services. In Column 2, it is assumed that all institutes, including both first-named institutes and not-first-named institutes, offered language services. In these cases, the sample increases to 86 countries (see map in Figure 3.A2). In both columns, the coefficients are slightly smaller but do not change qualitatively.

Column 3 shows an unbalanced sample, including all years for all countries with at least ten consecutive observations. This enlarges the sample in two ways. First, we no longer drop all observations of a country if in some years joint reporting occurred. Second, additional countries are added to the sample which came into existence later than 1977, like successors of the Soviet Union or Yugoslavia. This sample then contains 153 countries (see map in Figure 3.A3). Again, we find that this does not change our results qualitatively, but the estimate is slightly less precise.

Finally, we check if the association is driven by individual countries. Therefore, we run our preferred specification 69 times, each time dropping one country. Table 3.4 shows the results for two countries – USA (Column 1) and India (Column 2) – which, when dropped, lead to a change of the coefficient of language institutes.[34] Running the estimation without the USA increases the size of the coefficient, while the coefficient becomes smaller and insignificant when India is dropped. These two countries are special in two ways. First, they are the two countries with the largest populations. As we weight our regressions by population size, the influence of India and the USA on our results is relatively large. We present unweighted specifications in Columns 3 and 4. When dropping the USA, the coefficient of language institutes is no longer different from the coefficient for the whole sample. The same holds when dropping India. Second, India and the USA are outliers with respect to the maximum number of language institutes. While 66 countries have at most

---

[34] Results do not change for all other countries and are available upon request.

three language institutes, there are three countries that have more – two of them are India with a maximum of eight and the USA with a maximum of nine language institutes. In Column 5, both countries are dropped, and again, the coefficient is no longer different from the whole sample. The third country with an above-average number of institutes is South Korea with a maximum of five language institutes. In Column 6, South Korea is additionally dropped. The size of the coefficient becomes larger. This shows that the effect is not driven by countries with a larger number of institutes.

Table 3.4: Robustness checks: special countries

DV: Migration rate (log)	(1) without USA	(2) without IND	(3) without USA	(4) without IND	(5) without IND, USA	(6) without IND, USA, KOR
Language institutes	0.255***	0.051	0.184**	0.128*	0.165**	0.211**
	(0.062)	(0.032)	(0.081)	(0.067)	(0.075)	(0.084)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,584	2,584	2,584	2,584	2,546	2,508
Number of countries	68	68	68	68	67	66
Years	1977-2014	1977-2014	1977-2014	1977-2014	1977-2014	1977-2014
Adjusted R-squared	0.681	0.646	0.686	0.686	0.654	0.654
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	No	No	Yes	Yes
Weighted	Yes	Yes	No	No	Yes	Yes

Cluster-robust standard errors in parentheses. Observations are weighted by population size, except in Columns 3 and 4. Other controls (log(GDP/capita), free movement (limited), conflict, freedom rights, log(trade), log(migrant stock/population)) and constant are included but not shown. *** p<0.01, ** p<0.05, * p<0.1.

3.5.2.2 Heterogenous effects

As our sample covers a long time period and a large number of countries with distinct characteristics, we test if there are heterogenous effects of the number of language institutes on migration to Germany for different groups of countries and how these effects change over time.

First, our results might be driven by different country groups due to different time-invariant characteristics, which are absorbed by our fixed effects. To better understand whether this is the case, we estimate our preferred specification (Column 4 of Table 3.1) with an additional interaction term for the GI variable and a dummy variable that captures geographic and linguistic distance, respectively. The dummy for geographical distance (Mayer and Zignago, 2011) takes the value one if the distance to Germany is larger than or equal to the median distance to Germany. In an analogous way, linguistic distance is captured by a dummy variable that takes the value one if the major language spoken in the country is a non-Germanic language, which holds for 60 countries in our sample (Adserà and Pytliková, 2015).

We also explicitly consider the possible interactions between the GI variable and economic distance. We construct a dummy that indicates for each year if the log GDP/capita is smaller than or equal to the median log GDP/capita.

Table 3.5 presents the results, where Column 1 shows the preferred specification for better comparison. We can see that economic distance matters, i.e. the association is significantly stronger for countries with lower income even though the association is still positive and significant for both groups of countries. This might be related to the different reasons for language learning: Given the costs of language acquisition, in countries with lower incomes, the human capital motive might be more important, i.e. language learning might be seen as an investment whose returns realize later – possibly due to migration. In higher income countries, on the contrary, language learning might be more often linked to a consumption motive and thus related to cultural reasons or a general interest in languages. We find similar results for geographic and linguistic distance, but the interaction term for geographic distance is only significant at the 10 % level. Both measures might indirectly capture a negative correlation between distance from Germany or the German language, respectively, and German language-learning opportunities outside the GI – as in geographically or linguistically distant countries– where the German language can be assumed to be overall of less importance. In those countries, the additional

Table 3.5: Heterogenous effects

| DV: Migration rate (log) | (1) | (2) Interacted with: Geographic distance | (3) Linguistic distance |
	Economic distance		
Language institutes	0.115**	0.063***	0.074**	0.051**
	(0.054)	(0.023)	(0.031)	(0.020)
Language institutes * interaction		0.274***	0.180*	0.219***
		(0.057)	(0.092)	(0.061)
Other Controls	Yes	Yes	Yes	Yes
Observations	2,622	2,622	2,622	2,622
Number of countries	69	69	69	69
Years	1977-2014	1977-2014	1977-2014	1977-2014
Adjusted R-squared	0.671	0.679	0.674	0.676
Year FE	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes

Cluster-robust standard errors in parentheses. Observations are weighted by population size. Eco Other controls (log(GDP/capita), free movement, free movement (limited), conflict, freedom rights, log(trade), first lag of log(migrant stock/population)) and constant are included but not shown. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

effect of language-learning opportunities provided by the GI on migration is larger, pointing toward diminishing marginal benefits of language-learning opportunities.

Second, the relation might change over time. Therefore, we estimate our preferred specification for sub-samples of 20 years. Figure 3.7 plots the coefficients for the number of language institutes and the 90 % confidence intervals for different 20-year sub-samples where the indicated year marks the first year of the respective 20-year period. While the relation is quite large in the sub-sample of 1977 (0.197), it decreases in the subsequent sub-samples until the sub-sample of 1981 (0.071). The relation in the following sub-samples stays roughly constant until the sub-samples starting at the end of the 1990s. There is a further, slight decrease until the last sub-sample (0.033). The coefficients are significant on the 10 % level in 18 of the 19 sub-samples; only in one sub-sample is there no significant relation. Overall, this points toward a smaller role of language-learning opportunities provided by the GI for migration in recent years, which can be related to a more global world and, among others, to the advent of online language learning platforms in the late 2000s.
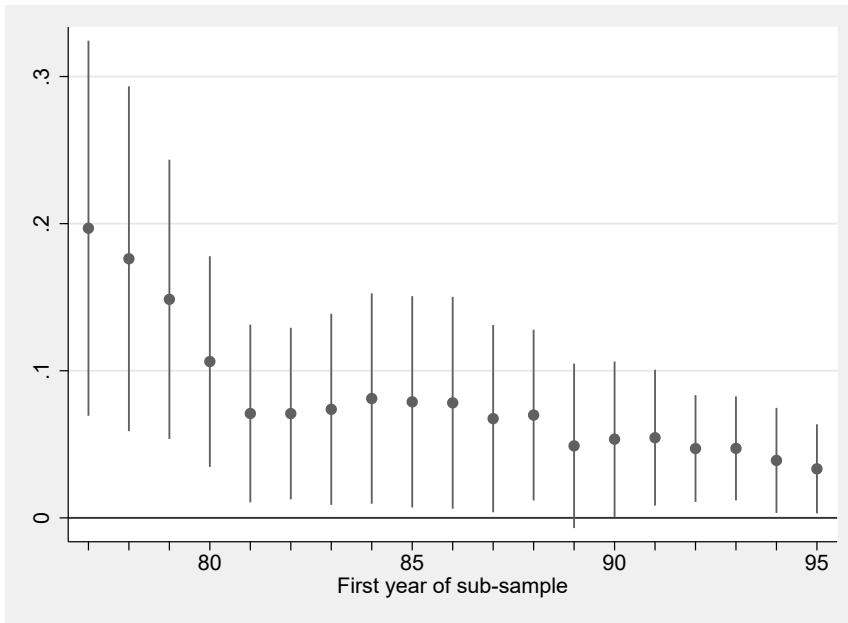
Figure 3.7: Coefficient plot for the number of language institutes: 20-years sub-samples (90 % confidence intervals)

3.6 Threats to exogeneity

In the previous section, we have shown that there is a robust positive relationship between language-learning opportunities and migration flows to Germany. In order to interpret that relation causally, we have to deal with two potential threats to exogeneity: reverse causality and multilateral resistance to migration.

3.6.1 Reverse Causality: The Case of Switzerland

We have found a positive relationship between the number of language institutes and the migration rate. However, this does not tell us the direction of causality. While we focus here on the direction of causality from language-learning opportunities to migration flows, we cannot rule out the opposite direction of causality, i.e. that the GI is more likely to open institutes in countries with larger migration to Germany. This makes it impossible to disentangle the positive correlation of institutes and the migration rate, as estimated in Chap-

ter 3.5, in the "migration effect" caused by the opening of an institute on the migration rate and the "selection effect" caused by the migration rate on the location decision for a new institute. If the latter effect were indeed relevant, we would overestimate the effect of the number of institutes on the migration rate, i.e. the migration effect would be biased upwards.

To address this issue, we follow two approaches: First, based on an analysis of strategies and overall objectives of the GI, we argue that there is evidence that the location decision is based on a large number of considerations, all of which are unrelated to migration (see Chapter 3.2). Second, we contrast our estimates for Germany with estimates for Switzerland. Even if we cannot fully rule out that the decision process for opening and closing institutes is affected by already existing migration flows to Germany, migration flows to Switzerland do not play any role in this process. The GI and the FFO neither consult with the Swiss government, nor are they influenced by Swiss representatives.[35] The number of institutes can therefore be considered exogenous to migration flows to Switzerland. At the same time, it is possible that language-learning opportunities affect migration flows to Switzerland, at least to the German-speaking part. Considering Switzerland thus allows us to isolate the "migration effect" from the "selection effect".

In more detail, we replace the dependent variable used so far by the migration rate to Switzerland. We distinguish between the German-speaking part and the non-German-speaking part of Switzerland.[36] By focusing on Switzerland, we can exclude the issue of reverse causality, as language institutes are exogenous to migration flows to Switzerland. This allows us to see if language is the driver behind the observed relation. We expect that language institutes have a positive and significant effect on migration flows to the German-speaking part but neither to the Italian- nor the French-speaking part.

We estimate Equation (3.1) with data for Switzerland. Migration flows to Switzerland are provided by Bundesamt für Statistik (2016, 2017). We use the same sample of countries as for Germany as destination country, but only for the period 1992 – 2014 due to data availability. To control for origin-specific characteristics that are not bilaterally related to Switzerland, we use the same variables as for the German specification, i.e. GDP/capita, population, conflict

---

[35] The GI cooperates with the Swiss and Austrian embassies in some host countries once institutes have been opened

[36] We define a canton as German-speaking if the main language for the majority was German in 2010 according to the Bundesamt für Statistik (2012) and then aggregate the data for German- and non-German-speaking cantons.

and war. Furthermore, we construct a variable for the free movement of persons between the EU/EFTA (Liechtenstein, Iceland and Norway) and Switzerland (Staatssekretariat für Migration, 2017). The variable takes the value one if there is some relaxation of the immigration rules and zero otherwise. These relaxations include different steps toward free movement, like quotas, national worker priority or an invocation of the safeguard clause. Data on the migrant stocks is based on migration flow data. Exports and imports to Switzerland are also included but are not disaggregated for the German-speaking and non-German-speaking parts of Switzerland (Barbieri and Keshk, 2016; Barbieri et al., 2009).

Table 3.6 shows the results. Columns 1-3 present the results for the German-speaking part of Switzerland and Columns 4-6 present the results for the non-German-speaking part of Switzerland. In Columns 1 and 4, we re-run our preferred specification for the shorter period from 1992 - 2014 for migration to the German-speaking and non-German-speaking parts of Switzerland, respectively. In both specifications, language institutes have a positive and significant effect on migration flows; however, the size of the coefficient is twice as large for the German-speaking part (0.100) than for the non-German-speaking part (0.047).

There is one further concern, however: migration flows to Germany and to Switzerland could be correlated. The decision to open or close a language institute by FFO and GI would then no longer be exogenous to migration to Switzerland. Therefore, we include as an additional control the log migration rate from the origin country to Germany in Columns 2 and 5. The coefficient for language institutes remains unchanged for both parts of Switzerland, while the migration rate to Germany is significantly correlated with migration to the non-German-speaking part of Switzerland only. In order to further address the issue of potential correlation of migration flows, we re-run the estimations for Switzerland but exclude countries from the sample with significantly (at least on the 10 %-level) and positively related migration flows to Germany and to the German-speaking part of Switzerland with a variation in the number

of language institutes in the period 1992–2014. The results are presented in Columns 3 and 6.[37]

Column 3 shows that the size of the effect of language institutes on the migration rate to the German-speaking part increases for this reduced sample. Overall, the correlation between migration flows to Germany and to the German-speaking part of Switzerland, which can be observed for some countries, does not affect our main results in a qualitative way. For the non-German-speaking part of Switzerland, Column 6 shows that the coefficient for language institutes turns insignificant and so does the coefficient for migration to Germany. The results for the reduced sample for the non-German-speaking part of Switzerland can also be interpreted as a placebo-test for the treatment of language institutes.[38]

As we exclude any direct effect of migration flows to Switzerland on the opening and closing decision for institutes, and as there is also no evidence of an indirect effect once we exclude the countries with correlated migration flows to the German-speaking part of Switzerland and to Germany, we interpret the results for the German-speaking part of Switzerland as evidence of a causal effect of the presence of language-learning opportunities abroad on migration from those countries to Switzerland. We take these results as support of a causal interpretation of the relation of language-learning opportunities abroad and migration as well for Germany as a destination country.

---

[37] This is based on country-specific regressions of the log of the migration rate to the German-speaking part of Switzerland on the log of the migration rate to Germany. The following 11 countries are excluded: Costa Rica, Denmark, Finland, India, Iraq, Mexico, Norway, Poland, Saudi Arabia, USA and Vietnam. For robustness, further specifications are run. See Tables 3.B1 and 3.B2 in the Appendix. This relation is negative only for three countries. As none of those countries has a variation in the number of language institutes, dropping them from the regression does not alter the results.

[38] Eugster et al. (2017) show that the French- and Italian-speaking parts of Switzerland are indeed culturally different from the German-speaking part. Therefore, it can be seen as a "different" country and can be used as a placebo-test.

Table 3.6: Estimation results: Switzerland

DV: log migrate	(1)	(2) CHE (German)	(3)	(4)	(5) CHE (non-German)	(6)
Language institutes	0.100***	0.098**	0.135**	0.047**	0.044**	0.034
	(0.037)	(0.038)	(0.067)	(0.020)	(0.021)	(0.037)
Migration rate Germany (log)		0.126	0.000		0.132**	0.082
		(0.084)	(0.084)		(0.053)	(0.053)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,587	1,587	1,334	1,587	1,587	1,334
Number of countries	69	69	58	69	69	58
Years	1992-2014	1992-2014	1992-2014	1992-2014	1992-2014	1992-2014
Adjusted R-squared	0.651	0.655	0.547	0.652	0.659	0.578
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes

Cluster-robust standard errors in parentheses. Observations are weighted by population size. Other controls (log(GDP/capita), EU/EFTA member, conflict, freedom rights, log(trade), first lag of log(migrant stock/population)) and constant are included but not shown. In (3) and (6), countries are excluded with significantly (at least on the 10%-level) and positively related migration flows to Germany and to the German-speaking part of Switzerland that have a variation in the number of language institutes in the period 1992 – 2014. *** p<0.01, ** p<0.05, * p<0.1.

3.6.2 Multilateral resistance to migration

As described in Chapter 3.4, if the IIA fails, the relative probabilities between two options, i.e. between the origin and one destination country, depend on a third option, such as an alternative destination country which constitutes a substitute to the chosen destination country. Bertoli and Fernández-Huertas Moraga (2013) call this "multilateral resistance to migration".[39] If the characteristics of alternative destinations are correlated with the regressors and we do not control for this, our estimates are biased. Bertoli and Fernández-Huertas Moraga (2013) show that the CCE estimator by Pesaran (2006) consistently corrects for multilateral resistance to migration, even if we do not have data on alternative destinations, as "the pattern of correlation in the error term, not only across destinations but also across origins, contains information about the unobserved attractiveness of other destinations, and the related unobserved bilateral migration rates" (Bertoli and Fernández-Huertas Moraga, 2013, p. 85). For the CCE estimator, cross-sectional averages of all dependent and independent variables interacted with heterogeneous coefficients for all countries have to be included. For our multiple-origin and single-destination setting we estimate the following equation:

$$y_{jt} = \alpha' GI_{jt} + \boldsymbol{\beta}' \boldsymbol{x_{jt}} + \boldsymbol{\phi_t'} \boldsymbol{d_t} + \boldsymbol{\phi_j'} \boldsymbol{d_j} + \boldsymbol{\phi_j'} \boldsymbol{d_{jT}} + \boldsymbol{\lambda_j'} \boldsymbol{\tilde{z}_t} + \eta_{jt} \qquad (3.2)$$

with the weighted cross-sectional average defined as

$$\boldsymbol{\tilde{z}_t} = \frac{1}{\sum_j \omega_{jt}} \left(\sum_j \omega_{jt} y_{jt}, \sum_j \omega_{jt} x_{jt} \right)$$

where ω_{jt} gives the weight for country j in t, for which we use population size.

In Table 3.7, we introduce stepwise our control variables in Columns 1-4 as we did in Table 3.1. In particular, we are interested in a comparison of the CCE-estimates and the fixed-effects (FE) estimates from our preferred specification.

---

[39] Bertoli et al. (2011) and Bratu et al. (2020) give illustrative examples of how immigration policies in destination countries might create spill-overs for migration flows to alternative destinations. Beverelli and Orefice (2019) provide a complementary view. They focus on the effects on migration flows to a given destination country which stem from competition among immigrants from different origin countries. Reductions in the costs of migration from a specific origin country to the destination country due to a bilateral preferential policy negatively affect migrating from a given origin country to the same destination country.

While the FE estimates capture both, the direct effect of a language institute on migration flows to Germany and the indirect effect due to the changed attractiveness of alternative destinations, the CCE estimates allow isolating the direct relation. We conduct a test for multilateral resistance. The CCE-test is a F-test on the joint significance of all cross-sectional averages included in the regression. The p-value for that test is 0.00 in all specifications. Hence, we conclude that multilateral resistance exits in our setting. We will analyse the economic relevance of it in the following.

Before interpreting the changes in our main variable of interest, we take a look at the changes that occur to the control variables. The absolute size of the negative coefficient for GDP/capita increases if we control for multilateral resistance. It indicates a positive indirect effect beside the negative direct effect. This implies that GDP/capita in the country of origin is negatively correlated with GDP/capita in alternative destinations. Put differently, an increase of GDP/capita in the country of origin makes migration to Germany less attractive on its own (direct effect), and at the same time, a decrease of GDP/capita in alternative destinations reduces their attractiveness, thereby increasing the attractiveness of Germany (indirect effect).[40]

While for limited free movement the difference in the coefficients between the CCE- and FE-estimates are not very large, we find that our FE-estimates are upward biased for the unlimited movement, pointing toward a significant and positive indirect effect. Unlimited freedom of movement to Germany thus goes hand in hand with a reduced attractiveness of alternative destinations. Bertoli et al. (2016) find the opposite for migration from EU countries to Germany, i.e. a negative indirect effect on migration to Germany. This captures a situation where unlimited freedom of movement to Germany is positively correlated with an increase in the attractiveness of alternative destinations and can be explained by the simultaneous removal of migration restrictions to those destinations, which are close substitutes to Germany, in their sample. In our sample, fewer European, and in particular, fewer Eastern-European countries are included. As free movement to Germany seems to go hand in hand with a reduced attractiveness of alternative destinations, those alternative destinations are likely outside the EU for most origin countries.

---

[40] This result is different from the one found by Bertoli and Fernández-Huertas Moraga (2013) for Spain as destination country as far as the indirect effect is concerned. For the sample of origin countries considered, the indirect effect is negative which indicates that GDP/capita in the origin country is positively correlated with GDP/capita in alternative destinations pointing at economically closer countries than in our sample.

In a similar manner, we can now interpret the change in the size of the coefficient of the number of language institutes. Controlling for multilateral resistance of migration, the point estimates reduce by half to 0.0576 and indicate an upward biased estimate with our FE specification. This means that roughly half of the total effect constitutes a direct effect on migration from the origin country to Germany. The other half is an indirect effect of redirected migration flows to Germany due to a relative decrease of attractiveness of alternative destinations. With the opening of a language institute, migration to Germany becomes more attractive for two groups of people. The first group are those who would not have migrated otherwise. This is the direct effect which we find when controlling for multilateral resistance. The second group are those who would have left the origin country even in the absence of a language institute but to another destination country.

Summing up, the CCE-estimates helps to identify the direct effect and to disentangle it from the indirect effect. In a broader sense, this allows the assessment of the relevance of new migration flows induced by the GI which is linked to the direct effect, as compared to the redirection of existing migration flows which is captured by the indirect effect. To determine the role which the GI could play in fostering migration and integration (see Chapter 3.2), both effects together, as derived with the FE-estimates, are of main economic and political importance.

Table 3.7: Estimation results - common correlated effects

DV: Migration rate (log)	(1) CCE	(2) CCE	(3) CCE	(4) CCE	(5) FE
Language institutes	0.088***	0.110***	0.056***	0.058***	0.115**
	(0.020)	(0.020)	(0.020)	(0.020)	(0.054)
GDP/capita (log)		-0.557***	-0.610***	-0.621***	-0.399***
		(0.049)	(0.056)	(0.057)	(0.143)
Free movement (limited)		0.689***	0.633***	0.631***	0.717***
		(0.267)	(0.230)	(0.231)	(0.177)
Free movement		0.921**	0.478	0.446	0.851***
		(0.443)	(0.352)	(0.362)	(0.294)
Conflict			0.010***	0.012***	0.016***
			(0.003)	(0.003)	(0.005)
Freedom rights			0.011	0.014*	0.043***
			(0.007	(0.008)	(0.014)
Trade (log)			0.105***	0.098***	0.091
			(0.028)	(0.028)	(0.099)
Migrant stock/pop. (log), lag=1				-0.001	0.387***
				(0.030)	(0.135)
Constant	-11.070	18.060	14.160	14.080	-7.014
	(7.924)	(25.810)	(33.640)	(33.720)	(75,879)
Observations	2,622	2,622	2,622	2,622	2,622
Number of countries	69	69	69	69	69
Years	1977-2014	1977-2014	1977-2014	1977-2014	1977-2014
Adjusted R-squared	0.959	0.965	0.979	0.979	0.668
Country FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes	Yes	Yes
Countries	69	69	69	69	69
Years	1977-2014	1977-2014	1977-2014	1977-2014	1977-2014
CCE-test (p-value)	0.00	0.00	0.00	0.00	

Observations are weighted by population size; CCE-results are estimated with the CCE-estimator (Pesaran 2006). *** p<0.01, ** p<0.05, * p<0.1.

3.7 Conclusion

In this chapter, we have analysed how the presence of German language-learning opportunities abroad affects migration to Germany. We find a significant and positive correlation between the number of language institutes of the GI and migration rates to Germany and show that the language-learning opportunities are indeed the relevant channel. This relationship is stronger for countries with lower income and where the linguistic and geographic distance is larger.

We further demonstrate that the identified relation can also be found for migration to the German-speaking part of Switzerland. While migration to Germany might be endogenous to the opening and closing of institutes, we argue that this is not the case for migration to Switzerland. This excludes the possibility of reverse causality and thus allows us to interpret the relationship between language-learning opportunities and migration as causal.

So far, similar effects have only been shown for foreign language learning at school for EU countries (Aparicio Fenoll and Kuehn, 2016). However, compulsory language learning is not within the reach of policy-markers in the destination country, much less are linguistic properties. This is different for language-learning opportunities for adults. In particular, cultural institutes like the GI are often part of the foreign cultural policy of potential destination countries. We have shown that the number of migrants can be affected by the supply of language-learning opportunities for adults. There is also evidence that language-learning opportunities abroad lead to positive (self-)selection of migrants related to education and integration outcomes after arrival (Jaschke and Keita, 2021).

Hence, we provide support for the strategy of the German government, as visible in the new Skilled Immigration Act effective since March 2020. This act aims at facilitating migration of skilled workers from third countries to Germany in order to cope with the skill shortage stemming from the aging of the population. As far as language proficiency of the migrants is concerned, an active role is assigned to the GI (Bundesregierung, 2018; Goethe-Institut, 2018). Our analysis not only provides information about the total effect of language-learning opportunities on migration to German, but it also disentangles the direct and indirect effects. While the direct effect captures additional outmigration from origin countries due to to additional language institutes, the indirect effect shows the redirection of already existing outmigration. This dis-

tinction is important for the relation of Germany with the origin countries. It facilitates assessment of the total effect against the background of origin countries' potential concerns of brain drain and at the same time helps Germany to better understand its own position in the global competition for talent.

Appendices to Chapter 3

3.A   Appendix A

Table 3.A1: Countries in the sample

Algeria	Haiti*	Paraguay*
Benin*	Honduras*	Peru
Bolivia	Hungary	Philippines
Bulgaria	Iceland*	Poland
Burkina Faso*	India	Romania
Burundi*	Iran	Rwanda*
Cambodia*	Iraq	Saudi Arabia
Cameroon	Ireland	Senegal
Chad*	Jordan	Sierra Leone*
Chile	Kenya	South Africa
Colombia	Liberia*	South Korea
Congo - Brazzaville	Madagascar*	Sri Lanka
Costa Rica	Malaysia	Sweden
Côte d'Ivoire	Mali*	Tanzania
Denmark	Mexico	Thailand
Dominican Republic*	Nepal	Togo
Ecuador*	New Zealand	Trinidad & Tobago*
El Salvador*	Nicaragua*	Tunisia
Ethiopia	Niger*	Turkey
Finland	Nigeria	United States
Ghana	Norway	Uruguay
Guatemala*	Pakistan	Venezuela
Guinea*	Panama*	Vietnam

* Countries which never had a language institute between 1977 and 2014

Table 3.A2: Summary statistics

Variable	Obs		Mean	Std. Dev.	Min	Max
Migration rate	2622	Overall	.0002317	.0007254	5.48e-07	.0111187
All institutes	2622	Overall	.8954996	1.398392	0	12
		Between		1.323684	0	8.447368
		Within		.477595	-3.551869	4.790236
Language institutes	2622	Overall	.8066362	1.158708	0	9
		Between		1.103029	0	6.5
		Within		.3782926	-1.798627	4.701373
Language institutes/population (in 1m)	2622	Overall	.0607665	.1070998	0	.6315869
Population (in 1m)	2622	Overall	39.01496	122.3834	.222142	1295.292
GDP/capita	2622	Overall	8904.741	11076.02	142.3924	74113.96
Free movement (limited)	2622	Overall	.0099161	.0991035	0	1
Free movement	2622	Overall	.0488177	.2155278	0	1
Trade (in 1m EUR)	2622	Overall	3898.953	11782.78	.075	145134.3
Migrant stock / population	2622	Overall	.0012884	.0038552	7.91e-06	.0348946
Conflict	2622	Overall	2.857742	7.701755	0	50
Freedom rights	2622	Overall	6.321129	3.817549	0	12

Table 3.A3: Estimation results: basic specification unweighted

DV: Migration rate (log)	(1)	(2)	(3)	(4)
Language institutes	0.128*	0.146*	0.143*	0.137**
	(0.073)	(0.076)	(0.075)	(0.067)
GDP/capita (log)		-0.525**	-0.577**	-0.405***
		(0.215)	(0.235)	(0.150)
Free movement (limited)		1.050***	1.048***	0.840***
		(0.231)	(0.237)	(0.169)
Free movement		1.517***	1.518***	0.970***
		(0.274)	(0.280)	(0.200)
Conflict			0.009*	0.009*
			(0.005)	(0.005)
Freedom rights			0.005	-0.007
			(0.024)	(0.020)
Trade (log)			0.061	0.060
			(0.056)	(0.066)
Migrant stock/pop. (log), lag=1				0.554***
				(0.133)
Constant	-10.660***	-6.384***	-7.165***	-3.467*
	(0.077)	(1.775)	(1.790)	(1.906)
Observations	2,622	2,622	2,622	2,622
Number of countries	69	69	69	69
Years	1977-2014	1977-2014	1977-2014	1977-2014
Adjusted R-squared	0.619	0.639	0.641	0.685
Year FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Country*10-year FE	69	69	69	69

Cluster-robust standard errors in parentheses. *** p¡0.01, ** p¡0.05, * p¡0.1.

Table 3.A4: Estimation results: basic specifications 1977-2010

DV: Migration rate (log)	(1)	(2)	(3)	(4)
Language institutes	0.125*	0.135*	0.119*	0.131**
	(0.067)	(0.069)	(0.060)	(0.062)
GDP/capita (log)		-0.430**	-0.540***	-0.460***
		(0.168)	(0.141)	(0.130)
Free movement (limited)		0.642***	0.657***	0.612***
		(0.175)	(0.182)	(0.155)
Free movement		0.942***	1.235***	1.006***
		(0.118)	(0.159)	(0.168)
Conflict			0.012**	0.014**
			(0.006)	(0.006)
Freedom rights			0.055***	0.044***
			(0.014)	(0.014)
Trade (log)			0.132	0.119
			(0.088)	(0.104)
Migrant stock/pop. (log), lag=1				0.348***
				(0.129)
Constant	-10.700	-7.062	-9.276	-6.518
	(672,274)	(707,586)	(659,790)	(541,189)
Observations	2,346	2,346	2,346	2,346
Number of countries	69	69	69	69
Years	1977-2010	1977-2010	1977-2010	1977-2010
Adjusted R-squared	0.573	0.589	0.608	0.633
Year FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes	Yes

Cluster-robust standard errors in parentheses. Observations are weighted by population size. *** p¡0.01, ** p¡0.05, * p¡0.1.
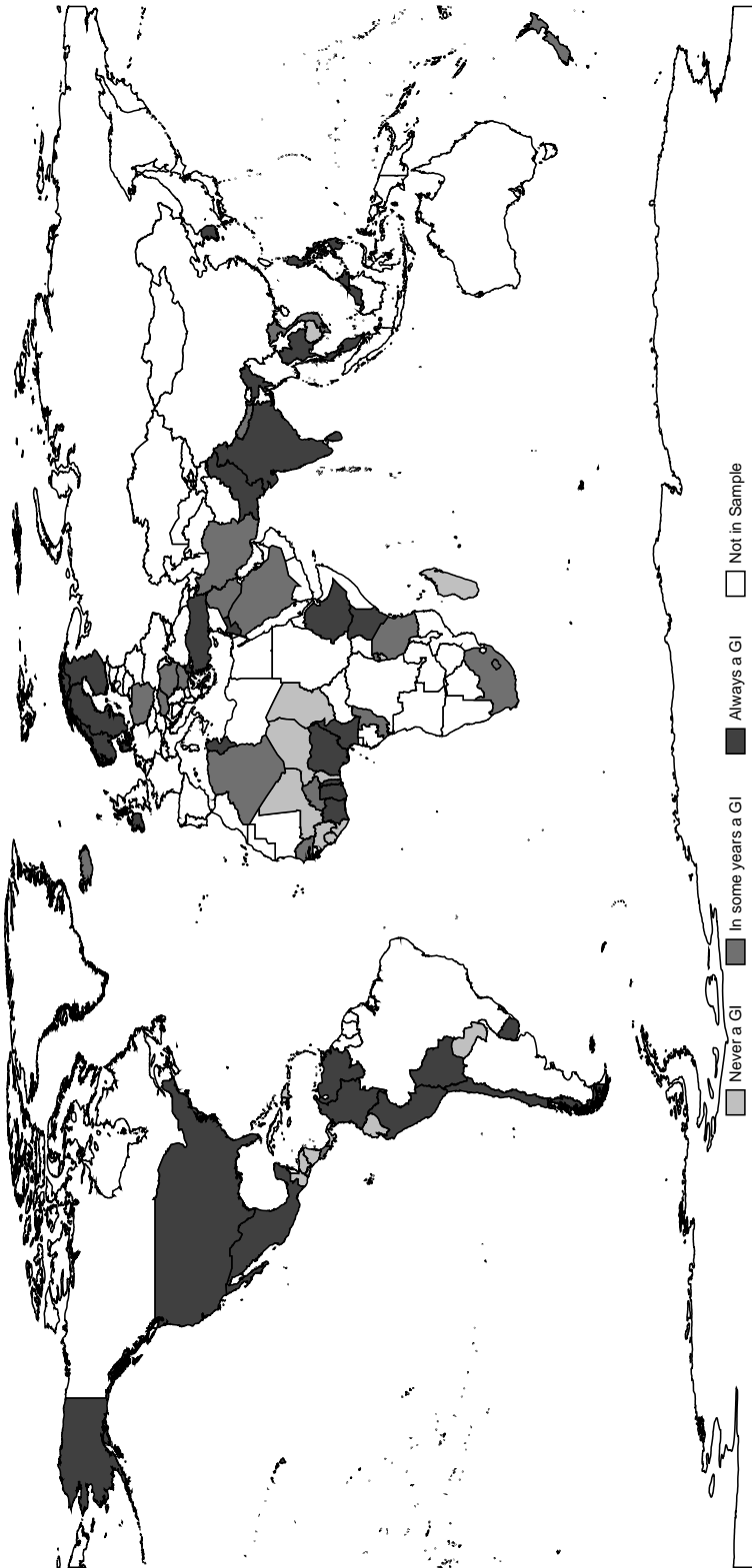
Figure 3.A1: The presence of the GI.

Never a GI | In some years a GI | Always a GI | Not in Sample

82
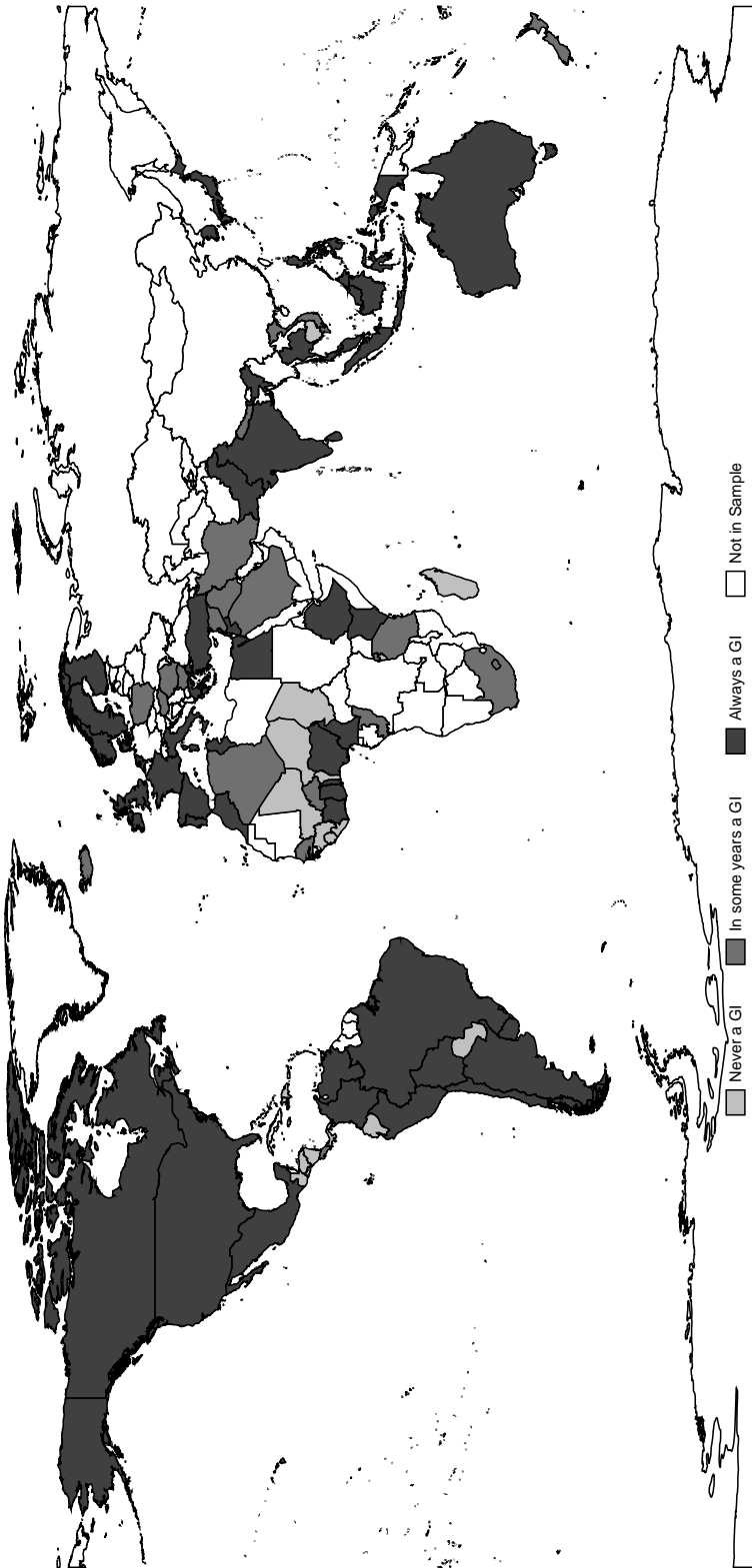
Figure 3.A2: The presence of the GI - sample with jointly reported institutes.
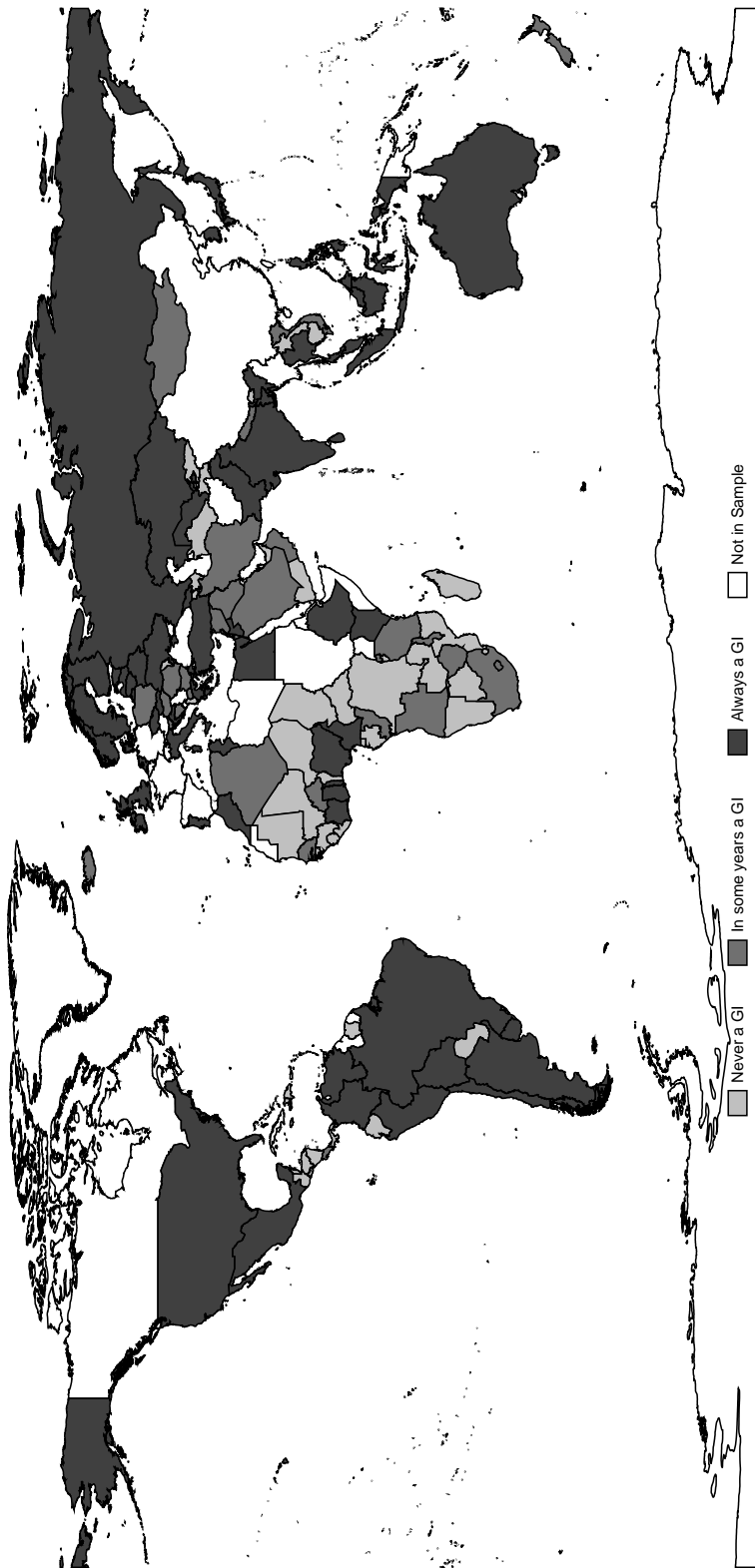
Figure 3.A3: The presence of the GI - unbalanced sample.

Never a GI   In some years a GI   Always a GI   Not in Sample

3.B   Appendix B

With Tables 3.B1 and 3.B2, we want to show that the results of Table 3.6, Columns 3 and 6 are robust to alternative ways of excluding countries with correlated migration flows to the German-speaking part of Switzerland and to Germany. Based on country-specific regressions of the log of the migration rate to the German-speaking part of Switzerland on the log of the migration rate to Germany, we determine the correlation based on the coefficient of the log of the migration rate to Germany. We focus on those countries with at least one language institute in the period 1992-2014. We consider eight different groups of correlated countries based on combinations of the following three dimensions:

- Only those countries with a variation in the number of institutes (V)/ those with and without a variation in the number of institutes (all, A)

- Only positive correlation (P)/ positive and negative correlation (PN)

- Correlation significant on the 10% level (10) / significant on the 5% level (5)

Table 3.B1 presents results for the German-speaking part of Switzerland and Table 3.B2 for the non-German-speaking part of Switzerland. Similar to Table 3.6, Columns 3 and 6, which show results for the group $V, P, 10$, the results are significant at least on the 10% level for the German-speaking part and insignificant for the non-German-speaking part. If we exclude only those countries with a variation in the number of institutes (V), we find that the correlation is always positive for those countries, i.e. there are no countries belonging to the groups $V, PN, 10$ and $V, PN, 5$. Therefore, we do not show these specifications, as they coincide with the specifications in Table 3.B1, Column 5, Table 3.B2, Column 5, and Table 3.6, Columns 3 and 6.

Table 3.B1: German-speaking part of Switzerland: Uncorrelated samples

DV: Migration rate (log)	(1)	(2)	(3)	(4)	(5)
			CHE (German)		
Language institutes	0.112*	0.114*	0.144**	0.145**	0.138**
	(0.065)	(0.064)	(0.072)	(0.071)	(0.067)
Migration rate Germany (log)	-0.055	-0.052	-0.089	-0.088	0.003
	(0.102)	(0.100)	(0.090)	(0.089)	(0.083)
Other controls	Yes	Yes	Yes	Yes	Yes
Observations	1,081	1,173	1,150	1,219	1,380
Number of countries	47	51	50	53	60
Years	1992-2014	1992-2014	1992-2014	1992-2014	1992-2014
Adjusted R-squared	0.467	0.467	0.506	0.505	0.545
Year FE	Yes	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Excluded correlated countries:					
Countries with var. in language inst.	A	A	A	A	V
Direction correlation	PN	PN	P	P	P
P-value ¡	10	5	10	5	10

Cluster-robust standard errors in parentheses. Observations are weighted by population size. Other controls (log(GDP/capita), EU/EFTA member, conflict, freedom rights, log(trade), first lag of log(migrant stock/population)) and constant are included but not shown. *** p¡0.01, ** p¡0.05, * p¡0.1.

Countries with var. in language inst.: Only those countries with a variation in the number of institutes (*V*)/ those with and without a variation in the number of institutes (all, *A*); *Direction correlation*: Only positive correlation (*P*)/ positive and negative correlation (*PN*); *P-value*:Correlation significant on the 10% level (10) / significant on the 5% level (5)

Table 3.B2: Non-German-speaking part of Switzerland: Uncorrelated samples

DV: Migration rate (log)	(1)	(2)	(3)	(4)	(5)
			CHE (Non-German)		
Language institutes	0.049	0.049	0.048	0.048	0.036
	(0.036)	(0.035)	(0.034)	(0.033)	(0.036)
Migration rate Germany (log)	0.042	0.044	0.045	0.046	0.083
	(0.063)	(0.062)	(0.055)	(0.054)	(0.053)
Other controls	Yes	Yes	Yes	Yes	Yes
Observations	1,081	1,173	1,150	1,219	1,380
Number of countries	47	51	50	53	60
Years	1992-2014	1992-2014	1992-2014	1992-2014	1992-2014
Adjusted R-squared	0.579	0.579	0.562	0.563	0.577
Year FE	Yes	Yes	Yes	Yes	Yes
Country*10-year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Excluded correlated countries:					
Countries with var. in language inst.	A	A	A	A	V
Direction of correlation	PN	PN	P	P	P
P-value ¡	10	5	10	5	10

Cluster-robust standard errors in parentheses. Observations are weighted by population size. Other controls (log(GDP/capita), EU/EFTA member, conflict, freedom rights, log(trade), first lag of log(migrant stock/population)) and constant are included but not shown. *** p¡0.01, ** p¡0.05, * p¡0.1.

Countries with var. in language inst.: Only those countries with a variation in the number of institutes (V)/ those with and without a variation in the number of institutes (all, A); *Direction correlation*: Only positive correlation (P)/ positive and negative correlation (PN); *P-value*:Correlation significant on the 10% level (10) / significant on the 5% level (5)

Chapter 4

Language-learning opportunities and exports

Abstract.[41] Using data on the presence of the Goethe Institutes (GI) in 134 importer countries between 1978 and 2014, we study the effect that language learning opportunities abroad have on German exports. We employ a gravity model of trade with a single exporter and use the Poisson-Pseudo Maximum Likelihood (PPML) estimator to measure the relationship of interest. To gauge the importance of potential reverse causality, we also estimate the effect that institutes have on Swiss exports. Our findings for both Germany and German-speaking Swiss cantons show that institutes do stimulate exports to GI-hosting countries but that this effect is confined to institutes offering language training services. This finding suggests that language requirements and acquisition underlie the positive link found between institutes and exports. This reading of our findings receives further support in additional explorations, where we study exports differentiated by Rauch (1999) product categories to account for differing communication requirements in trading.

4.1 Introduction

Language skills may be an important prerequisite for and driver of international trade. By reducing transaction costs that otherwise impede trade, language skills may make trade viable, or at least more profitable. Hence, the exchange of goods and services between countries may become intensified. Language-related transaction costs[42], of course, are far from uniform across goods and trading partners. They are particularly high for economic activities and exchanges that require extensive written and verbal communication (Selmier and Oh, 2013). They can also be prohibitively high, in which case trade will not be viable at all, and otherwise mutually beneficial transactions are not realized (Fidrmuc and Fidrmuc, 2016).

Although language commonality appears to be a rather simple measure to relate language skills to trade costs, its performance is rather impressive as Head and Mayer (2014) show in a meta-analysis. A more fine-grained way of including language skills in gravity models are measures of linguistic distance. Different indices of linguistic distance have been utilized in empirical research, which shows that such measures are generally negatively related to trade flows (Isphording and Otten, 2013b; Melitz and Toubal, 2014). Indices of linguistic distance, however, lack variability over time and also ignore potential foreign language skill acquisition, which might help to overcome trade impediments arising from linguistic distance. While Egger and Toubal (2016) solve the latter problem by estimating the effect of common acquired language on bilateral trade, their measure still is invariant over time. Exploiting the fall of the iron curtain as a natural experiment, Fidrmuc and Fidrmuc (2016) produce evidence that foreign language skills (here English) affect trade flows.

These pieces of evidence support the view that language and language acquisition matter for trade flows. Policies that increase domestic language skills abroad hence offer a potential tool for countries to promote their exports. In this chapter, we explore and test this possibility of export promotion. Specifically, using data on the openings and closings of institutes provided by the German Goethe-Institut (GI) in different countries (see Chapter 2), we study the effect that language-learning opportunities (LLOs) provided by such institutes have on German exports to countries hosting these institutes. The GI is

---

[42] The importance of transaction costs for economic activity and market exchanges more generally was first recognized and studied by Coase (1937) in the 1930s (Williamson, 1985).

a German cultural institute that offers language services, information on the German culture and hosts cultural events in countries all over the world.[43] Using a single-exporter gravity model of trade, we estimate the effect that the number of such institutes in a country has on German exports to that country. Our main sample contains annual observations on 134 countries for the period 1978 to 2014. Our findings show a significant positive but diminishing marginal impact of institutes on trade flows to a country, i.e. a non-linear effect of institutes on German exports.These results are corroborated in a battery of robustness checks.

Closest to our research are Ghosh et al. (2017), Lien and Lo (2017), and Lien et al. (2012, 2019) who show a positive relationship between the establishment of different cultural institutes (GI for Germany, Cervantes Institute for Spain, Confucius Institute for China, British Council for the United Kingdom) and trade flows. We can add two contributions to that particular strand of literature.

First, we explore and provide evidence for a particular causal pathway of this effect and also address the issue of reverse causality that may affect empirical studies in this area, as the openings and closings of cultural institutes might not be exogenous to export flows from the country that runs such institutes. We deal with this problem by estimating the effect of German institutes on Swiss exports. The German government decides jointly with the GI but not jointly with the Swiss government or any other Swiss institution, where to establish institutes. Moreover, parties participating in the decision do not take Swiss exports into account. The decision to operate an institute is hence exogenous to Swiss export flows. We are also able to distinguish Swiss exports from German-speaking cantons and from non-German-speaking cantons. Using this information, we provide evidence that institutes stimulate exports from the German-speaking part of Switzerland but not from the non-German-speaking part of Switzerland to countries hosting these institutes.

Second, while Ghosh et al. (2017), Lien and Lo (2017), and Lien et al. (2012, 2019) see cultural institutes as a way of increasing soft power (Nye, 2004), we further investigate language as a channel which facilitates communication and therefore reduces transaction cost. The results for Switzerland already provide some suggestive evidence that language might function as a mechanism facili-

---

[43] In this chapter, we maintain the following convention: when referring to the association of the Goethe-Institut we use the abbreviation "GI". When talking about specific branches of the GI abroad, we refer to them as "institutes".

tating increased trade flows. Additionally, by distinguishing between institutes that offer language services and those that do not, we show that language services are the driver of the results by distinguishing between institutes that offer language services, and those that do not. Furthermore, we use Rauch's (1999) product categories in the categorization of our export flow data to account for the different need for communication and different search costs when trading certain products. While homogeneous goods are traded on an organized exchange where only little communication is needed, differentiated goods do not have reference prices and a certain degree of communication between the trading entities is required. We find that there is no significant association between the number of institutes and the export of homogeneous goods, while the export of differentiated goods increases with the number of institutes present in the country. We conclude that language is the relevant channel.

Concerning the literature on language and trade, our findings provide further and more general evidence for the importance of acquired language skills for international trade flows. This is of potential policy relevance, as acquired language skills (but not the sharing of a common language or linguistic distance) can be targeted and manipulated by policy as a means of export promotion. An investigation of heterogeneous effects shows that trade-promoting effects are largest when initial trade barriers are high.

The chapter is organized in the following way. Section 4.2 gives background information on the literature on language in the context of international trade and on the GI. Section 4.3 describes the data and outlines our estimation strategy. In Section 4.4, we show our results for German and Swiss exports, and we demonstrate that language is the driver of the results by distinguishing exports according to the Rauch (1999) product categories. Section 4.5 concludes.

4.2 Background

4.2.1 Language and international trade

As trading is an act that heavily depends on written and verbal communication, language plays an integral part in facilitating economic interactions between trading partners (Selmier and Oh, 2013). Trade frictions caused by linguistic differences can be analysed in the transaction cost framework, since such frictions represent the costs of using the market (e.g., costs for contracting and

safeguarding)(Williamson, 1981). Apart from mere communication, language also functions as a carrier of cultural and social norms. Not only does this function of language potentially enhances trust between trade partners, it could also shape preferences for certain (cultural) goods (Disdier et al., 2010; Egger and Toubal, 2016). Language as a component of these transaction costs has been used in various applications of gravity models explaining international trade patterns. Most studies make use of binary indicators of language commonality, i.e. common official languages or common native languages, as well as measures for fractional spoken language overlap, such as common spoken languages (Egger and Toubal, 2016). In a meta-analysis, Egger and Lassmann (2012) collect coefficients for common languages (official or spoken) and provide an estimate which implies a 44% increase in trade flows between countries due to a common language.[44] Helpman et al. (2008), in turn, estimate a common language to increase the probability of bilateral trade between two countries by 10 %.

Deviating from the dichotomous common-language perspective, measures of linguistic proximity facilitate a more fine-grained analysis of the impact of communication difficulty on trade between countries. Popular measures for linguistic proximity consist of cardinal measures that make use of the level that languages share on a language family tree (Adserà and Pytliková, 2015; Guiso et al., 2009) or of continuous scales that express lexicographic and/or phonetic similarities, such as the Levenshtein distance as used by Isphording and Otten (2013b). A continuous measure of linguistic proximity allows for the calculation of elasticities in the context of the gravity model of trade. Isphording and Otten (2013b) find that a 1 percentage point increase in the Levenshtein distance leads to a significant decrease in trade by about 0.6%. However, these concepts are not free of criticism. A major disadvantage of these indices is the necessity of symmetric linguistic distances between languages which implies a similar difficulty in foreign language acquisition that goes both ways (Van der Slik, 2010).

The majority of the previously mentioned measures (apart from common spoken languages) that frequently enter gravity models of trade are time-invariant and often not policy-relevant variables. In reality, however, there are several ways to alleviate the language barrier between two trading partners: migration (Melitz and Toubal, 2014; Rauch and Trindade, 2002), automated translation (Brynjolfsson et al., 2019) or foreign language acquisition (Fidrmuc and Fidrmuc, 2016). Language acquisition is central to this chapter, as the GI offers

---

[44] Head and Mayer (2014) provide comparable effects in a similar analysis of trade costs.

German language courses worldwide. According to Ginsburgh et al. (2017), one of the driving forces behind learning a foreign language is trade with a country where the respective language is spoken. This evokes a severe problem of reverse causality, which Fidrmuc and Fidrmuc (2016) tackle by using the fall of the iron curtain as a natural experiment and exploiting differences in foreign language proficiency in Eastern Europe that have been exogenous to trade. For the analysis of the impact of language-learning opportunities abroad offered by the GI, another layer of this problem arises, since the opening of an institute could be also partially driven by trade with Germany. Studies on the impact of cultural institutes on foreign trade are still of quite recent vintage (Ghosh et al., 2017; Lien and Lo, 2017; Lien et al., 2012, 2019) and have not specifically explored language courses as a potential causal pathway between cultural institutes and trade flows.

Intuitively, closing the cultural gap between trading partners as well as improving communication should be beneficial for foreign trade between countries. Nevertheless, there are products which require more communication effort and/or mutual trust than others. The theoretical basis for this argument is best reflected in Williamson (1979) and Williamson (1985) which consider asset-specificity to be the driving force for the governance cost[45] necessary to trade products. Non-specific assets therefore cause only few governance costs when traded on a market compared to more specific assets. With an increasing asset specificity, more hierarchical types of organizations tend to generate much lower governance costs than the market.

Similar to this transaction cost view, Rauch (1999) argues that trade frictions (e.g. search costs) increase with the degree of differentiation of the products. In order to formulate categories to distinguish the products, Rauch (1999) offers a tripartite system: homogeneous goods, listed goods, and differentiated goods. Homogeneous goods are products that are traded on an organized exchange; listed goods define products that are listed in trade publications and therefore have a reference price; differentiated goods describe products that do not have a reference price and are potentially bought from a specific supplier. According to Rauch (1999), the latter category of differentiated goods is expected to be most affected by search costs (as an obstacle to trade). Rauch (1999) finds that sharing colonial ties and/or a common language has a positive effect on

---

[45] According to Williamson (1989), governance is defined as the means to create order, mitigate conflict and realize mutual gains. The costs of the usage of a type of governance (market, hybrid, hierarchy) are therefore of administrative and bureaucratic nature.

all three categories but is most beneficial for trade with differentiated products. Melitz and Toubal (2014) add to this research by analysing the impact of different concepts of common languages (common native/spoken/official language) on trade. They find that trade with differentiated goods benefits especially from common languages. Egger and Toubal (2018) additionally account for acquired language commonality and show that acquired language commonality has a positive effect on trade with differentiated goods, whereas the effects for homogeneous were mostly statistically insignificant. Investigating the special case of Switzerland as a multilingual country, Egger and Lassmann (2015) apply a spatial regression discontinuity design to show the causal impact of common native languages in foreign trade. Their findings demonstrate that common native languages impact the extensive margin in foreign trade more than the intensive margin. Here also do the effect sizes differ across the three Rauch (1999) categories, suggesting a particular importance for trading differentiated products. These results suggest that the role of language is more prominent when it comes to the trading of more complex products. Finally, Felbermayr and Toubal (2010) show that cultural proximity positively impacts international trade with differentiated goods but find no effect on homogeneous goods.

4.2.2   The Goethe-Institut

As a cultural institute, the GI (and its branches, i.e. the individual institutes) acts on the behalf of the German government and contributes to Germany's foreign cultural policy. The main duty of the GI is to promote German language and culture worldwide. The GI is closely connected to the Federal Foreign Office (FFO) which provides the main funding for the GI. Only language services are funded by fees (Goethe-Institut, 2014).

Regarding the promotion of language, the GI offers language courses and standardized exams, provides scholarships and trains local German teachers to improve the quality of teaching. Furthermore, the GI builds on cultural exchange and cooperation by offering cultural events and providing information on German culture and society, e.g. by maintaining libraries. These services are provided by institutes distributed all over the world. Figure 4.1 shows countries with institutes in 2014. In that year, 143 institutes were operating in 93 countries. 126 of these institutes offered language services. Chapter 2 and Chapter 3 provide a comprehensive overview on the history of the GI and its

Figure 4.1: The presence of the GI in 2014.

institutes, paying particular attention to the development of language learning at the GI.

When it comes to the decision of opening and closing institutes, the GI works together with the FFO. These decisions are influenced by many factors, such as legal, social and political aspects of the host countries, recent global and regional changes and by the main objective of the GI to promote German culture and language (see also Chapter 2 and 3). Analysing the determinants of the openings and closings of institutes and the stock of institutes in operation, Jaschke and Keita (2021) find that the share of German exports to a country does not have a significant influence on any of these three measures, i.e. the stock of institutes and their turnover. Lien and Oh (2014) show for the case of the Confucius Institute that among FDI flows, the geographical distance and development status of a country and trade are important determinants for the establishment of Confucius Institutes. Thus, a reasonable suspicion arises that German trade also affects the openings of institutes, and therefore we need to deal with the issue of reverse causality (see Section 4.4.2).

4.3 Data and empirical strategy

4.3.1 Data

4.3.1.1 Dependent variable: exports

To study trade flows from Germany to other countries, we use export data (measured in current British pounds) provided by Fouquin and Hugot (2016), which mainly draw on the Trade Statistics by the International Monetary Fund. In additional explorations investigating Swiss trade flows, we use Swiss export data provided by the Swiss Federal Customs Office.

In order to distinguish German exports according to different trade categories that are potentially heterogeneously affected by linguistic and cultural differences, we use data provided by Comtrade (2019). From these data, we were able to extract German exports according to their Standard International Trade Classification (SITC) Revision 2 and 3 on a four-digit level. We then aggregated them to homogeneous and differentiated goods, following the classification suggested by Rauch (1999). Trade values can potentially differ depending on the reporting country. A German export is reported by Germany as well as by the importer country. For the purpose of this study, there is a need only for data on the single exporter, and we rely on the data reported by Germany. These data have been reported by Germany starting in 1978. A potential disadvantage of using the Comtrade data is that zero-trade flows and missing flows in product types cannot be distinguished from each other but this problem is not relevant for our sample if we only use data reported by Germany.

The three Rauch (1999) categories are ranked according to their degree of differentiation. In the analysis, only homogeneous and differentiated goods will be used to compare the extremes of the spectrum and to avoid vagueness related to the category of listed products. Furthermore, we use the conservative way to aggregate the data, which, in case of ambiguities about which category applies, is designed to minimize the number of homogeneous and listed entries. Thus, compared to the liberal categorization, on average homogeneous goods but require more and differentiated goods less communication to be traded. To use the conversion tables provided by Rauch (1999), the reported data have to be categorized according to the second or third revision of SITC on a 3- or 4-digit level.

4.3.1.2 Independent variable: language-learning opportunities

Our independent variable of prime interest captures language-learning opportunities abroad. We use data on the presence of institutes based on the annual reports by the GI (see Chapter 2). The data provide information on institutes at the city-year level, which we aggregate to the country level. For our main specification, we consider the number of institutes a country hosts in a particular year. The dataset also provides information on whether the institutes provide language services, i.e. standardized language certificates and language courses. In order to investigate whether language services are the driver of our results, we make use of this information. However, we have to keep two limitations in mind: first, there is a measurement error in this information, as in some cases the annual reports do not provide information if language services are offered. In particular, this is the case when the number of course and exam participation are reported jointly with other institutes. For example, institutes that are subsidiaries of other main institutes have reported their numbers jointly in the annual reports by the GI (for more information on the different types of institutes in the dataset, refer to Chapter 2). Second, there are relatively few institutes without language services (see Figure 4.4).

4.3.1.3 Other control variables

We control for several factors that are associated with export flows and that may confound, if ignored, the relation between language-learning opportunities abroad and trade. First, the GDP of the importer country controls for the economic (business-cycle) condition in and the economic size of the importer country. Second, to proxy trade openness of an importer country, we consider the relation of its total imports to its GDP. Third, we include a control for the population size of an importer country and an indicator for its membership in the General Agreement on Trade and Tariffs (GATT). All these data are provided by Fouquin and Hugot (2016). Finally, we control for trade agreements with two dummy variables provided by Mario Larch's Regional Trade Agreements Database from Egger and Larch (2008). The first one indicates for each year if both Germany and the importer country are members of a customs union. For the Swiss case, there is no variation in this variable, and we omit it in our regressions. Second, we include a dummy variable that indicates whether the importer country has a free trade agreement (FTA) and/or an economic integration agreement (ETA) with Germany (Switzerland).
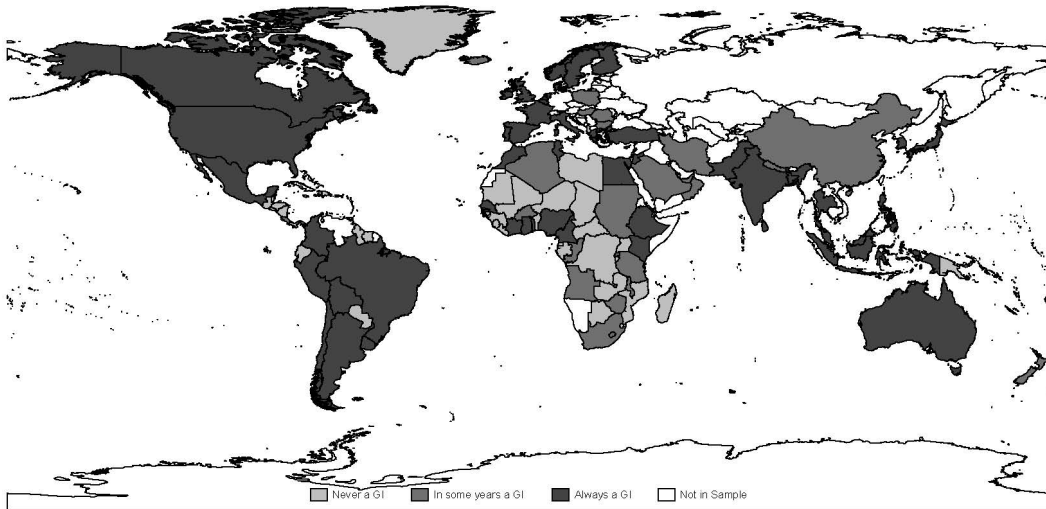
Figure 4.2: Map of countries and the presence of institutes in the sample.

4.3.1.4 Sample construction

We construct our sample to attain a balanced sample for 1978-2014. After excluding Switzerland, Luxembourg and Austria as (partly) German-speaking countries, we end up with a balanced sample of 134 countries for the aggregated export flows.[46] For the analysis of homogenous and differentiated goods (Section 4.4.3), we balance our sample over the same time period (1978-2014) across the two categories, which results in 95 importer countries with positive trade flows from Germany. In order to keep the samples comparable, we start the observation period for both samples in 1978, as for the conversion tables provided by Rauch (1999) the reported data have to be categorized according to the second or third revision of SITC on a 3- or 4-digit level. These have been reported by Germany starting in 1978. Additionally, the problem of joint reporting of two or more institutes, and therefore the uncertainty about the supply of language services, was especially large starting in the beginning of the 1970s until 1978.
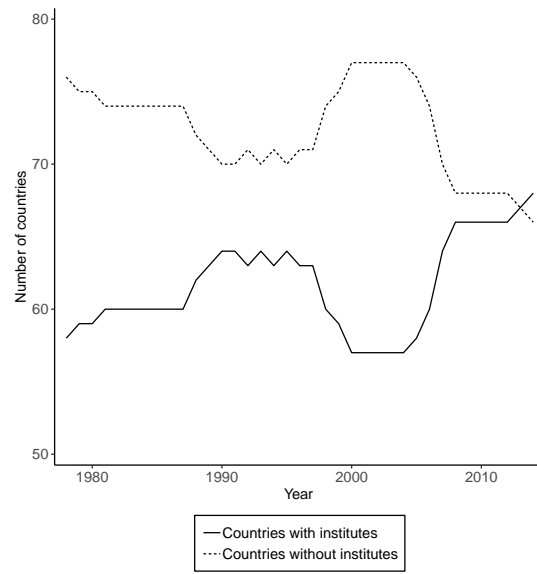
Figure 4.3: Number of countries with institutes in the sample.
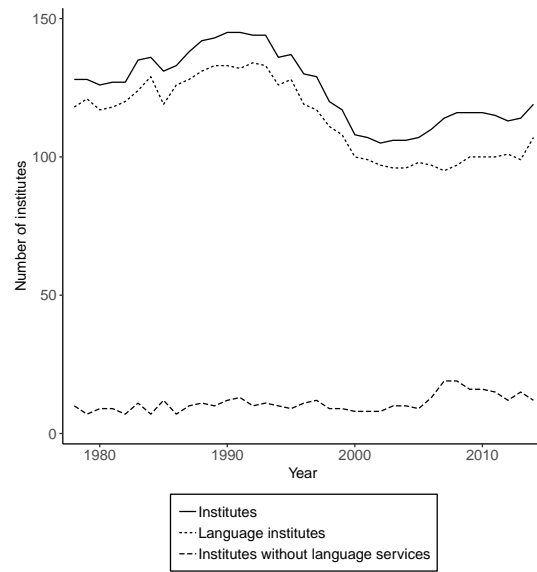
Figure 4.4: Number of institutes in the sample.

99
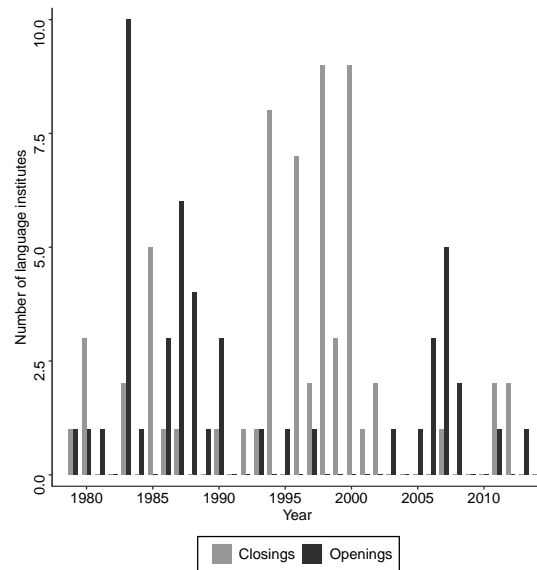
Figure 4.5: Openings and closings of institutes in the sample.

4.3.1.5 Descriptive statistics

Our main sample consists of 134 countries. The GI was never present in 60 countries in our period of observation, while 24 countries had institutes in some years and 50 countries had institutes throughout the whole period. Figure 4.2 shows a map of the countries in our main sample. Figure 4.3 presents the number of countries with and without institutes for each year in the observation period. While the number of countries with institutes increased until the 1990s, at the end of the 1990s the number started to decrease to 57 between 2000 and 2004. In the latter years, the number rose again and peaked in 2014 with 68 countries. While the number of countries with institutes has increased, the number of institutes itself has decreased since the beginning of the 1990s, as can be seen in Figure 4.4. The number of institutes that do not offer language courses stays fairly constant and rather small throughout the observation period.[47] For our estimation, we use the variation of number of institutes per

---

[46] We balance our sample to assure comparability across different specifications. In Section 4.4.1.1, we show that results do not only depend on balancing the sample but also hold for an unbalanced sample.

[47] In Figure 4.4, we assume that all institutes that are jointly reported offer language services.

100

country, which includes the openings and closings of institutes. An overview of openings and closings over time is presented in Figure 4.5.

In our sample, the countries that on average receive the highest German exports are France, the US, Italy, Great Britain and Belgium. Only 22 country-year observations exhibit zero trade flows. Average exports to a country in our data are worth around 2.3 billion British Pounds. Median exports, however, are only about 7.7 million British Pounds, which indicates that the distribution of exports is highly left-skewed, with relatively few but very large importers. Table 4.1 provides summary statistics for the variables we use in the analysis.

Table 4.1: Summary statistics

Variable	Obs		Mean	Std. Dev.	Min	Max
German exports in Mill. GBP	4958	Overall	2298.73	7522.31	0	84331.43
Number of institutes	4958	Overall	0.93	1.62	0	12
		Between		1.56	0	8.57
		Within		0.45	-2.64	5.50
Number of language institutes	4958	Overall	0.80	1.33	0	9
		Between		1.26	0	6.4865
		Within		0.43	-3.21	5.04
Number of institutes	4958	Overall	0.08	0.43	0	6
without lang. serv.		Between		0.32	0	2.92
		Within		0.29	-1.84	4.49
Population (in 1000)	4958	Overall	37630.76	137105.80	49.2	1364270
GDP in Mill. GBP	4958	Overall	135012.90	615356.80	12.02	10700000
Total Imports/GDP	4958	Overall	0.48	1.34	0.04	45.81
Customs Union	4958	Overall	0.10	0.30	0	1
FTA/EIA	4958	Overall	0.13	0.34	0	1
GATT	4958	Overall	0.76	0.43	0	1

4.3.2   Theoretical background and empirical strategy

Originating from Newton's law of gravity, the gravity model in economics provides a useful framework to analyse international trade patterns. In simple terms, the gravity model explains trade as an increasing function of the economic size of trading partners and a decreasing one with respect to trade frictions. In the field of international trade, it has been used by Tinbergen (1962) for the first time. It soon developed into a widely used tool in empirical analyses which exhibit a sound foundation in trade theory (Anderson and Van Wincoop, 2003). In the current application of the gravity model, the analysis

centers around a single exporter (Germany) and its trade relations with other countries. While this setting does not allow for checking multilateral resistance, the consideration of only trade flows from one country has the essential advantage of permitting the distinction of exports from imports (Földvári, 2006) and, further, an analysis of the effects of the opening of a cultural institute as a tool for export promotion. Our choice of control variables (see Section 4.3.1.3) is mainly motivated by the extensive overview by Head and Mayer (2014) of the gravity model as a workhorse in international trade.

In the presence of heteroskedasticity, estimates of the gravity model are biased and inconsistent if estimated with fixed-effects ordinary least squares, including a log-linearized version of the dependent variable. Santos Silva and Tenreyro (2006) show that estimating gravity models of trade with the Poisson pseudo-maximum likelihood (PPML) estimator has the advantage of being consistent and unbiased in the presence of heteroskedasticity. Additionally, the PPML estimator solves the problem of including zero trade flows, as the dependent variable is included in levels (not logs). By using the PPML estimator in our empirical analysis, we follow other researchers who estimate trade flows from but a single exporter (Johnston et al., 2015; Lien et al., 2019). Drawing from the theoretical specification of the gravity model, we estimate the following equation:

$$y_{jt} = exp\left[\boldsymbol{\alpha'GI_{jt}} + \boldsymbol{\beta'x_{jt}} + \boldsymbol{\phi'_t d_t} + \boldsymbol{\phi'_j d_j}\right] + \eta_{jt} \tag{4.1}$$

where y_{jt} represents exports from Germany to the importer country j in year t. $\boldsymbol{GI}$ represents the vector of GI related variables, our main variables of interest. For each specification we estimate, we always control for the number of institutes and its square to capture the possibly non-linear effect of another additional institute. In our main specification, we use the number of all institutes. $\boldsymbol{x_{jt}}$ is the vector of control variables. It includes the log of the population in the importer country j, dummy variables indicating joint membership of the European custums union, an EIA or FTA between Germany and the importer country j in year t, and whether the importer country j is a member of the GATT. Furthermore, we control for economic conditions in the importer country with the log of GDP. As a measure for trade openness, we include the share of total imports in GDP. We also include two sets of dummies to control for importer country effects, and year fixed effects. Specifically, year dummies d_t control for importer-invariant effects, and importer dummies d_j control for time-invariant characteristics of the importer country j and its relations

Table 4.2: Estimation results

DV: Exports	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Institutes	0.317***	0.170***	0.144***	0.139***			0.122***
	(0.099)	(0.052)	(0.043)	(0.043)			(0.036)
Institutes sq.	-0.017***	-0.011***	-0.010***	-0.010***			-0.009***
	(0.005)	(0.003)	(0.002)	(0.002)			(0.002)
Inst./pop. in mill.						1.059***	
						(0.320)	
Inst./pop. in mill. sq.						-0.259***	
						(0.081)	
log GDP		0.846***	0.772***	0.786***	0.804***	0.794***	0.796***
		(0.053)	(0.052)	(0.052)	(0.058)	(0.054)	(0.048)
log population		-0.555*	-0.428	-0.425*	-0.344	-0.405*	-0.421
		(0.312)	(0.260)	(0.255)	(0.294)	(0.244)	(0.265)
CU			0.555***	0.550***	0.573***	0.573***	0.454***
			(0.111)	(0.110)	(0.141)	(0.115)	(0.117)
FTA/EIA			0.167*	0.167*	0.219**	0.182**	0.153*
			(0.088)	(0.086)	(0.088)	(0.082)	(0.079)
GATT			0.265***	0.255***	0.299***	0.279**	0.138
			(0.098)	(0.096)	(0.114)	(0.114)	(0.085)
Import share				0.088***	0.092***	0.087***	0.095***
				(0.022)	(0.023)	(0.020)	(0.026)
Constant	19.100***	7.845	7.283*	6.870	5.139	6.387	3.211
	(0.041)	(4.904)	(4.314)	(4.251)	(4.879)	(4.098)	(2.977)
Observations	4,958	4,958	4,958	4,958	4,958	4,958	6,206
R-squared	0.970	0.991	0.992	0.992	0.991	0.991	0.992
Countries	134	134	134	134	134	134	184
Years	1978-2014	1978-2014	1978-2014	1978-2014	1978-2014	1978-2014	1978-2014
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Cluster-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

with Germany. η_{jt} denotes the error term. We cluster standard errors at the importer country level.

4.4 Results

4.4.1 German exports

4.4.1.1 Main results

Table 4.2 shows PPML regression results for seven basic specifications of trade flows from Germany to 134 importer countries. The interpretation of coefficients estimated by PPML is equivalent to that of coefficients estimated by OLS when the dependent variable is in logs: coefficients of independent vari-

ables in logs can be interpreted as elasticities, and coefficients of independent variables in levels capture semi-elasticities. We sequentially introduce control variables in columns (1)-(4). Column (5) shows results when GI variables are omitted as regressors. As is evident from the table, estimated coefficients remain statistically significant (with one exception) and also of similar magnitude to those reported in column (4). The one exception is the coefficient on log population, which is now slightly smaller and also imprecisely estimated. Column (6) relates the number of institutes to the population size of an importer country, instead of considering their absolute number, and column (7) uses an unbalanced (and hence larger) sample.

Before focusing on our main variables of interest, we briefly discuss the coefficients of the control variables. The estimated coefficient for GDP in the importer country is positive and significant in all specifications. The estimated coefficient for population size in the importer country is negative and significant at the 10 %-level in column (4). Being a member of the European Union Customs Union (EUCU), i.e. the only customs union in which Germany is a member, increases German exports to a country. The positive effect of an EIA and/or FTA membership is only significant in PPML estimations at the 10 %-level. GATT membership increases exports significantly. This effect vanishes in the unbalanced sample in column (7). Finally, an importer country's openness to trade, measured by total imports/GDP, has a positive and significant effect on German exports.

We next turn to our main variable, which captures German language-learning opportunities in importer countries. We hypothesized that an increase in German language-learning opportunities in a country increases German exports to that country. And indeed, this is what we find. As shown in column (4), the coefficient for an additional institute is 0.139 and highly significant. However, the marginal effect of an additional institute is highest for the first institute in a country and decreases with an increasing number of institutes, which can be seen by the negative and significant effect of the associated squared term of institutes.

Column (6) tests if the estimated effect depends on the size of the population (or market) an institute serves in an importer country, as proxied by the ratio of language institutes to total population (in million inhabitants). With this specification, we account for differences in openings between large (e.g. India) and small (e.g. Hungary) countries. As it turns out, the estimated effect re-

mains positive and significant and therefore robust to the use of this alternative measure.

Finally, we want to test if our results depend on our choice of a balanced sample whereby we consider only countries with non-missing observations in all years, 1978-2014. Column (7) shows results for an unbalanced sample with 184 countries. As can be seen, our results turn out to be robust to this change in sample but the sizes of our two coefficients become slightly smaller than in column (4).

4.4.1.2 Heterogenous effects for country groups

Table 4.3: Heterogenous effects for country groups

DV: exports	(1) advanced economies	(2) non-advanced economies	(3) Germanic language	(4) non-Germanic language	(5) members of EUCU	(6) non-members of EUCU
Institutes	0.106**	0.416***	0.045	0.197***	0.049*	0.169**
	(0.044)	(0.115)	(0.037)	(0.058)	(0.028)	(0.073)
Institutes sq.	-0.008***	-0.076***	-0.004*	-0.020***	-0.012***	-0.011***
	(0.003)	(0.009)	(0.002)	(0.004)	(0.002)	(0.004)
Observations	2,479	2,479	629	4,329	506	4,452
R-squared	0.992	0.997	0.994	0.994	0.994	0.991
Years	1978-2014	1978-2014	1978-2014	1978-2014	1978-2014	1978-2014
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes

Cluster-robust standard errors in parentheses. Other controls (log (GDP), log population, CU, FTA/EIA, GATT, import share and constant are included but not shown. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The effect of institutes on German exports to a country might also be sensitive to and vary with certain importer country characteristics. We therefore perform subgroup analyses for groups of importer countries differentiated by their economic development, their linguistic distance to German and their membership status in the EUCU. Table 4.3 reports the results. We find that the results of all three subgroup analyses go in the same direction: countries with already lower trade barriers benefit less from an institute than those countries with higher trade costs. This indicates that economies that are facing higher trade barriers initially benefit most from the establishment of institutes.

First, we split the sample according to importer countries' economic development. We classify a country as an advanced (non-advanced) economy if its average GDP/capita over the observation period is larger (smaller) than the

median GDP/capita of importer countries in our sample. Columns (1) and (2) show for both subsamples a significant and positive but decreasing effect of the institute numbers. This effect is much more pronounced, however, for non-advanced economies. These findings are similar to those of Lien et al.'s (2019).

Second, we split the sample according to the linguistic distance between the language spoken by the majority of the population in the importer country and the German language (Adserà and Pytliková, 2015). In column (3), we consider only countries with a Germanic language and in column (4) only countries where the majority of the population speaks a non-Germanic language. As is evident from a comparison of column (3) and column (4), institutes only exert an effect on German exports in countries where the majority of the population speaks a non-Germanic language. This result might be explained either by the fact that in these countries German is already a more common language or that the knowledge of English is even better and therefore used as the lingua franca for economic exchanges. As described in Section 4.2.1, linguistic distance increases the costs of trade, in particular communication costs. Consequently, an institute helps to overcome these larger costs and creates a larger potential for cost reductions than for linguistically closer languages where communication costs are already less.

Third, when splitting up the sample into members and non-members of the EUCU, we find that the effect for non-members is much larger and also more precisely estimated than the effect for members of the EUCU. A potential explanation for this finding is the following. While trade barriers within the EUCU are already very low, language skills might help to overcome existing language barriers and increase trade in non-member states of the EUCU.

All three subgroup analyses suggest that the benefit from the openings of institutes is larger for economies that face higher trade barriers.

4.4.2 Reverse causality: the case of Switzerland

One of the reasons for opening or closing an institute may be trade promotion. If so, then our main explanatory variable is not exogenous to German exports, and the relationship we estimate between institutes and trade flows suffers from reverse causality. The GI together with the German government decide whether and where to open or close institutes. However, neither of them consults the Swiss government or any other Swiss institution, nor does any Swiss institution

Table 4.4: Estimation results: Swiss exports

	(1)	(2)	(3)	(4)	(5)	(6)
DV: Exports	CHE (German-speaking)			CHE (non-German-speaking)		
Institutes	0.094*	0.134***	0.137***	0.104	0.100	0.100
	(0.050)	(0.047)	(0.047)	(0.080)	(0.064)	(0.064)
Institutes sq.	-0.008***	-0.010***	-0.010***	-0.007	-0.007*	-0.007*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
log GDP	0.574***	0.615***	0.618***	0.589***	0.599***	0.599***
	(0.090)	(0.080)	(0.079)	(0.104)	(0.088)	(0.087)
log population	0.060	-0.068	-0.062	-0.029	-0.013	-0.010
	(0.208)	(0.185)	(0.179)	(0.250)	(0.199)	(0.198)
FTA/EIA = 1	0.167***	0.150***	0.149***	0.090	0.089	0.088
	(0.051)	(0.051)	(0.051)	(0.085)	(0.083)	(0.083)
GATT	0.252	0.149	0.145	-0.170	-0.225	-0.225
	(0.218)	(0.123)	(0.123)	(0.343)	(0.172)	(0.172)
Import share	0.094**	0.098**	0.099**	0.119**	0.120**	0.120**
	(0.044)	(0.045)	(0.045)	(0.059)	(0.059)	(0.059)
Constant	1.286	2.519	3.036	1.013	0.566	2.537
	(3.297)	(3.201)	(2.315)	(2.956)	(2.752)	(1.898)
Observations	2,680	3,340	3,584	2,680	3,340	3,584
R-squared	0.984	0.983	0.984	0.987	0.987	0.988
Countries	134	167	183	134	167	183
Years	1995-2014	1995-2014	1995-2014	1995-2014	1995-2014	1995-2014
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes

Cluster-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

have any say in this regard. The openings and closing of institutes should hence be exogenous to Swiss exports. Following Chapter 3, which first made use of this reasoning and identification strategy in the context of migration, we replace German exports with Swiss exports as our dependent variable. Furthermore, distinguishing between Swiss exports from German-speaking and non-German-speaking cantons in Switzerland allows us to perform a placebo test for the German-speaking part of Switzerland. In other words, we see this part of Switzerland as non-treated by German language-learning opportunities and therefore expect no significant effect of the GI variables. The two groups of cantons should not differ substantially in institutional factors. Therefore, the non-German-speaking cantons should be close to a suitable counterfactual for the German-speaking cantons.

Export data for Switzerland at the canton level are available from 1995. We aggregate export flows by cantons that speak primarily German (German cantons) and cantons that do not (non-German cantons).[48] We estimate our preferred specification[49] (Table 4.2 Column 4) for the three regions the German-speaking part of Switzerland (Table 4.4 Column 1-3), the non-German-speaking part of Switzerland (Table 4.4 Column 4-6), Germany (Table 4.A1 Column 1-3) and Switzerland (Table 4.A1 Column 4-6). For each region, we run regressions using three different estimation samples for the period 1995-2014. The first sample contains exactly the same countries as the main sample and is therefore balanced. The second includes all 167 countries for which we have balanced observations for all three regions. The third sample includes all available observations for all three regions, i.e. this third sample is unbalanced, and it contains observations on 183 importer countries.

As shown in Table 4.A1 columns (1)-(3), which consider German exports for the shorter period 1995-2014, the size of the coefficient on the institute count variable is around 0.06 for all three samples, which is less than half the size of the coefficient estimate in column (4) in Table 4.2. The estimated effect is significant for the larger samples in columns (2)-(3) but only imprecisely estimated in column (1). The size of the coefficient for the squared term is also about half of the size as in column (4) in Table 4.2 and is significant in

---

[48] The canton with the lowest share of German native speakers (68.3 %) classified as German-speaking is Graubünden. The canton with the highest share of German native speakers (29.2 %) that is classified as non-German-speaking canton is Fribourg. 18 out of the total of 26 cantons are classified as German-speaking cantons.

[49] The control variable indicating membership in a customs union is omitted for Switzerland, as there is no variation in this variable in the data.

all three specifications. Table 4.A1 Column 4-6 show the same estimations for total Swiss exports. The relationship between the number of institutes and Swiss total exports is larger than 0.1 and significant in all three specifications, at least at the 5 % level.

When we split exports into exports from the German and the non-German speaking parts of Switzerland, we see that the relationship is only due to exports from the German-speaking parts. Table 4.4 columns (1)-(3) present estimates with exports from the German-speaking part of Switzerland. For the two larger samples, the size of the coefficient for the number of institutes is about the same as column (4) in Table 4.2. For the smaller sample in column (1), it is slightly smaller and only significant at the 10 % level. These results suggest a causal effect of the number of institutes on export flows, as the number of German Goethe institutes is arguably exogenous to Swiss trade flows. However, the results in column (1) and (2) hint at heterogeneous effects for different groups of countries, which seem to be more present in the larger samples. A potential reason for this effect heterogeneity may be that German exports vary in type across importer countries, requiring lesser or greater language skills for their international trade. In the next subsection, we investigate the language skill requirement of different German export flows, and thereby also provide evidence on the importance of language for international trade and our main finding of a link between institutes and German exports. With respect to the non-German-speaking part of Switzerland, the placebo test is successful with all three samples: there is no significant positive effect of opening an institute on exports. The size of the coefficient for the number of institutes is around 25 % smaller, and the estimations are much more imprecise than in the German-speaking part. This is similar for the squared number of institutes.

4.4.3 Rauch product categories and the language channel

In light of previous results by Rauch (1999) and other scholars we reviewed in Section 4.2.1, we expect that the establishment of cultural institutes may aid international trade. More specifically, we hypothesize that the supply of language courses particularly facilitates the trade of differentiated goods, as it requires greater communication and hence higher language proficiency of trading parties. In this section, we test these expectations by re-estimating our main specification from column (4) in Table 4.4 with export flows aggregated accord-

ing to the classification suggested by Rauch (1999). We are using Comtrade data reported by Germany for the longest time period available (1978-2014) that allow the conversion to Rauch-categories and a balance of our observations. The final sample covers 95 countries. For comparison, we also estimate aggregated exports for the same sample as in column (1) in Table 4.5, using the data provided by Fouquin and Hugot (2016).

Institutes that offer German courses and institutes that do not can both contribute to the establishment of networks and help to bridge the cultural gap between countries. Cultural institutes such as the GI may have a general trade-promoting effect. However, institutes that offer language courses may benefit trade even more by increasing trade in communication-intensive goods. To explore this possibility, we modify our main specification in column (4) in Table 4.2 by splitting the total number of institutes into those which offer language services and those which do not.[50] This allows us, as noted, to further investigate language as the potential driver of the effect of institutes on exports. Both specifications make use of the whole set of control variables and employ PPML for estimation. As additional outcomes in this exercise, we consider exports of homogeneous goods and of differentiated goods, which represent the two 'extremes' in the classification by Rauch (1999).

Similar to the previous specifications, in the reduced sample with only 95 countries (see column (1)), the number of institutes has a positive but decreasing impact on overall exports. The coefficients for the other control variables have the expected signs and are similar in size to those we obtained in earlier regressions. In column (2), we split the number of institutes into those that offer language services (i.e., language institutes) and those which do not. Keeping in mind that there are relatively few institutes who do not offer language services, the results show that the effect seems to be driven by language institutes, which exert a statistically significant effect. The estimated coefficient for non-language institutes, in contrast, turns out to be both much smaller and also imprecisely estimated.

---

[50] Column (2) in Table 4.5 shows results only for the reduced sample of 95 countries. In Table 4.A2 in the Appendix, however, we show that results do not differ for the full sample of 134 countries. Because of joint reporting of two or more institutes in the annual year books, it is not clear for some institutes if they actually do offer language courses or not. The specification in column (1) in Table 4.A2 assumes that only the main institute offers language services in these cases. Column 2 instead assumes that all institutes with joint reporting do offer language services. As it turns out, the results do not depend on which assumption is made. In Table 4.5, we report results from specifications where we assume that only the main institute offers language services.

Table 4.5: Estimation results: Rauch categories

DV: exports	(1) Aggr.	(2) Aggr.	(3) Hom.	(4) Hom.	(5) Diff.	(6) Diff.
Institutes	0.130***		0.057		0.106**	
	(0.040)		(0.120)		(0.041)	
Institutes sq.	-0.010***		-0.003		-0.009***	
	(0.002)		(0.007)		(0.002)	
Lang. Institutes		0.132***		0.059		0.086**
		(0.051)		(0.068)		(0.042)
Lang. Institutes sq.		-0.016***		0.000		-0.012***
		(0.005)		(0.005)		(0.004)
Institutes w/o lang. serv.		0.024		-0.005		0.006
		(0.034)		(0.045)		(0.029)
Institutes w/o lang. serv. sq		-0.004		-0.003		-0.003
		(0.006)		(0.006)		(0.005)
log GDP	0.791***	0.802***	0.664***	0.568***	0.902***	0.918***
	(0.053)	(0.051)	(0.118)	(0.120)	(0.054)	(0.057)
log population	-0.369	-0.375	-0.021	-0.294	-0.347	-0.328
	(0.248)	(0.255)	(0.417)	(0.314)	(0.258)	(0.268)
CU	0.383***	0.361***	0.658***	0.567***	0.283***	0.266***
	(0.079)	(0.088)	(0.122)	(0.132)	(0.066)	(0.079)
FTA/EIA	0.192*	0.202*	0.145	0.108	0.194	0.213
	(0.116)	(0.118)	(0.147)	(0.136)	(0.131)	(0.136)
GATT	0.250**	0.288***	-0.255	-0.166	0.273**	0.307**
	(0.100)	(0.105)	(0.221)	(0.213)	(0.114)	(0.124)
Import share	0.088***	0.090***	0.133	0.086*	0.088***	0.091***
	(0.021)	(0.022)	(0.081)	(0.052)	(0.019)	(0.021)
Constant	5.621	6.103*	0.946	8.242	2.749	2.366
	(3.498)	(3.614)	(6.943)	(5.320)	(3.627)	(3.937)
Observations	3,515	3,515	3,515	3,515	3,515	3,515
R-squared	0.992	0.993	0.950	0.957	0.991	0.991
Years	1978-2014	1978-2014	1978-2014	1978-2014	1978-2014	1978-2014
Countries	95	95	95	95	95	95
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes

Cluster-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Columns (3) and (4) consider exports of homogeneous goods, while columns (5) and (6) consider exports of differentiated goods. Institutes do not appear to affect exports of homogeneous goods. The same holds true for the two subgroups of institutes with and without language-learning opportunities. However, for differentiated export goods, institutes exert an effect similar to the one observed for overall (aggregate) exports, albeit one that is slightly smaller in magnitude. As shown in column (6), this effect is driven by institutes with language-learning opportunities. Institutes which do not offer language services do not impact trade in differentiated goods. Therefore, as was expected and as the previous literature suggests, trade in differentiated goods benefits from cultural institutes particularly. Furthermore, our evidence shows that this effect is driven by institutes that offer language services, which suggests that language is indeed the driver behind the effect of such institutes on exports.

4.5   Conclusion

This chapter adds to the literature on cultural institutes and their impact on foreign trade in several ways. We show that language-learning opportunities (and not mere cultural services) offered by the GI abroad play an important role in export promotion. This strongly suggests that language (and more specifically foreign language-acquisition) is a relevant channel for foreign trade relations.

We also find that the strength of the export-promoting effect of institutes differs across countries. The effect is more pronounced for less advanced economies, countries with a non-Germanic native language (i.e., a language that is linguistically further away from German) and countries that are not part of the same trade union as Germany.

Concerning ¡otential reverse causality, we find that institutes abroad also have a positive impact on Swiss exports to countries hosting these institutes, although Switzerland has no say in the decision of where institutes are operated. Distinguishing between exports by non-German-speaking cantons and German-speaking cantons shows that only the latter benefit from establishing institutes abroad.

Finally, in line with the previous literature, we also find evidence for effect heterogeneity by type of export product. Using the product classification suggested by Rauch (1999), we find that institutes, specifically those that offer language-

learning opportunities, benefit the export of differentiated goods, which require more communication for their trade. Trade in homogeneous goods, in contrast, appears to be unresponsive to the presence and activities of institutes.

Appendix to Chapter 4

4.A Appendix A

Table 4.A1: Estimation results: German and Swiss total exports

	(1)	(2)	(3)	(4)	(5)	(6)
DV: Exports		DEU			CHE	
Institutes	0.057	0.062**	0.063**	0.103**	0.130***	0.132***
	(0.039)	(0.029)	(0.029)	(0.042)	(0.039)	(0.039)
Institutes sq.	-0.005**	-0.005***	-0.005***	-0.008***	-0.010***	-0.010***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
log GDP	0.726***	0.747***	0.745***	0.544***	0.582***	0.584***
	(0.055)	(0.048)	(0.048)	(0.078)	(0.068)	(0.067)
log population	-0.184	-0.185	-0.185	0.106	0.024	0.027
	(0.311)	(0.276)	(0.275)	(0.206)	(0.175)	(0.172)
CU = 1	0.377***	0.225***	0.224***			
	(0.077)	(0.078)	(0.078)			
FTA/EIA = 1	0.056	0.008	0.005	0.155***	0.143***	0.142***
	(0.074)	(0.066)	(0.066)	(0.049)	(0.047)	(0.047)
GATT	0.400***	0.165*	0.166*	0.143	0.050	0.047
	(0.090)	(0.089)	(0.089)	(0.262)	(0.137)	(0.137)
Import share	0.085***	0.092***	0.092***	0.126*	0.127**	0.128**
	(0.024)	(0.027)	(0.027)	(0.064)	(0.065)	(0.065)
Constant	3.767	3.493	1.368	1.422	1.970	3.274*
	(4.846)	(4.433)	(2.996)	(2.763)	(2.666)	(1.839)
Observations	2,680	3,340	3,584	2,680	3,340	3,584
R-squared	0.995	0.994	0.995	0.990	0.990	0.990
Countries	134	167	183	134	167	183
Years	1995-2014	1995-2014	1995-2014	1995-2014	1995-2014	1995-2014
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes

Cluster-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4.A2: Estimation results: institutes with and without language services

DV: Exports	(1)	(2)
Lang. Institutes	0.136***	0.142***
	(0.041)	(0.054)
Lang. Institutes sq.	-0.015***	-0.017***
	(0.003)	(0.005)
Institues w/o lang. serv.	-0.034	0.026
	(0.037)	(0.035)
Institues w/o lang. serv. sq	0.005	-0.004
	(0.007)	(0.006)
Observations	4,958	4,958
R-squared	0.993	0.993
Countries	134	134
Years	1978-2014	1978-2014
Other controls	Yes	Yes
Year FE	Yes	Yes
Country FE	Yes	Yes

Cluster-robust standard errors in parentheses. Other controls (log GDP, log population, custom unions, FTA/EIA, GATT, import share) and constant are included but not shown. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Part II

Language Learning: Human Capital Investment or Consumption

Chapter 5

Language Learning: Human Capital Investment or Consumption

Abstract.[51] 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 adding consumption motives. Based on a novel dataset collected in 14 countries worldwide from language course participants, 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 the professional use, which emerge mostly in a "tied-mover" context. This provides guidance for targeted migration and integration policies.

117

5.1 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 (Isphording and Otten, 2013b; Toomet, 2011) 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 using 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. Effects of immigrants' skills of the host country's language have been widely discussed. Researchers and policy makers alike emphasize the importance of those language skills for integration into the labour market of the host country. More specifically, language skills improve labour market outcomes of migrants by increasing earnings (see e.g Dustmann and Van Soest (2001); Chiswick and Miller (1995)) and employment probabilities (Budría et al., 2019; Dustmann and Fabbri, 2003) and by improving occupational choices (Aldashev et al., 2009).

All this evidence can be put in the light of the human capital theory (Becker, 1964; Schultz, 1960). 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 increasing 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 individuals' 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 (Canton and Jong, 2005; Oosterbeek and Ophem, 2000; Oosterbeek and Webbink, 1995). If we ignore irrational choices, expected labour market returns seem to be not the only determinant of those choices. One explanation

for the observed patterns is that education or, broadly speaking, learning has a consumption value and generates direct utility (Kodde and Ritzen, 1984; Lazear, 1977; Schaafsma, 1976). 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 chapter, we focus on foreign language learning and different motives based on learning as human capital investment or as consumption. Language learning leads to a particular form of skills that can be acquired in many different contexts, e.g. at school, at universities, but also in language courses. While choices about language acquisition at school are often determined by the school curriculum or parents' preferences, participating as an adult in a language course offered by a university or a private provider is more directly related to determinants which can be linked to the human capital motive or the consumption motive.

To address the question, why people decide to acquire foreign language skills, we use survey data which we collected in 14 countries worldwide from language course participants at the Goethe-Institut (GI), a German cultural institute. To the best of our knowledge, we are the first to apply the human capital framework enlarged by adding consumption motives to foreign language learning. By doing so, we want to contribute to a better understanding of foreign language acquisition in a cross-country perspective.

We proceed in three steps: First, we present descriptive evidence on all reasons for foreign language acquisitions by countries. Based on the participants' main reasons, we build categories which we then aggregate into the human capital investment motive and the consumption motive. We use binary probit estimations to find individual and country-level determinants of the human capital investment motive. In order to identify heterogeneities, we have a closer look at subgroups based on age, gender and countries. Second, we provide descriptive evidence on a second measure – the expected probability of using the foreign language skills in a professional environment – and show that a given human capital investment motive does not necessarily match with a high probability of professional use. Therefore, we apply again binary probit estimations to study individual and country-level determinants of a professional use and identify heterogeneities in a subgroup analysis based on age, gender and countries. Last, by looking at language learning by motives, we are 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. We find that the former case mostly emerges in a "tied-mover" context. Understanding these relations is important from a policy-point of view as it provides guidance for targeted migration and integration policies. As special focus will be on possible difference between European and non-European countries.

The remainder of the chapter is structured as follows. In Section 5.2, we describe how the survey was conducted and present the individual characteristics of the language course participants. Section 5.3 provides descriptive evidence of the reasons and their distributions across countries as well as by age and gender subgroups. It also introduces the concepts of categories and motives. Section 5.4 presents estimation results for the determinants of language learning based on the human capital motive. In Section 5.5, we show descriptive evidence of the expected probability of professional use, present estimation results about its determinants and discuss the spillover effects. Section 5.6 concludes.

5.2 Survey at the Goethe-Institut

5.2.1 Selection of countries and institutes

The GI is a German cultural institute that offers language courses worldwide and is an important part of the foreign cultural policy of the German government. In addition to language courses, 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 2019 the GI was present in 98 countries with a total of 156 institutes (Goethe-Institut, 2020).[52]

We selected 19 institutes in 14 countries and conducted a survey among language course participants between June and December 2018. The content of the pen and paper questionnaire included questions on socio-demographic characteristics, education, professional background, language skills, previous migra-

---

[52]   For more details on language learning at the GI, see Chapter 2.

tion experience and future migration plans as well as questions on the reasons of learning the German language.

Table 5.1 gives an overview of the selected countries in the survey and the main characteristics on which we based this selection to assure a heterogeneous sample of countries based on the following characteristics: geographic distance to the German-speaking region, linguistic distance to the German language, average income (GNI/capita) level as categorized by the World Development Indicator in 2018 (World Bank, 2021) and the absence (or presence) of migration barriers defined by the (lack of) freedom of movement within the European Union (EU) and the European Economic Area. 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.

In order to increase participation in our survey, we have taken several measures. First, the survey has been 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, 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). Third, we took several measures to encourage participation, which differed between European and non-European countries. In European countries, a team member of the research project was present for at least one unit of each course offered at the institutes during a given week and handed out the questionnaire to each participant present at that unit. Most of the participants filled-in the questionnaire 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, envelopes were prepared 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. In total, 6,664 language course participants submitted valid questionnaires. Table 5.1 gives an overview of the number of participants and the response rates by country. In European countries, the response rate ranged

Table 5.1: Country characteristics and response rates. GI sample

Countries	Ling. close [1]	Geogr. close [2]	Income (GNI/capita) [3]	Absence of migr. barriers [4]	Notes	Partici- pants	Response rate
Netherlands	Yes	Yes	High	Yes	Linguistically close country	139	0.67 [5]
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 [5]
South Korea	No	No	High	No	Developed, non-Europ. country	470	0.65 [5]
Mexico	No	No	Upper-middle	No	Emerging markets	491	0.60
Indonesia	No	No	Lower-middle	No	Developing country	883	0.55 [5]
India	Yes/No (English)	No	Lower-middle	No	Developing country	830	0.72 [5]

[1] Germanic languages.
[2] Countries in Europe.
[3] Categorized as by the World Development Indicator in 2018.
[4] Freedom of movement for workers within the EU and the European Economic Area.
[5] Response rates based on registered course participants, not actual attendance.

from 67 % to 99 %. In these countries (except the Netherlands), the response rate is based on the actually distributed number of questionnaires; in non-European countries (and the Netherlands), the response rate is based on registered course participants. The number of registered course participants is by definition equal or larger than course attendance in each unit. 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. In non-European countries, the response rate ranged from 59 % to 72 %. A further exception is Mexico where the response rate (60 %) is based on the number of distributed questionnaires during the process of course inscription.

5.2.2 Individual characteristics of language course participants

The individual characteristics of language course participants are very different across countries. Table 5.2 shows the means by country groups (see Table 5.A1 in the Appendix for the means by all 14 countries and Table 5.A2 for variable descriptions). We distinguish between European countries, which are members of the EU, European countries which do not belong to the EU and

non-European countries.[53] Of the 6,664 language course participants that took part in the survey, we exclude 2,308 observations because of missing information in the variables utilized in our analysis. Our main sample therefore contains 4,356 individuals. Among all participants 78 % are younger than 35 years. This share is higher for course participants in non-EU countries. The majority of course participants is female (59 %) and has no partner (60 %) 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 10 %, but in particular due to Great Britain with 28 % and the Netherlands with 18 %).

The young average age of the participants might be responsible for the low share of those with children (only 13 %). This might also explain the high share of those who indicate that they are in education (43 % overall, over 50 % in non-EU countries, but only 28 % in EU countries). Most of the other participants are active in the labour market (46 %). The majority of course participants already has a university degree (64 %). 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. On average, the level of international applicability of their education as evaluated by the participants is highest in the EU (4.02) on a scale from 1 to 5. This value is slightly smaller in non-European countries, but with much more variation between the countries ranging from a rather low international applicability (2.83) in Japan to a very high one in Mexico (4.32).

---

[53] The group of European, EU countries comprises the Netherlands, Great Britain, Spain, Italy, Czech Republic, Poland and Romania; the group of European, non-EU countries includes Ukraine and Bosnia and Herzegovina; finally Japan, Korea, Mexico, India and Indonesia belong to the group of non-European countries.

Table 5.2: Descriptive statistics: means of individual characteristics by country-groups

Variable	European (EU) n=1822	European (non-EU) n=683	Non-European n=1851	Total n=4356
Age: under 35 years	0.68	0.82	0.87	0.78
Gender: female	0.59	0.60	0.59	0.59
Children	0.18	0.17	0.07	0.13
No partner	0.44	0.58	0.77	0.60
Partner (native German)	0.10	0.03	0.03	0.06
Partner (other native)	0.46	0.39	0.20	0.34
Occ.: in education	0.28	0.51	0.54	0.43
Occ.: in lab. market	0.60	0.37	0.35	0.46
Other occ./no answer	0.12	0.12	0.11	0.12
University degree	0.76	0.58	0.55	0.64
Risk attitude	6.09	6.77	6.47	6.36
Patience	5.80	6.62	6.28	6.13
Intern. applic. of education	4.02	3.85	3.74	3.87
Migration intention	0.63	0.82	0.68	0.68

Notes: All variables except risk attitude, patience and international applicability of education are binary variables that take the value 1 for the indicated statement. Risk attitude is measured from 0 (risk-averse) to 11 (risk-loving). Patience is measured from 0 (very inpatient) to 10 (very patient). International applicability of education is measured from 1 (none of the education is internationally applicable) to 5 (all of the education is internationally applicable).

Participants in the EU are on average slightly less risk prone, but also slightly less patient. Intention to migration is quite high (overall 68 %). In particular this is the case in the non-EU countries Bosnia and Herzegovina (88 %), Ukraine (81 %) and Mexico (80 %). The opposite is the case in Japan where only 38 % have migration intentions.

5.3 Reasons for language learning

In this section, we present the reasons for learning language indicated by the course participants and explain how we aggregate them into categories and ultimately into the human capital motive and the consumption motive. A special focus is on cross-country differences with a special focus on possible difference between European and non-European countries and on heterogeneities based on age and gender.

5.3.1 Categorization of reasons

In the survey among course participants at the GI all respondents have in common that they decided to learn the German language. We now examine the reasons behind this decision. First, all participants answered the following multiple-response question: *Why are you learning German?* Afterwards, participants were asked to choose the main reason among the chosen reasons.

We categorize the main reasons according to Figure 5.1 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 second step, we further aggregate these categories into the two motives *human capital investment* and *consumption good*.

The categories *education* and *educational and labour migration* include only one single reason each. *Domestic labour market* gathers all reasons that include potential advantages in the labour market of the country where the survey took place. *Personals reasons* refers to *family* and *friends* as well as *other migration*, i.e. migration which is not related to labour market or education. The last category sums up *cultural interest* in a broad sense and includes holidays, cultural interests (films, literature,...) as well as general interests in languages.

On the basis of these categories, we group the reasons into motives. 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 one via *labour migration.* 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 affect the accumulation of other human capital positively, e.g. by opening up better education possibilities in destination countries where language skills are a requirement for education.

One 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 consumption goods which belong to the categories *personal reasons* and *cultural interest.*

When we use motives and categories we always build on the indicated main reason. However, the question about the main reason led to a substantial number of missing data points, as not all participants indicated their main reason among their chosen reasons. 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.[54,55]

5.3.2 Reasons, categories and motives

In the following, we present the distribution of the main reasons, the categories and the motives. The main sample is constructed in a way that we drop all observations where we neither have information about the respondent's main reason nor are able to impute the category or motive.

Figure 5.1: Categorization of reasons to learn languages.
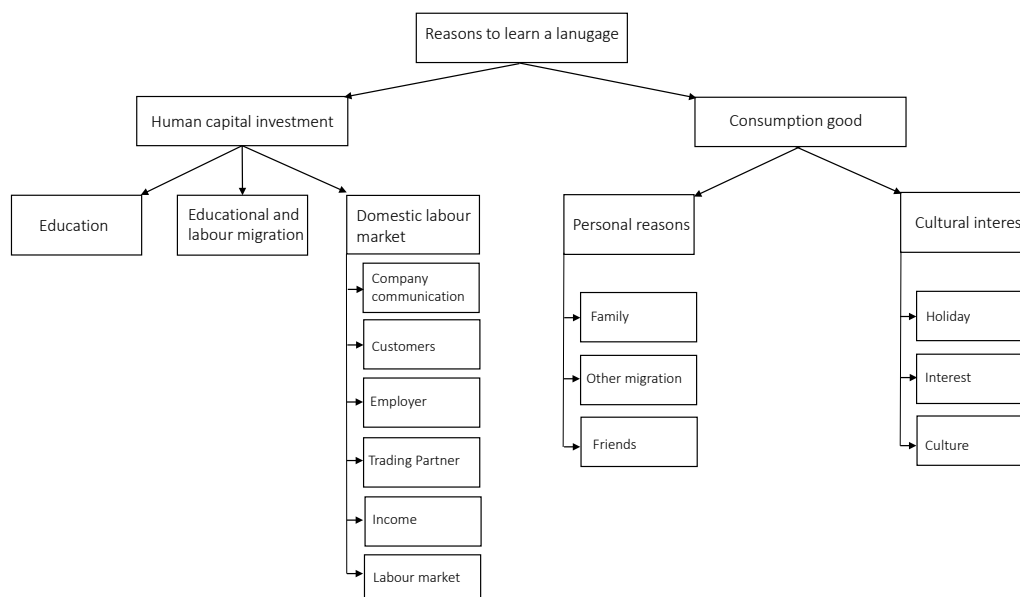
Note: See Table 5.A3 in the Appendix for exact the wording of the question in the questionnaire and Table 5.A4 for an overview of the abbreviated terms we use in the text and graphs for the different reasons. Note that in Japan, Bosnia and Herzegovina, Great Britain and Poland we have not distinguished 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. In Indonesia, we split the category "Study/education/training/PhD" into "Study/education/PhD" and "(Vocational) Training", but re-merged it for our analysis. Other reasons were categorized according to the free-text field if possible.

5.3.2.1 Reasons and main reasons

Figure 5.2 shows the share of participants for each of the single reasons. *Interest* is top of the list chosen by 58.7 % of the participants. Other frequently mentioned reasons are *Education* (46.1 %) and *Educational and labour migration* (41.3 %). Half of the respondents picked three or less reasons (on average 3.15 reasons) and only 25 % picked more than four.

Our categorization as presented in Figure 5.1 does not imply that the single reasons within each category need to be correlated. Also the relevance of cross-

---

54 The results are available from the authors upon request.

55 In order to increase sample size, we impute categories if the respondents gave reasons that belong only to one category. We apply the same procedure for the motives if all reasons belong either to the human capital motive or the consumption motive.

Figure 5.2: Reasons for learning German (n=4356).
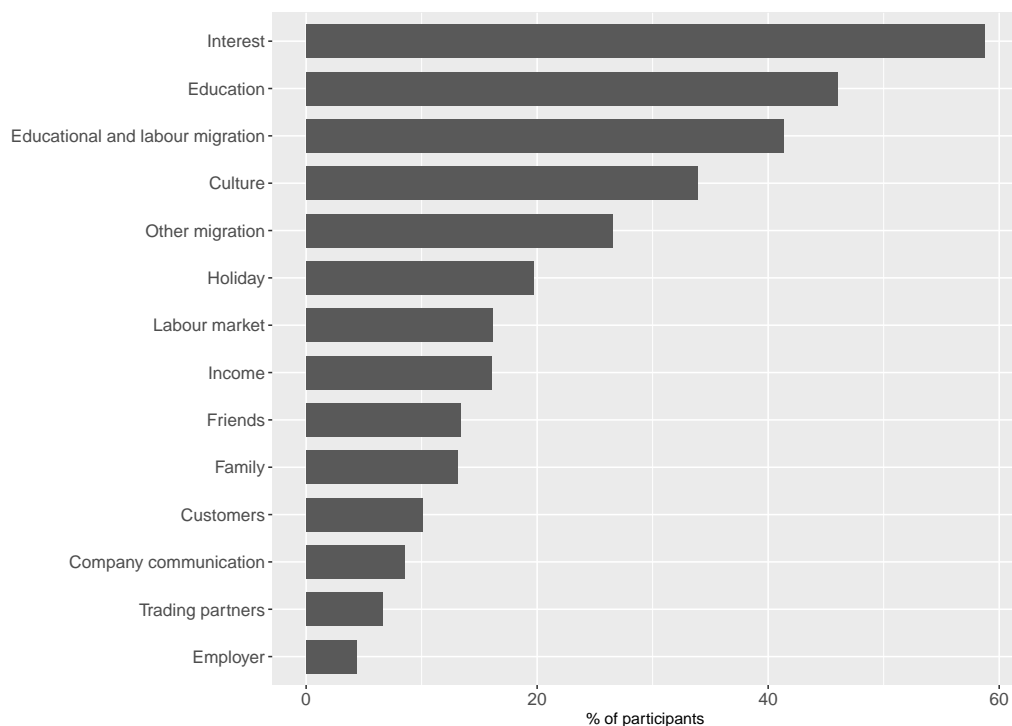
category correlation is not obvious. Therefore, a further look is warranted. We ignore those 809 participants who only indicated one single reason for learning German and focus on those with at least two reasons. Figure 5.3 maps the correlation coefficients between all reasons (see Table 5.A5 in the Appendix for the numbers and their statistical significance). We see that there is some correlation within some of our categories, while there is none for others. The correlation coefficients between all reasons range from -0.19 to 0.39.

For reasons within categories, we always find positive correlations. These are particularly large for reasons belonging to the *domestic labour market* and *cultural interest* categories (see the upper-left and the lower-right corner of Figure 5.3) and a bit less pronounced for the category *personal reasons*. At the same time, correlations are not necessarily large for reasons belonging to different categories even within the same motive. This shows for example when looking at the reason *family* and the reason *culture*. They are both part of the consumption motive; nevertheless their correlation is even negative. This holds a

forteriori when considering reasons belonging to different motives. Between the category *domestic labour market*, which is part of the human capital motive, and the category *cultural interest*, which belongs to the consumption motive, and their respective reasons, there is almost no statistically significant correlation. In some cases the relationship turns negative.

It is also interesting to note, that *family* is only positively correlated with *friends* and *other migration* while it is uncorrelated or even negatively correlated with all other reasons outside the category *personal reasons*. *Education* is a similar case which has only a positive and somewhat larger correlation with *educational and labour migration*. Overall, the correlations provide support for the way we aggregated reasons to categories and also to motives.

Figure 5.3: Correlations between reasons for learning German (n=4356).
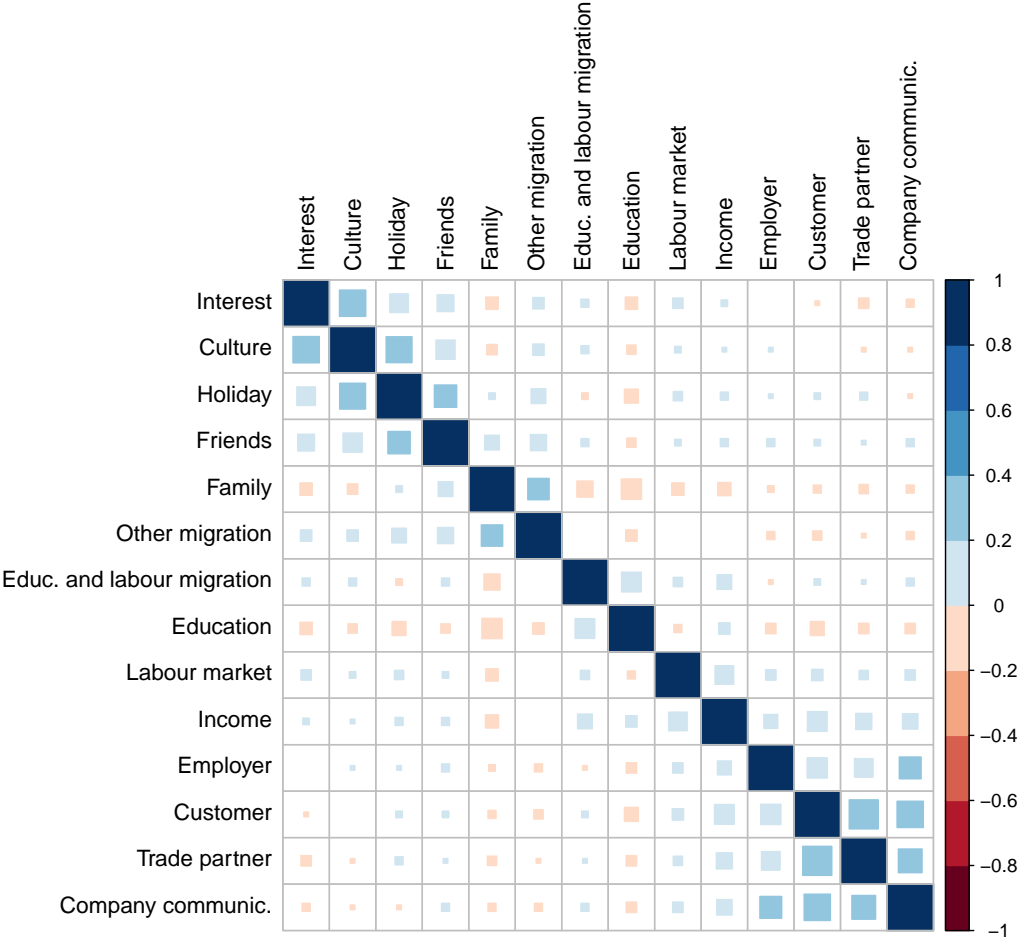
Notes: The size of the squares indicates the size of the correlation coefficient. Red squares indicate a negative correlation, blue squares a positive correlation. For statistical significance see Table 5.A5 in the Appendix.

We now have a look at the main reasons. Figure 5.4 shows their distribution in the sample. The largest group among the participants chose *education* (24.7 %) as main reason, which is more than half of the respondents that indicated *education* as one of possibly more reasons for learning a language (see Figure 5.2). This ratio is different for *interest* which comes second as main reason: while 58.7 % mentioned this as one of possibly more reasons, only 18.0 % indicated it as their main reason. Both reasons related to migration jointly constitute the largest group of main reasons, while separately viewed *educational and labour migration* ranks third (17.2 %) and *other migration* ranks fourth (7.6 %). The relative importance of *employers* and *trading partners*, on the contrary, seems to be rather limited both as reasons in general (4.4 % and 6.6 %) and as main reasons (0.8 % and 0.9 %). In addition, there seem to be reasons which play a role in general when learning a language, but are rarely chosen as main reasons. *Holiday* (19.7 % and 1.0 %) and *culture* (33.9 % and 4.1 %) are examples for this.

Figure 5.4: Main reasons for learning German (n=4356).
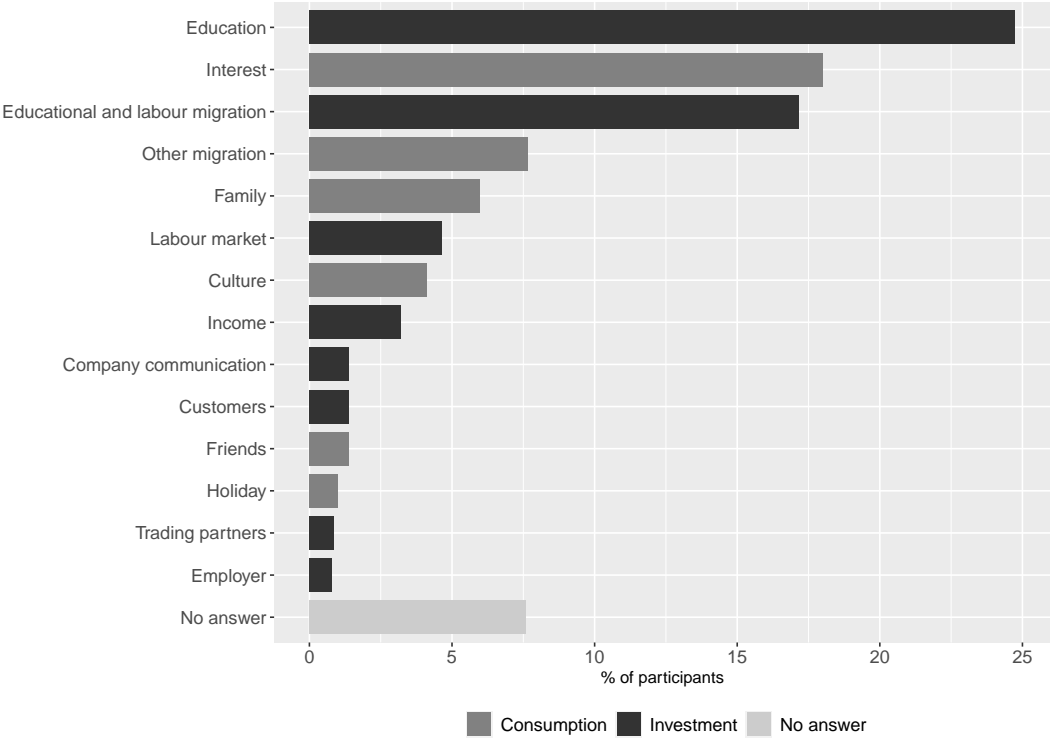
Figure 5.5: Main reasons for learning German by countries (n=4356).
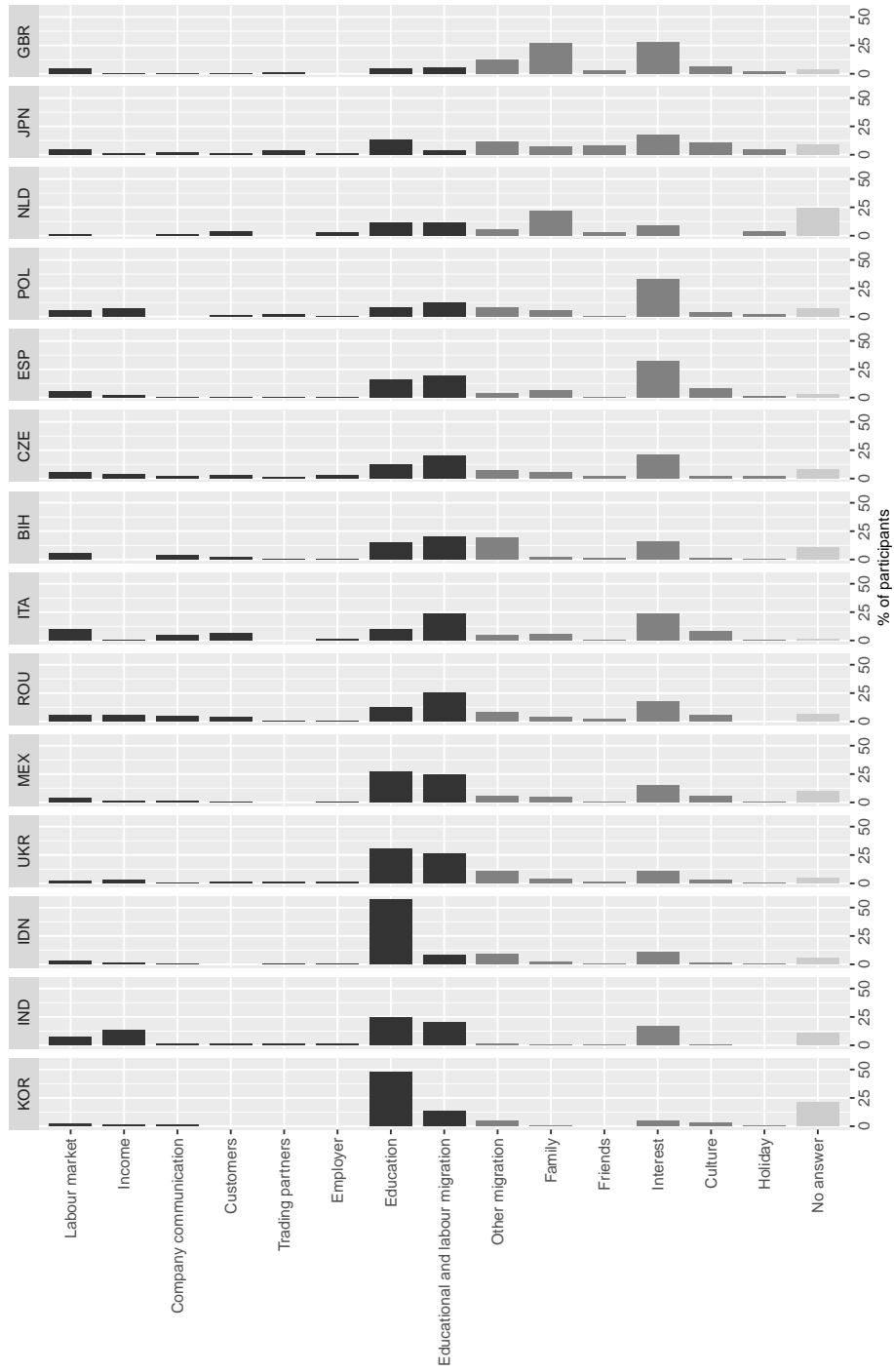
In Figure 5.5, we present an overview of the main reasons by countries and show that there is a large heterogeneity. This is in particular obvious when looking at *education* and *family*. The share or participants that indicated *education* ranges from 4.6 % to 57.2 %, the share that indicated *family* from 0.2 % to 26.8 %. This makes it evident that the overall shares are driven by only a few countries: the overall share of *family* would decrease by 2 percentage points if Great Britain was not in the sample. This is likely related to the relatively high share of partners who have German as native language in the British sample (see Table 5.A1 in the Appendix). Dropping Indonesia, the country with the highest share of *education*, would result in an even larger decline from 24.7 % to 20.0 % of the overall share.

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 the 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* as their main reason have the opposite in common: 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*, in the Netherlands it is *family*.

5.3.2.2 Categories

We have categorized the 14 main reasons into the five categories *education, educational and labour migration, domestic labour market, personal reasons* and *cultural interest*. Figure 5.6 shows the share for each of the categories in the aggregated sample. One quarter each indicated that either *education* or *cultural interest* was their motivation behind their decision to study German, followed by *educational and labour migration* (17.2 %), *personal reasons* (15.5 %) and *domestic labour market* (12.9 %).

In Figure 5.7, we can again see the heterogeneity across the countries similar to what we observed for the main reasons: shares of *personal reasons* range from 2.5 % in India to 42.7 % in Great Britain for example. Also the distribution of shares within countries is very diverse. We find countries where the five

Figure 5.6: Categories for learning German (n=4356).
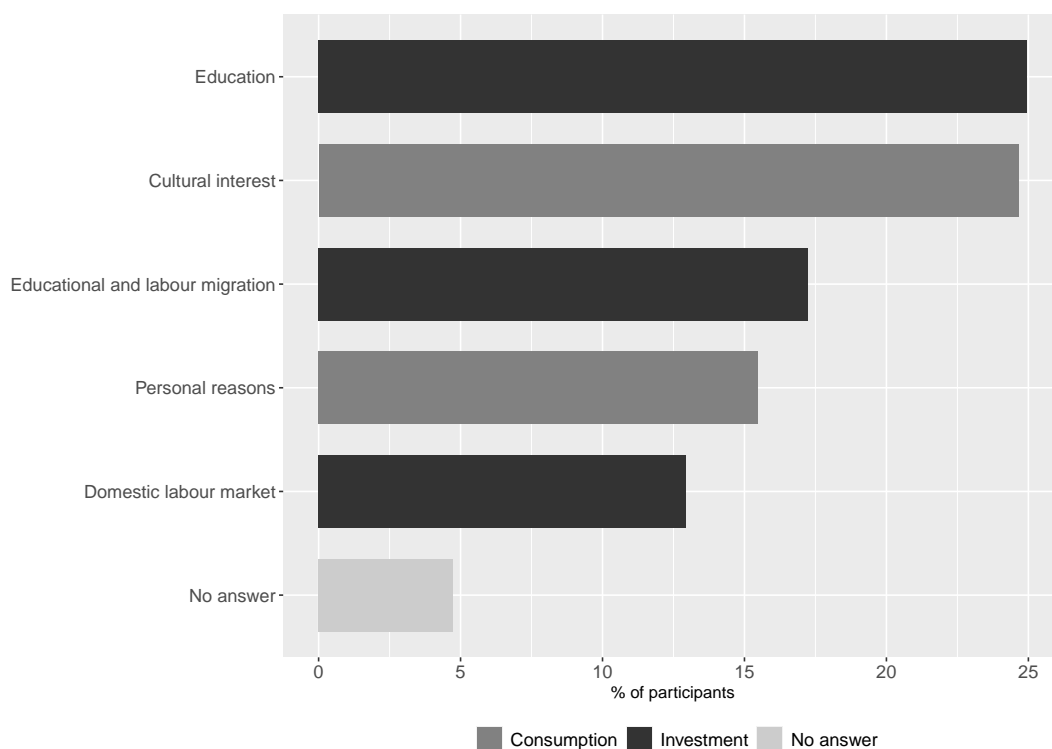
categories are relatively evenly distributed like Bosnia and Herzegovina, the Czech Republic or Romania. The opposite is the case for India, Great Britain and Korea, where the shares for one or two categories are much larger than for the other categories.

Figure 5.7 confirms mainly the results we have found in Figure 5.5 when looking at the most important categories for each country. 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 for *interest* has translated into the category *cultural 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.

Figure 5.7: Categories for learning German by countries (n=4356).
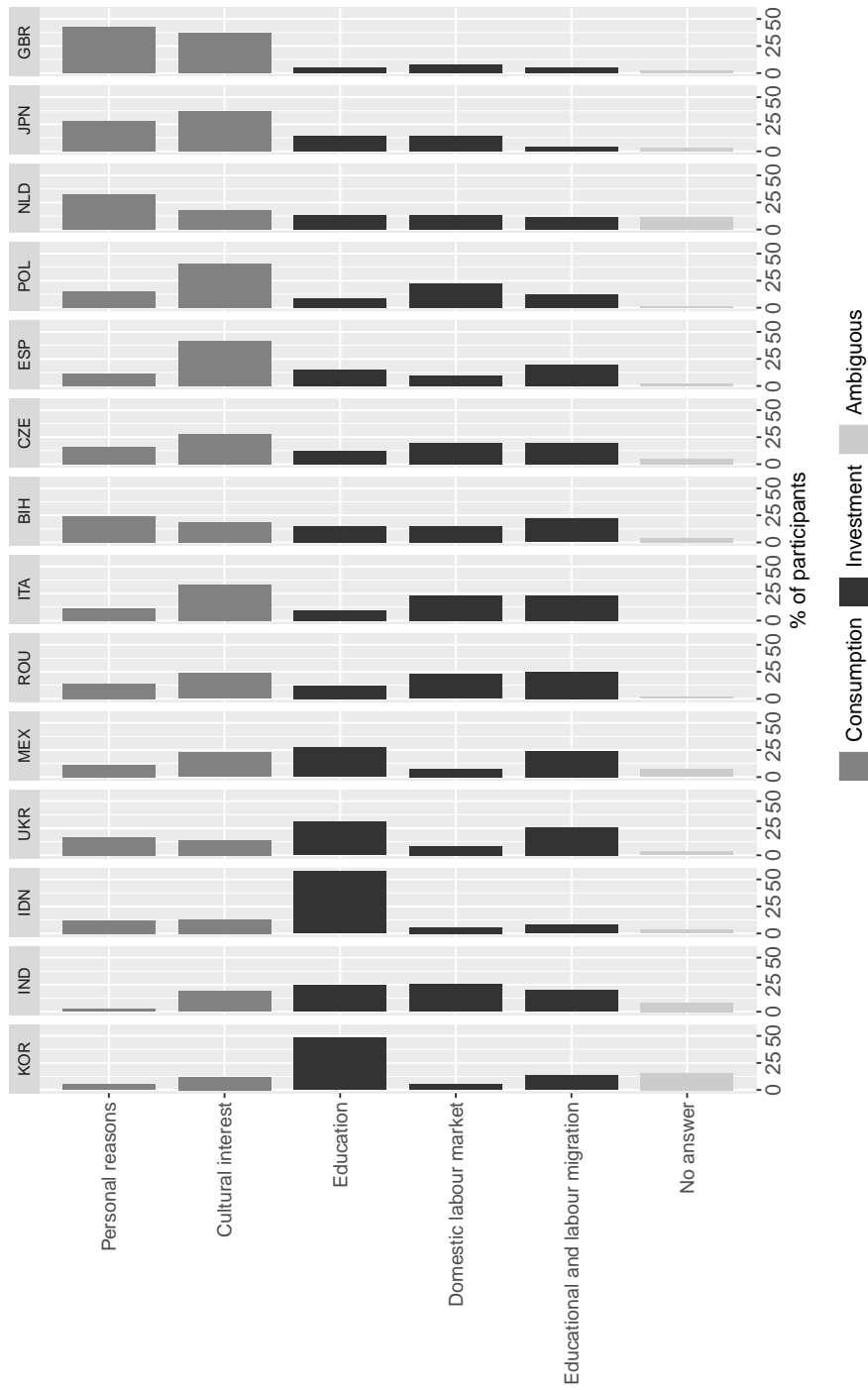
5.3.2.3 Motives

As the final step, we focus on the two motives: human capital investment and consumption. While language learning because of a main reason that is categorized as *cultural interest* or *personal reasons* is seen as a *consumption good*, a reason which belongs to the categories *education, educational and labour migration* or *domestic labour market* makes language learning a human capital investment. Figure 5.8, upper part, gives the distribution of the investment and consumption motives by countries. The countries are arranged in descending order according to the share of participants indicating a main reason categorized as human capital investment. The variation across countries is large and the share of human capital investment as main motive ranges from 19.2 % to 82.1 % (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. In these countries, the largest categories *education* (in Korea, India, Indonesia, Ukraine and Mexico) and *educational and labour migration* (in Romania) translate into the importance of language learning as a human capital investment. 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. There, the most important categories *cultural interest* (in Italy and the Czech Republic) and *personal reasons* (in Bosnia and Herzegovina), which belong to the consumption motive, are of slightly less importance than the investment motive. Third, in Spain, Poland, the Netherlands, Japan and Great Britain the consumption motive is more important than the investment motive. The most important categories *cultural interest* (in Spain, Portugal and Japan) and *personal reasons* (in the Netherlands and Great Britain) reflect this.

Before we look in more detail at the different determinants behind the different motives of the decision to learn German in Section 5.4, we give a first descriptive overview of the role of age and gender for the cross-country differences in Figures 5.8 where the middle part presents the distribution of the two motives across countries by age groups and the lower part by gender.

We start with possible 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 cases, 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 that the share of older people is relatively low in some countries (see Table 5.A1 in the Appendix). This holds for example for India with a share of only 5 % compared to the average share of 22 % in the entire sample.[56]

When looking at the younger age group, the general pattern is relatively close to the full sample for the three groups of countries. The investment motive dominates in most cases with the exception of Great Britain.

The pattern for age cannot be transferred to gender. There are no strong differences in the distribution of the human capital motive and the consumption motive between male and female participants in most countries. If at all, the investment motive seems to be slightly more important for men. This shows up in the three countries of the second group (Italy, Bosnia and Herzegovina and the Czech Republic). Another exception is Poland, where the majority of female participants indicates consumption motives while the opposite is the case for male participants.

In contrast to age, the composition of participants with respect to gender across countries is much more homogeneous. Overall 59 % of the participants are female. While the share of women is a bit larger e.g. in Japan (68 %), Korea (67 %) and Romania (67 %), in Bosnia and Herzegovina, Great Britain, Indonesia and Mexico the shares are almost equal for men and women.

Overall, we conclude that differences in the gender 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 next section 5.4, we will investigate the within-country variation.

---

[56] Shares of older people are also below average in Indonesia with 5 % and in Korea with 7 %, while in Mexico and Ukraine, the share of older participants is slightly smaller than the average with 17 %. Shares are particularly large in Japan (58 %) and Great Britain (48 %).

Figure 5.8: Human capital and consumption motives by countries, age and gender (n=4356).
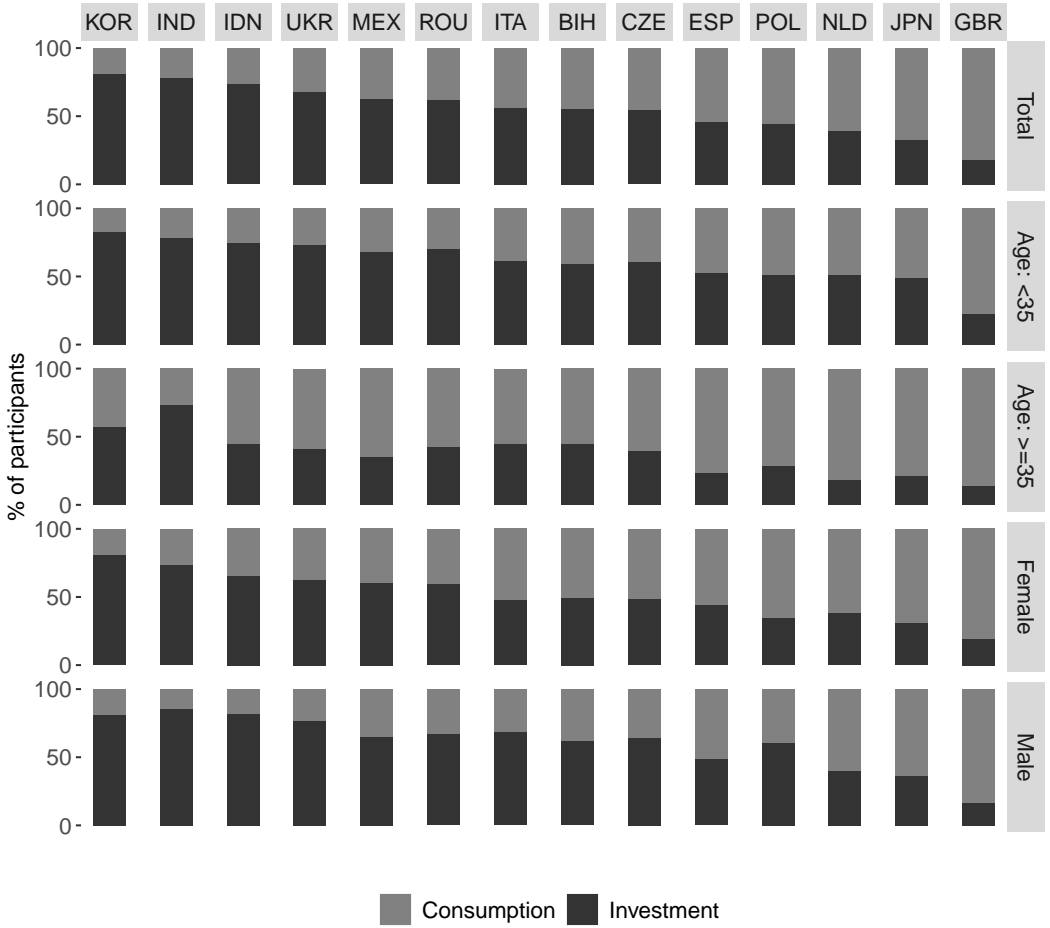
5.4 Determinants of the human capital motive

5.4.1 Estimation strategy

We explore individual-specific and country-specific determinants of the human capital investment motive. In particular, we are interested if the descriptive evidence found above for age and gender continues to hold after controlling for other factors. Furthermore, we check the robustness of our results for age and gender subgroups. We estimate the probability of human capital investment motive via maximum likelihood method in a binary probit model:

$$Pr(investment_i \mid X_i, C_i) = \alpha + \beta \,' X_i + \gamma \,' C_i + \varepsilon_i \qquad (5.1)$$

where $investment_i$ takes a value of 1 if respondent i states to have a human capital investment motive and 0 otherwise. X_i represents a set of individual-specific explanatory variables of respondent i as presented in Table 5.A2 in the Appendix. C_i captures either country-level factors or country fixed effects in order to control for heterogeneity between countries. Country-level characteristics are binary control variables that are based on the characteristics as described in Table 5.1 and include linguistic closeness by indicating if the official language is a Germanic language, European, non-EU countries and non-European countries (with European, EU countries as reference),[57] and upper-middle and high income countries (with lower-middle income countries as reference). ε_i is an idiosyncratic error term. Standard errors are heteroscedasticity robust White standard errors.

In the following, we discuss the association between the individual-specific and country-specific explanatory variables on the one hand and the probability of the human capital investment motive as the main motive versus the consumption motive 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).

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

---

[57] See Table 5.1 for the countries and country-groups.

is learning for consumption purposes, and not "no learning" at all. In the absence of monetary benefits and for given monetary costs, as we have it with the consumption motive, the individual financial means are important. This is an important factor we cannot control for on the individual level, but we consider this on the country level.

5.4.2 Estimation results

Table 5.3: Estimation results: basic specifications

	Dependent variable: Human capital investment					
	(1)		(2)		(3)	
Age: under 35 years	0.190***	(0.024)	0.131***	(0.024)	0.154***	(0.024)
Gender: female	−0.074***	(0.014)	−0.087***	(0.013)	−0.077***	(0.014)
Children	−0.019	(0.026)	−0.060**	(0.026)	−0.039	(0.026)
Partner (native German)	−0.451***	(0.026)	−0.371***	(0.033)	−0.406***	(0.030)
Partner (other native)	−0.050***	(0.017)	−0.027	(0.017)	−0.031*	(0.017)
Occ.: in education	0.146***	(0.021)	0.113***	(0.021)	0.134***	(0.021)
Other occ./no answer	−0.009	(0.022)	−0.009	(0.021)	−0.010	(0.021)
University degree	0.014	(0.018)	0.037**	(0.018)	0.036**	(0.018)
Risk attitude	0.007**	(0.003)	0.003	(0.003)	0.002	(0.003)
Patience	0.011***	(0.003)	0.006*	(0.003)	0.008***	(0.003)
Intern. applic. of educ.	0.029***	(0.006)	0.039***	(0.007)	0.035***	(0.007)
Migration intention	0.070***	(0.015)	0.059***	(0.015)	0.062***	(0.016)
Germanic language					−0.100***	(0.019)
European (Non-EU)					0.002	(0.029)
Non-European					0.080***	(0.020)
Upper-middle income					−0.075***	(0.024)
High income					−0.110***	(0.024)
Country FE	No		Yes		No	
McFadden Pseudo R2	0.14		0.17		0.15	
Percent. correctly predicted	69.7		71.6		70.2	
Observations	4,356		4,356		4,356	

Average marginal effects. Heteroscedasticity robust White standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

We present our main results in Table 5.3. 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. Furthermore, some variables turn insignificant, e.g. risk attitude. This hints at a large variation in risk attitude across countries which does not translate into a significant relationship within countries.

Before focusing on the individual-specific 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. Considering that English is almost worldwide a Lingua Franca and the most spoken foreign language in the world, the benefits of learning a foreign language seem to be rather small. This also holds to a certain extent for Dutch which allows its speakers an easier access to English (and also German).

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 European countries close to Germany which are members of the EU and for which migration restrictions are non-existent. The second group are European countries which are not members of the EU. The geographic distance to Germany is still rather small, but migration to Germany, Austria and Switzerland is much more restricted. Similar migration restrictions hold for the third group of non-European countries. In addition, the geographic distance to German-speaking countries is much larger. For the latter group of countries the probability of human capital investment motives is significantly larger than for 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 (or non-European) countries. This makes it more likely for language course participants to acquire language skills for investment purposes, as we saw in Figure 5.8. Second, geographic proximity, which is given for European EU and non-EU countries can be a proxy for two other measures: cultural proximity and a larger existing migrant stock in German-speaking countries due to migration flows in the past. Both make the consumption motive of language learning more likely for these countries. There are thus possible explanations for the observed differences between non-European and European (EU or non-EU) countries.

Finally, the country-wide income level plays an important role: participants from upper-middle and higher income countries are on average less likely to learn German for investment motives. 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 5.1) which is reflected in the results here.

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 of the results in the following and for the estimations by age and gender subgroups in Table 5.4.

In the human capital theory, age is an important factor with a negative effect on human capital investment. To say it differently, the older the individual is the less time there is to recoup the investment and benefit from it. Analogously, the older the participant in a language course, the less time for the returns to realize. In addition, the costs of learning a language increase with age as the required effort increases. This can be seen in the literature on the language acquisition of immigrants in their host country (see the three Es in Chiswick and Miller, 2015, (especially the "E" standing for "efficiency")) and explains the 13.1 % point increase in the probability of investment motives for younger age groups. This relationship is robust within gender subsamples.

Female participants have on average a lower probability of an investment motive than male participants. The incentive to acquire foreign language skills as human capital investment depends on the potential benefits. The situation of women on the labour market is often worse in terms of labour market participation and wages. Furthermore, in the migration context women are more likely to be the tied mover who joins the male labour migrant with a lower own probability of labour market participation. This relationship is stronger for the younger age group.

Having a native German partner reduces the probability of an investment motive in comparison to singles and those with a partner with another native language. The size of the average marginal effect is 37.1 % point and thus almost three times as large as the age effect of the younger-age group. A native German partner increases the opportunities where consumption of German language skills seems to be more likely, e.g. communication with their partners and their families as well as with friends. This relationship is robust within all subsamples.

The probability of having an investment motives is larger by 11.3 % point for course participants in education compared to those who are in the labour mar-

ket. Being used to study likely decreases the cost of acquiring language skills by increasing the efficiency of learning. Furthermore, foreign language skills might increase the productivity of other skills in a complementary way and therefore positively affect the overall benefits of this human capital investment. Not surprisingly, this relationship does not hold for participants in the older age group. Having a university degree increases the probability of having an investment motive. This is mainly driven by younger course participants. International applicability of education makes it more likely that opportunities on the foreign and domestic labour market emerge where language skills lead to benefits. The results point towards a robust positive relationship with the probability of having investment motives.

Human capital investment is a risky endeavour as the benefits are not certain. However, risk attitude it is not significantly related to the probability of an investment motive in the main specifications once we control for country-fixed effects (or country characteristics). Looking at the subsamples, we find a positive relation only for the older age group. Patience is an important characteristic with regard to human capital investments, where benefits realize much later – if they realize at all. Therefore, a positive correlation between patience and the probability of having an investment motive is not a surprise. However, this seems to be due only to male participants.

Last but not least, migration intention is positively associated with the probability of investment motives. This stresses the role of investment-related migration in contrast to consumption-related migration. This correlation seems to be due to males and younger participants. Both groups are more likely to be migrants with a labour-market focus.

Table 5.4: Estimation results: age and gender subsamples

	Dependent variable: Human capital investment							
	(1)		(2)		(3)		(4)	
	female		male		age: ¡ 35 years		age: ≥ 35	
Age: under 35 years	0.115***	(0.032)	0.147***	(0.036)				
Gender: female					−0.096***	(0.015)	−0.054*	(0.029)
Children	−0.083**	(0.036)	−0.034	(0.038)	−0.015	(0.045)	−0.057*	(0.032)
Partner (native German)	−0.365***	(0.038)	−0.379***	(0.064)	−0.407***	(0.045)	−0.257***	(0.034)
Partner (other native)	−0.029	(0.022)	−0.026	(0.026)	−0.024	(0.019)	−0.019	(0.035)
Occ.: in education	0.131***	(0.028)	0.090***	(0.031)	0.144***	(0.022)	0.365	(0.270)
Other occ./no answer	−0.003	(0.028)	−0.006	(0.032)	0.062**	(0.026)	−0.117***	(0.034)
University degree	0.037	(0.024)	0.041	(0.027)	0.061***	(0.020)	−0.058	(0.055)
Risk attitude	0.003	(0.004)	0.003	(0.005)	−0.001	(0.004)	0.015**	(0.006)
Patience	0.003	(0.004)	0.010**	(0.004)	0.004	(0.003)	0.010	(0.006)
Intern. applic. of educ.	0.034***	(0.009)	0.047***	(0.010)	0.039***	(0.008)	0.039***	(0.013)
Migration intention	0.039*	(0.020)	0.086***	(0.024)	0.053***	(0.018)	0.060*	(0.031)
Country FE	Yes		Yes		Yes		Yes	
McFadden Pseudo R2	0.15		0.21		0.12		0.15	
Percent. correctly predic.	69.1		74.8		71.31		70.1	
Observations	2,572		1,784		3,395		961	

Average marginal effects. Heteroscedasticity robust White standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Figure 5.A3 graphically displays the results by country-groups. At first glance, the coefficients and their level of significance look very similar. A closer look reveals some differences, however. The investment motive in Germanic countries is not correlated to the age group and neither to the degree of international applicability of education, differently from all the other groups. Focusing on country-groups by income and geographic closeness or ease of movement, respectively, we find a large similarity between lower-middle income countries and non-European countries. Also upper-middle and high income countries share many similarities. These are also the only two groups where are university degree is positively and significantly related to the investment motive.

5.5 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 a professional environment. More precisely, they answered to the question "How likely is it that you will use your German language skills in your job?" on a scale from 1 (very unlikely) to 5 (very likely). Foreign language skills have a pro-

ductive value if opportunities arise to use them in a professional environment, independent from the main reason behind the decision to learn a language. In Section 5.3, we showed that the reasons for learning are very diverse and refer to different motives. However, within these motives the probability of using the language skill in a job is not always ex-ante clear, as this is very much driven by opportunity and expectation.

For illustration, let us look at the two migration related reasons *educational and labour migration* and *other migration* where the probability of a professional use might diverge. On the one hand, *other migration*, which belongs to the consumption motive (see Figure 5.1), might lead to a professional use, if individuals migrate as tied movers. At first, the reason to acquire foreign language skills follows from a consumption motive, because the skills increase the utility by facilitating family life or partnership. Only later, when it comes to labour-market participation, the foreign language skills might also be used in a professional environment. Then, the investment aspect of language learning comes into the picture, which we can interpret as spillover effects 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 investment motive.

In order to identify what makes the use of German language skills in the labour market more likely we estimate its determinants and try to identify potential spillovers from the consumption motive in Section 5.5.2. Before, we present descriptive evidence on the heterogeneity of the probability of professional use of German by countries and subgroups.

5.5.1 Descriptive evidence

The probability of using German in the labour market is on average quite high with 3.68 on a scale from 1 to 5. As shown in Figure 5.A1 in the Appendix, the distribution in the overall sample is left-skewed with a median of 4 and the mode of 5 indicated by 36,9 % of the participants. In the following we aggregate the five point scale to a binary variable where values 1 to 3 are aggregated to "unlikely" and values 4 and 5 to "likely".

Around 60 % indicated that they will likely use German in the labour market and 40 % that the will likely not do so. 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 subsamples in the following. Deviations can be interpreted as spillover effects or they can be due to multiple reasons for language learning among which only one is indicated as main reason. In Section 5.5.2, we try to shed more light on that.

The distribution again varies across countries as presented in Figure 5.9, upper part. Remember that the countries are ordered by their share of the human capital 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 (see also Figure 5.8.) 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 likelihood 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, has with Spain and Poland two countries, where we expect the share of those with a high likelihood 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 it is the probability of professional use is smaller in Japan than expected. There seem to be factors in play that hinder those with a human capital investment motive to think that they will be able to use German in an work-related context, and vice versa make those with the consumption motive expect a professional use of their foreign language skills. Before we examine in detail possible determinants with binary probit regressions, we again look at the distributions by age and gender and by motive.

Figure 5.9, middle part, shows that the pattern for the younger age group closely follows the pattern for the total sample. This does not hold for older participants. In the non-European countries Korea, India, Indonesia and Mexico the share of older participants indicating a low likelihood of using German in the labour market is larger than the share of those with a high likelihood. This stands in contrast to the overall sample. On the contrary, in the European

Figure 5.9: Use of German on the labour market by countries, age and gender.
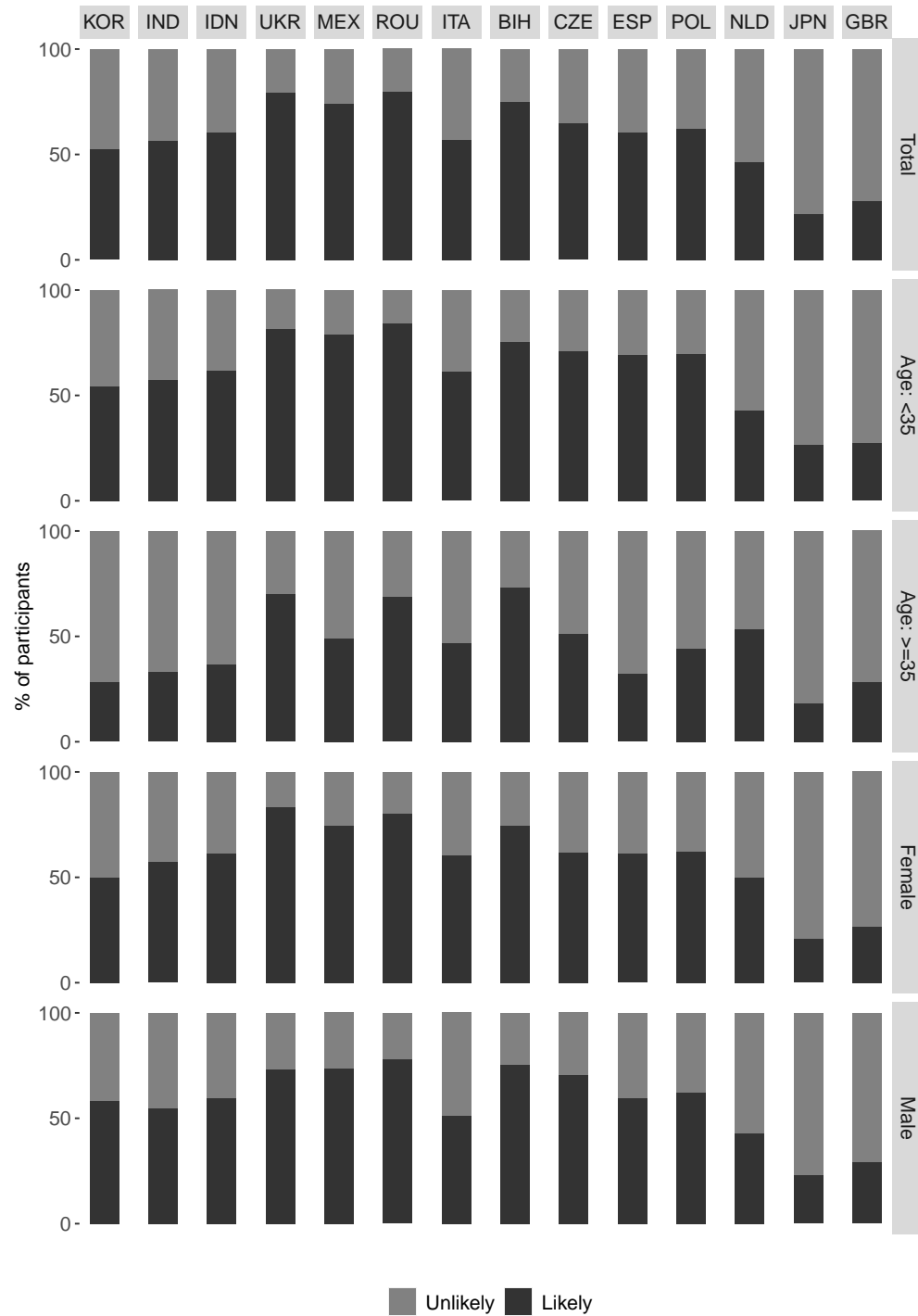
countries where the share of human capital investment motives is predominant (Ukraine and Romania), the pattern is similar among both age groups. While the same holds for Bosnia Herzegovina, the share of a low likelihood for the older age group is larger in Italy and the Czech Republic. Finally, in the countries where consumption motives are predominant the likelihood of using German in the labour market has a negative association with age in Spain and Poland , but not in Japan, Great Britain and the Netherlands. This emphasizes the role of age for the likelihood of using German in the labour market. 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 (Figure 5.9, lower part).

At the beginning, we assumed as a benchmark that there is a perfect relationship between having a human capital investment motive and a high likelihood of using German in the labour market. The graphical analyses above provided some evidence that there is no perfect correlation between those two variables. Figure 5.A2 shows, however, that there is indeed a positive correlation between those two. In all countries, the share of those who have human capital investment motives is larger among the participants with high likelihood for professional use than among the participants with low likelihood. We can see that the correlation is, in particular, strong in Mexico, Japan and Great Britain, as different motives lead to very different probabilities of professional use. On the contrary, the correlation is very weak in Korea, India and Indonesia; there is hardly a difference in the indicated probability of professional use between those with an investment motive and those with a consumption motive.

In the following section, we will investigate the determinants of this imperfect relationship, or to put it differently, we want to understand 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 consumption motives to the labour market.

5.5.2 Determinants of professional use of German

5.5.2.1 Estimations strategy

We explore individual-specific and country-specific determinants of the professional use of German. As in Section 5.4.1, we examine if the descriptive evidence we have seen for age and gender is robust also after controlling for

other relevant factors. Furthermore, we check the robustness of our results for age and gender subgroups. Additionally, we focus on reasons behind language learning as an additional determinant and as the basis for a further subgroup analysis. For this, we estimate the probability of having a high probability of professional use of the German language in a binary probit model via maximum likelihood method:

$$Pr(use_i|\ X_i, C_i) = \alpha + \beta\ '\ X_i + \gamma\ '\ C_i + \varepsilon_i \qquad (5.2)$$

where use_i takes a value of 1 if respondent i states a high probability for professional use of German and 0 otherwise. X_i represents a set of individual-specific explanatory variables of respondent i as used before in Section 5.4.2 and presented in Table 5.A2 in the Appendix. C_i are either country-level factors or country fixed effects to control for heterogeneities across countries. Country-level characteristics are binary control variables that are based on the characteristics as used in Section 5.4.2 and described in Table 5.1. ε_i is an idiosyncratic error term. Standard errors are heteroscedasticity robust White standard errors.

In the following, we discuss the association between the individual-specific and country-specific explanatory variables and the probability of a professional use of German and compare the results to the determinants of human capital investments in Section 5.4.

5.5.2.2 Estimation results

We present our main results in Table 5.5. Column 1 includes the same individual-specific characteristics and country fixed effects as in Table 5.4, Column 2 additionally adds the investment motive dummy and Column 3 includes dummies for the categories *educational and labour migration, domestic labour market, personal reasons* and *cultural interest* with *education* as reference category instead of the investment dummy. Column 4 re-estimates Column 2 with country characteristics instead of country fixed effects.[58]

When adding motives in Table 5.5 Column 2 and 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

---

[58]   Note that the sample is slightly smaller in Columns 3 due to the imputation of the categories as described in Section 5.3.1.

Table 5.5: Estimation results: basic specifications

	Dependent variable: Professional use of German							
	(1)		(2)		(3)		(4)	
Investment			0.191***	(0.016)			0.200***	(0.016)
Domestic labour market					0.054**	(0.025)		
Educational and labour migr.					0.005	(0.023)		
Personal reasons					−0.163***	(0.026)		
Cultural interest					−0.176***	(0.022)		
Age: under 35 years	0.083***	(0.023)	0.056**	(0.022)	0.050**	(0.023)	0.070***	(0.022)
Gender: female	0.017	(0.014)	0.034**	(0.014)	0.034**	(0.014)	0.040***	(0.014)
Children	−0.036	(0.026)	−0.025	(0.026)	−0.027	(0.026)	−0.012	(0.025)
Partner (native German)	0.027	(0.032)	0.083***	(0.030)	0.094***	(0.032)	0.058*	(0.029)
Partner (other native)	0.030*	(0.018)	0.037**	(0.017)	0.035**	(0.018)	0.032*	(0.017)
Occ.: in education	0.115***	(0.021)	0.092***	(0.021)	0.105***	(0.021)	0.100***	(0.021)
Other occ./no answer	0.026	(0.022)	0.027	(0.022)	0.031	(0.022)	0.022	(0.021)
University degree	−0.008	(0.019)	−0.015	(0.018)	−0.008	(0.019)	−0.010	(0.018)
Risk attitude	0.018***	(0.003)	0.017***	(0.003)	0.017***	(0.003)	0.016***	(0.003)
Patience	0.005*	(0.003)	0.004	(0.003)	0.004	(0.003)	0.005*	(0.003)
Intern. applic. of education	0.049***	(0.007)	0.042***	(0.007)	0.042***	(0.007)	0.044***	(0.007)
Migration intention	0.085***	(0.016)	0.073***	(0.016)	0.074***	(0.017)	0.072***	(0.016)
Germanic lang.							−0.131***	(0.020)
European (Non-EU)							−0.012	(0.031)
Non-European							−0.162***	(0.019)
Upper-middle income							0.083***	(0.023)
High income							−0.127***	(0.024)
Country FE	Yes		Yes		Yes		No	
McFadden Pseudo R2	0.13		0.15		0.15		0.11	
Percent. correctly predicted	68.6		70.8		70.8		70.6	
Observations	4,356		4,356		4,150		4,356	

Average marginal effects. Heteroscedasticity robust White standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

significant, e.g. being female or having a native German partner. As we expected, having an investment motive increases the probability of a professional use of German significantly by 19.1 % points. Categories that belong to the consumption motive are negatively associated compared to the reference category *education*, which belongs to the investment motive. Within the investment categories, the category *domestic labour market* is significantly more correlated than education, i.e. 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* and *educational and labour migration* categories.

When including country characteristics instead of fixed effects, the coefficients for individual characteristics remain qualitatively unchanged. The dummy for Germanic language is negative and similar to the relationship with the human capital investment motive (cf. Table 5.3). The same holds for European,

non-EU countries, for which the relationship is insignificant, and high-income countries, which show a positive relationship. Non-European countries, on the contrary, show now a negative and significant correlation, while it was positive in the specification for the human capital motive (cf. Table 5.3). While on average, respondents from non-European countries are more likely to indicate a human capital investment motive, they, at the same time, assess the probability of a professional use of German to be lower than respondents from the reference group of European, EU countries. The same pattern can be seen, when comparing lower- and upper-middle income countries where the lower middle income countries, which constitute the reference group, are more likely to have an investment motive, but less likely to professionally use German.

The results for individual characteristics are mainly in line with previous results on the human capital investment motive (see also Table 5.A7 for subgroups by age and gender), 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. Second, while having a partner with German as native language is associated negatively with the investment motive, we find for the professional use of German a positive and significant relationship, but more pronounced – once again – among younger respondents. Third, risk attitude turns into a positive and significant relationship, while the coefficient was not significantly different from zero when estimating the investment motives. This result is is not there for older respondents, however.

These three individual characteristics are important determinants for spillovers from a consumption motive to a professional use of German (see Table 5.6). Focusing on Column (3), being in the younger age group, female and with a native German partner as well as being still in education where the international applicability of education is high (but without a university degree) leads 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.

Table 5.6: Estimation results: subsamples 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	0.119***	(0.031)	−0.052	(0.033)	0.084***	(0.030)	−0.052	(0.033)
Gender: female	0.033	(0.023)	0.027	(0.017)	0.044**	(0.022)	0.027	(0.017)
Children	−0.004	(0.034)	−0.038	(0.044)	0.013	(0.033)	−0.038	(0.044)
Partner (native German)	0.100***	(0.038)	0.037	(0.090)	0.118***	(0.036)	0.036	(0.090)
Partner (other native)	0.035	(0.027)	0.038*	(0.022)	0.024	(0.027)	0.038*	(0.022)
Occ.: in education	0.113***	(0.035)	0.097***	(0.026)	0.091***	(0.035)	0.097***	(0.026)
Other occ./no answer	0.021	(0.031)	0.054*	(0.030)	0.039	(0.030)	0.054*	(0.030)
University degree	−0.055*	(0.031)	0.011	(0.023)	−0.063**	(0.030)	0.011	(0.023)
Risk attitude	0.013**	(0.005)	0.019***	(0.004)	0.011**	(0.005)	0.019***	(0.004)
Patience	0.0005	(0.005)	0.006*	(0.004)	−0.0001	(0.004)	0.006*	(0.004)
Intern. applic. of education	0.053***	(0.010)	0.031***	(0.009)	0.049***	(0.010)	0.031***	(0.009)
Migration intention	0.077***	(0.024)	0.069***	(0.020)	0.049**	(0.023)	0.069***	(0.020)
Other investm. reason					0.214***	(0.024)	0.001	(0.018)
Country FE	Yes		Yes		Yes		Yes	
McFadden Pseudo R2	0.08		0.15		0.08		0.18	
Percent. correctly predicted	72.4		70.0		72.4		71.2	
Observations	1,821		2,535		1,821		2,535	

Average marginal effects. Heteroscedasticity robust White standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

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 those not the main one).

With respect to the different country-groups, some patterns emerge (see Figure 5.A4). First, Germanic countries stand out as individual characteristics there mostly do not play a role when it comes to the professional use of German language skills. Second, European, EU countries and upper-middle and high-income countries are very similar and the same can be said about European, non-EU and non-European countries, lower-middle income countries and to a certain extent for the group of non-Germanic countries. Especially being young and female and having a native German partner is of importance in the latter groups of countries and not in the former ones adding more insights to this observation.

5.6 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. To get the full picture, we enlarge the human capital framework by adding consumption motives for foreign language learning. Based on a new dataset collected in 14 countries worldwide from language course participants, we are able to 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, human capital motives are particularly interesting for course participants in the context of trade as well as educational and labour migration. Policy measures targeting at this group, such as subsidies for foreign language learning or, in general, more language-learning opportunities, should therefore be one focus.

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. These spillovers are particularly relevant for those who indicate *other migration* as their main reason. 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 to learn a foreign language, a professional use of the language is not unlikely. As a consequence, policy measures aiming at this group then 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 a visa 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 has a clear focus on language learning as human capital investment. It 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. 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 prospects for economic and social integration.

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, do we get the full picture.

Appendix to Chapter 5

5.A   Appendix A

Table 5.A1: Descriptive statistics: means of individual characteristics by countries

Variable	BIH	CZE	ESP	GBR	IDN	IND	ITA	JPN	KOR	MEX	NLD	POL	ROU	UKR	Total
Age: under 35 years	0.74	0.69	0.77	0.52	0.95	0.97	0.68	0.42	0.93	0.83	0.64	0.7	0.69	0.83	0.78
Gender: female	0.51	0.63	0.57	0.51	0.52	0.61	0.6	0.68	0.67	0.53	0.55	0.63	0.66	0.62	0.59
Children	0.21	0.24	0.11	0.2	0.05	0.04	0.14	0.25	0.03	0.08	0.18	0.17	0.24	0.17	0.13
No partner	0.51	0.35	0.62	0.34	0.79	0.94	0.4	0.37	0.91	0.64	0.38	0.43	0.42	0.6	0.6
Partner (native German)	0.03	0.05	0.05	0.28	0.03	0	0.05	0.11	0	0.04	0.18	0.03	0.02	0.03	0.06
Partner (other native)	0.46	0.59	0.33	0.38	0.18	0.06	0.55	0.52	0.09	0.33	0.44	0.55	0.56	0.37	0.34
Occ.: in education	0.39	0.29	0.42	0.06	0.62	0.75	0.33	0.15	0.66	0.31	0.23	0.29	0.26	0.54	0.43
Occ.: in lab. market	0.45	0.62	0.43	0.77	0.3	0.19	0.53	0.62	0.2	0.59	0.65	0.64	0.65	0.35	0.46
Other occ./no answer	0.16	0.09	0.14	0.17	0.08	0.06	0.13	0.24	0.15	0.1	0.12	0.07	0.09	0.11	0.12
University degree	0.64	0.68	0.7	0.96	0.34	0.6	0.71	0.87	0.45	0.71	0.71	0.73	0.75	0.56	0.64
Risk attitude	6.95	5.65	6.36	5.53	6.99	6.61	6.93	5.26	5.62	6.94	6.4	5.9	6.25	6.73	6.36
Patience	6.8	5.59	5.87	5.33	6.1	6.58	6.14	6.22	6.22	6.26	5.77	5.83	6.34	6.58	6.13
Intern. applic. of educ.	3.93	3.88	4.18	4	3.83	3.77	3.92	2.83	3.52	4.32	3.99	3.99	4.01	3.83	3.87
Migration intention	0.88	0.63	0.56	0.71	0.73	0.68	0.67	0.38	0.65	0.8	0.64	0.6	0.63	0.81	0.68
No. observ.	132	329	476	347	559	444	218	216	292	340	77	143	232	551	4356

Table 5.A2: 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: *under 18, 18 to 24, 25 to 34, 35 to 49, 50 to 64, 65 and older.*
Female	Binary	Indicates if respondent is female.
No partner	Binary	Indicates if respondent has no partner.
Partner: native German	Binary	Indicates if respondent has a partner with German as native language.
Partner: other native	Binary	Indicates if respondent has a partner with other native languages than German.
Children	Binary	Indicates whether respondent has any children.
University degree	Binary	Indicates whether respondent has an university degree based on highest educational qualification: *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.*
Occ.: in education	Binary	Indicates whether respondent's main occupation is in education based on the category on main occupation: *pupil, student, student apprentice/(unpaid) trainee or apprentice.*
Occ.: in labour market	Binary	Indicates whether respondent's main occupation is in labour market based on the category on main occupation: *employee/civil servant with non-highly skilled job, employee/civil servant with highly skilled job, self-employed graduate (lawyer, doctor, ...)/freelance, other self-employed.*
Occ.: other/no answer	Binary	Indicates whether respondent's main occupation is not answered or other based on the category on main occupation: *unemployed, housewife/househusband, retiree or other.*
Risk attitude	Numerical (0-10)	Measures respondents' willingness to take risks (*"Would you describe yourself as someone who tries to avoid risks (risk-averse) or as someone who is willing to take risks (risk-prone)?"*) on a 11-point scale from 0 for *"risk-averse"* to 10 for *"risk-prone"*.
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 *"very impatient"* to 10 for *"very patient"*.
International applicability of education	Numerical (1-5)	Measures respondents' self-evaluated international applicability of education (*"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? "*) on a 5-point scale from 1 for *"none"* to 5 for *"all"*.
Migration intention	Binary	Indicates whether respondent has no migration intention (*"I would not move to another country under any circumstances"* or *"In principle, I would move to another country, but I have not thought about it in the last 12 months"*), or migration intention (*"I have been thinking about moving to another country in the last 12 months, but have n specific plans.", "I am planning a move to another country."* or *"I already have a date for my planned move to another country."*).

Table 5.A3: Question on (main) reasons for learning German.

Why are you learning German? Multiple choices possible.

☐ Study/education/training/PhD
☐ Internal company communication
☐ German-speaking trading partners
☐ German-speaking customers
☐ Higher income in the [country]
☐ Requirement/support of the employer
☐ Other considerations regarding career/labour market in [country]
☐ Requirement for visa
☐ Partner or family
☐ Social environment/friends
☐ (Possible) move to a German-speaking country for professional reasons
☐ (Possible) move to a German-speaking country for other reasons
☐ Holiday
☐ Culture (film, literature,...)
☐ Interest in languages
☐ Other: [free-text]

Look at your answers and circle the main reason why you are learning German.

Table 5.A4: Overview of responses and reasons

Reason	Responses
Education	Study/education/training/PhD
Company communication	Internal company communication
Trading partners	German-speaking trading partners
Customers	German-speaking customers
Income	Higher income in the [country]
Employer	Requirement/support of the employer
Labour market	Other considerations regarding career/labour market in [country]
Family	Partner or family
Friends	Social environment/friends
Educational and labour migr.	(Possible) move to a German-speaking country for professional reasons
Other migration	(Possible) move to a German-speaking country for other reasons
Holiday	Holiday
Culture	Culture (film, literature,...)
Interest	Interest in languages

Table 5.A5: Correlation coeffiecents

	Interest	Culture	Holiday	Friends	Family	Other migrat.	Educ., labour migration	Educat.	Labour market	Income	Employer	Customer	Trade partner
Interest													
Culture	0.32***												
Holiday	0.16***	0.31***											
Friends	0.13***	0.17***	0.23***										
Family	-0.07***	-0.05***	0.02	0.10***									
Other mig.	0.06***	0.06***	0.10***	0.12***	0.21***								
Educ., lab. mig.	0.03*	0.03**	-0.02	0.03*	-0.12***	0.00							
Education	-0.07***	-0.04**	-0.09***	-0.04***	-0.19***	-0.06***	0.18***						
Lab. market	0.05***	0.02	0.04***	0.02	-0.07***	0.00	0.04***	-0.03**					
Income	0.02	0.01	0.03**	0.03*	-0.08***	0.00	0.10***	0.06***	0.16***				
Employer	0.00	0.01	0.01	0.03***	-0.02	-0.03**	-0.01	-0.05***	0.05***	0.09***			
Customer	-0.01	0.00	0.02	0.02	-0.03**	-0.04**	0.02	-0.09***	0.06***	0.18***	0.19***		
Trade partner	-0.05***	-0.01	0.03*	0.01	-0.04**	-0.01	0.01	-0.05***	0.04***	0.12***	0.16***	0.39***	
Company communic.	-0.03*	-0.01	-0.01	0.03*	-0.03*	-0.03*	0.03*	-0.05***	0.05***	0.11***	0.22***	0.32***	0.26***

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

Table 5.A6: Estimation results: country characteristics

	(1) Germanic	(2) Non-Germ.	(3) Lower-middle income	(4) Upper-middle income	(5) High income	(6) European (EU)	(7) European (non-EU)	(8) Non-European
Age: under 35 years	0.075	0.132***	0.176***	0.150**	0.166***	0.112***	0.119*	0.102**
	(0.047)	(0.027)	(0.056)	(0.059)	(0.029)	(0.031)	(0.066)	(0.044)
Gender: female	−0.052*	−0.094***	−0.125***	−0.063*	−0.108***	−0.067***	−0.074**	−0.119***
	(0.027)	(0.016)	(0.021)	(0.035)	(0.020)	(0.022)	(0.034)	(0.020)
Children	−0.061	−0.063**	0.011	−0.034	−0.012	−0.086**	−0.084	0.011
	(0.058)	(0.030)	(0.047)	(0.065)	(0.034)	(0.037)	(0.061)	(0.045)
Partner (nat. German)	−0.387***	−0.336***	−0.504***	−0.434***	−0.489***	−0.296***	−0.302***	−0.446***
	(0.056)	(0.046)	(0.081)	(0.092)	(0.035)	(0.036)	(0.104)	(0.066)
Partner (other native)	0.00003	−0.024	−0.032	−0.110***	−0.061***	0.003	0.001	−0.049
	(0.040)	(0.019)	(0.031)	(0.041)	(0.023)	(0.024)	(0.039)	(0.030)
Occ.: in education	0.127***	0.116***	0.123***	0.118**	0.116***	0.115***	0.078	0.214***
	(0.046)	(0.023)	(0.036)	(0.050)	(0.031)	(0.034)	(0.050)	(0.033)
Other occ./no answer	−0.016	−0.004	0.074**	0.004	0.045	−0.053	−0.055	0.061*
	(0.043)	(0.025)	(0.036)	(0.054)	(0.029)	(0.034)	(0.053)	(0.032)
University degree	0.005	0.026	0.034	0.108**	0.058**	0.023	0.0004	0.011
	(0.037)	(0.020)	(0.029)	(0.049)	(0.027)	(0.032)	(0.041)	(0.027)
Risk attitude	0.003	0.002	0.007	−0.003	0.004	0.002	−0.001	−0.008
	(0.007)	(0.004)	(0.006)	(0.008)	(0.005)	(0.005)	(0.009)	(0.005)
Patience	0.006	0.007**	0.006	0.011	0.007*	0.003	0.003	0.006
	(0.006)	(0.003)	(0.005)	(0.007)	(0.004)	(0.004)	(0.007)	(0.005)
Intern. applic. of educ.	0.021	0.042***	0.042***	0.054***	0.046***	0.034***	0.030*	0.056***
	(0.013)	(0.008)	(0.011)	(0.018)	(0.009)	(0.011)	(0.017)	(0.010)
Migration intention	0.074**	0.067***	0.043*	0.154***	0.074***	0.045*	0.068	0.118***
	(0.029)	(0.018)	(0.026)	(0.043)	(0.021)	(0.023)	(0.047)	(0.022)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	868	3,373	704	1,554	2,098	1,822	683	1,851

Average marginal effects. Heteroscedasticity robust White standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

160

Table 5.A7: Estimation results: subsamples age and gender

	Dependent variable: Professional use of German							
	(1) female		(2) male		(3) age: ¡35 years		(4) age: ≥ 35	
Investment	0.177***	(0.020)	0.221***	(0.026)	0.143***	(0.018)	0.361***	(0.033)
Age: under 35 years	0.066**	(0.029)	0.043	(0.035)				
Gender: female					0.033**	(0.016)	0.017	(0.029)
Children	−0.023	(0.034)	−0.036	(0.041)	−0.001	(0.048)	0.006	(0.032)
Partner (native German)	0.069*	(0.036)	0.101*	(0.052)	0.097***	(0.035)	0.037	(0.051)
Partner (other native)	0.012	(0.022)	0.067**	(0.028)	0.070***	(0.020)	−0.061*	(0.034)
Occ.: in education	0.085***	(0.026)	0.100***	(0.033)	0.116***	(0.022)	−0.105	(0.164)
Other occ./no answer	0.031	(0.027)	0.024	(0.037)	0.059**	(0.027)	0.001	(0.035)
University degree	−0.020	(0.023)	−0.016	(0.030)	0.0003	(0.020)	−0.048	(0.047)
Risk attitude	0.016***	(0.004)	0.019***	(0.005)	0.019***	(0.004)	0.007	(0.007)
Patience	0.002	(0.004)	0.008*	(0.005)	0.003	(0.003)	0.011*	(0.006)
Intern. applic. of educ.	0.047***	(0.009)	0.034***	(0.010)	0.042***	(0.008)	0.040***	(0.012)
Migration intention	0.054***	(0.020)	0.101***	(0.025)	0.061***	(0.018)	0.113***	(0.031)
Country FE	Yes		Yes		Yes		Yes	
McFadden Pseudo R2	0.16		0.15		0.12		0.23	
% correctly predicted	71.2		70.7		70.1		74.5	
Observations	2,572		1,784		3,395		961	

Average marginal effects. Heteroscedasticity robust White standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 5.A8: Estimation results: country characteristics

	(1) Germanic	(2) Non-Germanic	(3) Lower-middle income	(4) Upper-middle income	(5) High income	(6) European (EU)	(7) European (non-EU)	(8) Non-European
				Dependent variable: Professional use of German				
Investment	0.207***	0.204***	0.139***	0.182***	0.236***	0.211***	0.130***	0.194***
	(0.044)	(0.017)	(0.028)	(0.038)	(0.023)	(0.024)	(0.035)	(0.027)
Age: under 35 years	−0.056	0.072***	0.115**	0.063	0.029	0.029	0.036	0.125***
	(0.051)	(0.026)	(0.056)	(0.049)	(0.029)	(0.030)	(0.053)	(0.047)
Gender: female	−0.002	0.042***	0.063***	0.034	0.012	0.028	0.101***	0.015
	(0.034)	(0.016)	(0.024)	(0.032)	(0.021)	(0.022)	(0.031)	(0.023)
Children	−0.025	−0.022	0.008	0.012	−0.050	−0.058	0.027	−0.009
	(0.067)	(0.029)	(0.054)	(0.051)	(0.037)	(0.036)	(0.048)	(0.056)
Partner (native German)	0.030	0.124***	0.252***	0.047	0.064	0.053	0.151***	0.194***
	(0.062)	(0.038)	(0.043)	(0.101)	(0.039)	(0.040)	(0.050)	(0.060)
Partner (other native)	0.0002	0.024	0.096***	0.015	0.017	0.024	0.017	0.069**
	(0.053)	(0.019)	(0.032)	(0.037)	(0.025)	(0.025)	(0.036)	(0.032)
Occ.: in education	0.067	0.104***	0.162***	0.117***	0.057*	0.057*	0.089**	0.160***
	(0.051)	(0.023)	(0.036)	(0.044)	(0.031)	(0.032)	(0.044)	(0.035)
Other occ./no answer	−0.048	0.046*	0.078*	0.054	−0.003	−0.019	0.026	0.091**
	(0.058)	(0.024)	(0.045)	(0.045)	(0.030)	(0.032)	(0.050)	(0.038)
University degree	0.031	0.003	0.010	−0.034	−0.016	−0.043	−0.009	0.033
	(0.043)	(0.020)	(0.031)	(0.045)	(0.028)	(0.030)	(0.033)	(0.031)
Risk attitude	−0.004	0.017***	0.016**	0.020***	0.017***	0.016***	0.026***	0.014**
	(0.008)	(0.004)	(0.007)	(0.007)	(0.005)	(0.005)	(0.008)	(0.006)
Patience	−0.004	0.005	−0.003	0.003	0.009**	0.005	0.004	0.002
	(0.007)	(0.003)	(0.005)	(0.007)	(0.004)	(0.005)	(0.006)	(0.005)
Intern. applic. of educ.	0.008	0.041***	0.041***	0.024	0.049***	0.040***	0.051***	0.037***
	(0.015)	(0.007)	(0.013)	(0.015)	(0.010)	(0.010)	(0.013)	(0.012)
Migration intention	−0.033	0.068***	0.080***	0.081**	0.060***	0.061***	0.184***	0.048*
	(0.038)	(0.018)	(0.028)	(0.040)	(0.022)	(0.024)	(0.045)	(0.025)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	813	3,202	670	1,479	2,001	1,774	658	1,718

Average marginal effects. Heteroscedasticity robust White standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Figure 5.A1: Professional use of German: 1=very unlikely, 5=very likely.
(n=4356)
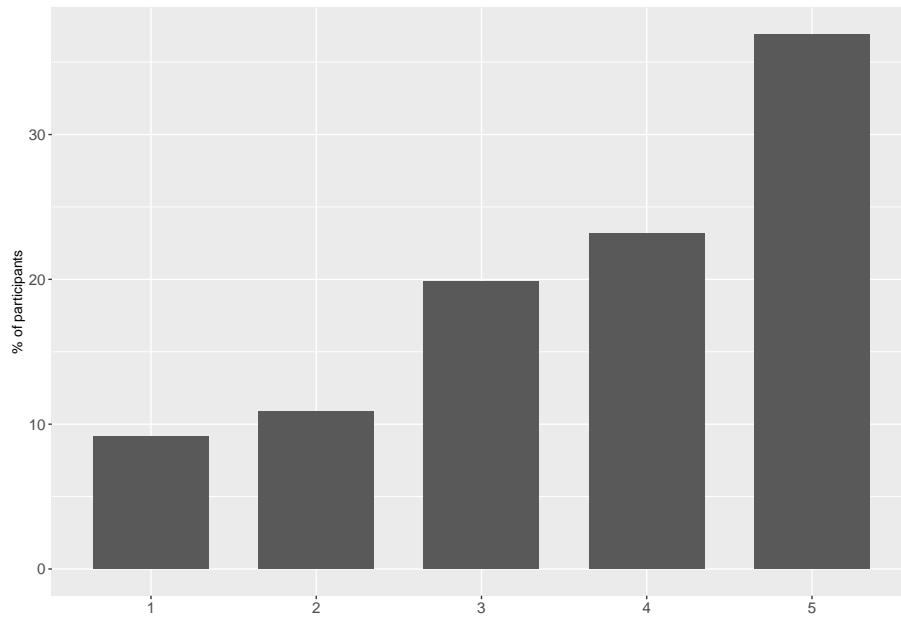
Figure 5.A2: Use of German on the labour market by countries and economic motive.
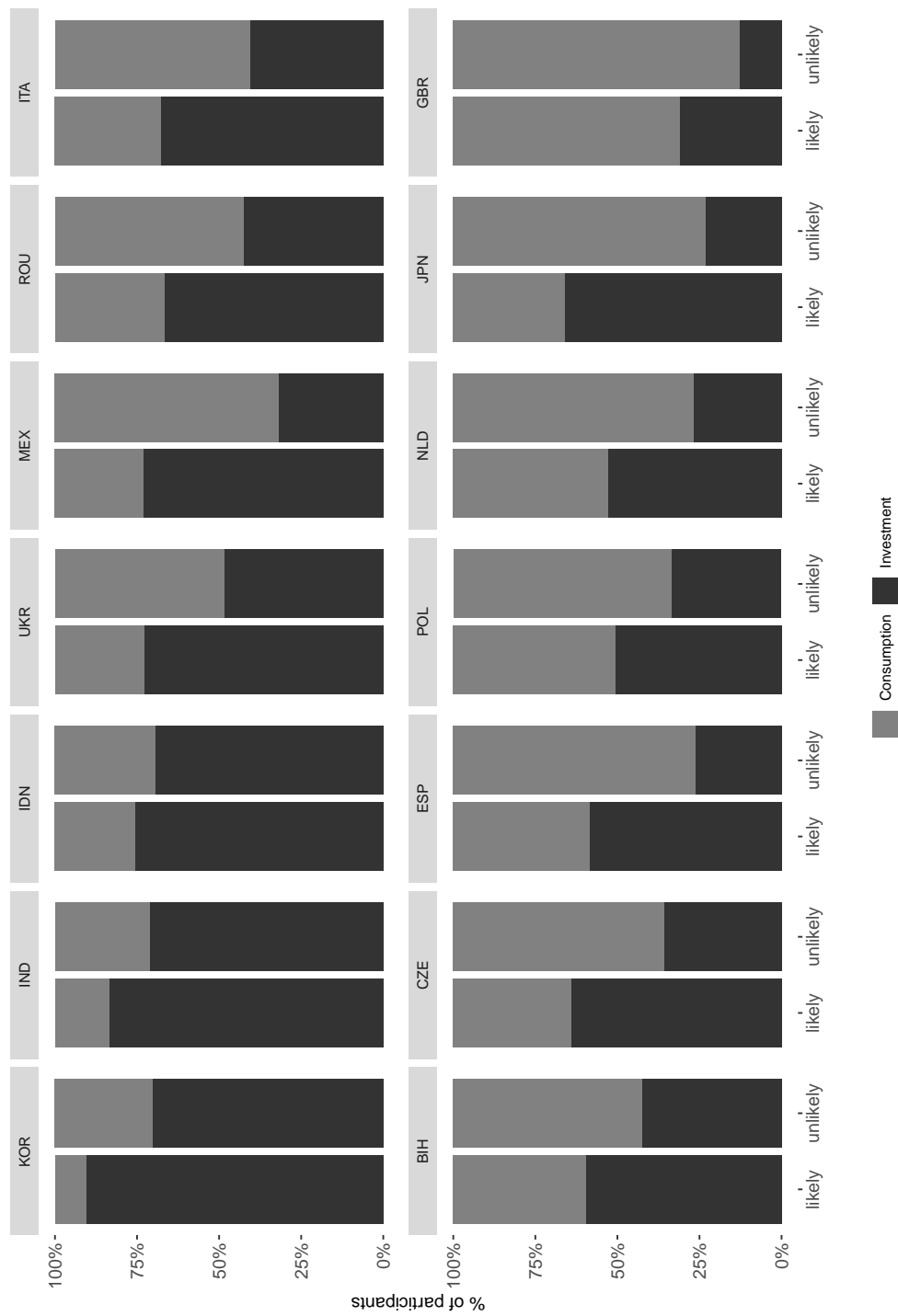
Figure 5.A3: Average marginal effects and 95 % CI for estimations with country group subsamples according to country characteristics.

Figure 5.A4: Average marginal effects and 95 % CI for estimations with country characteristics subsamples.
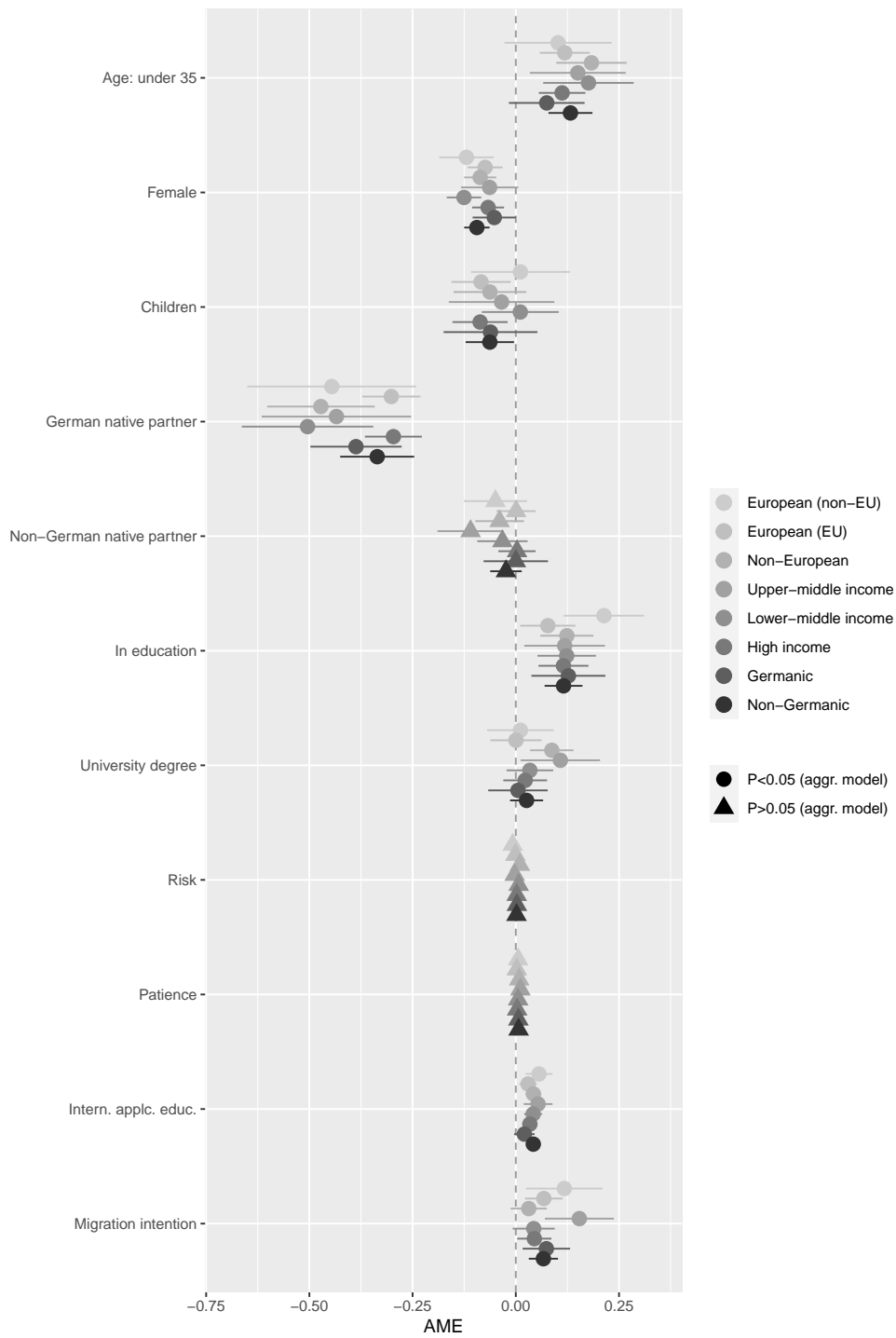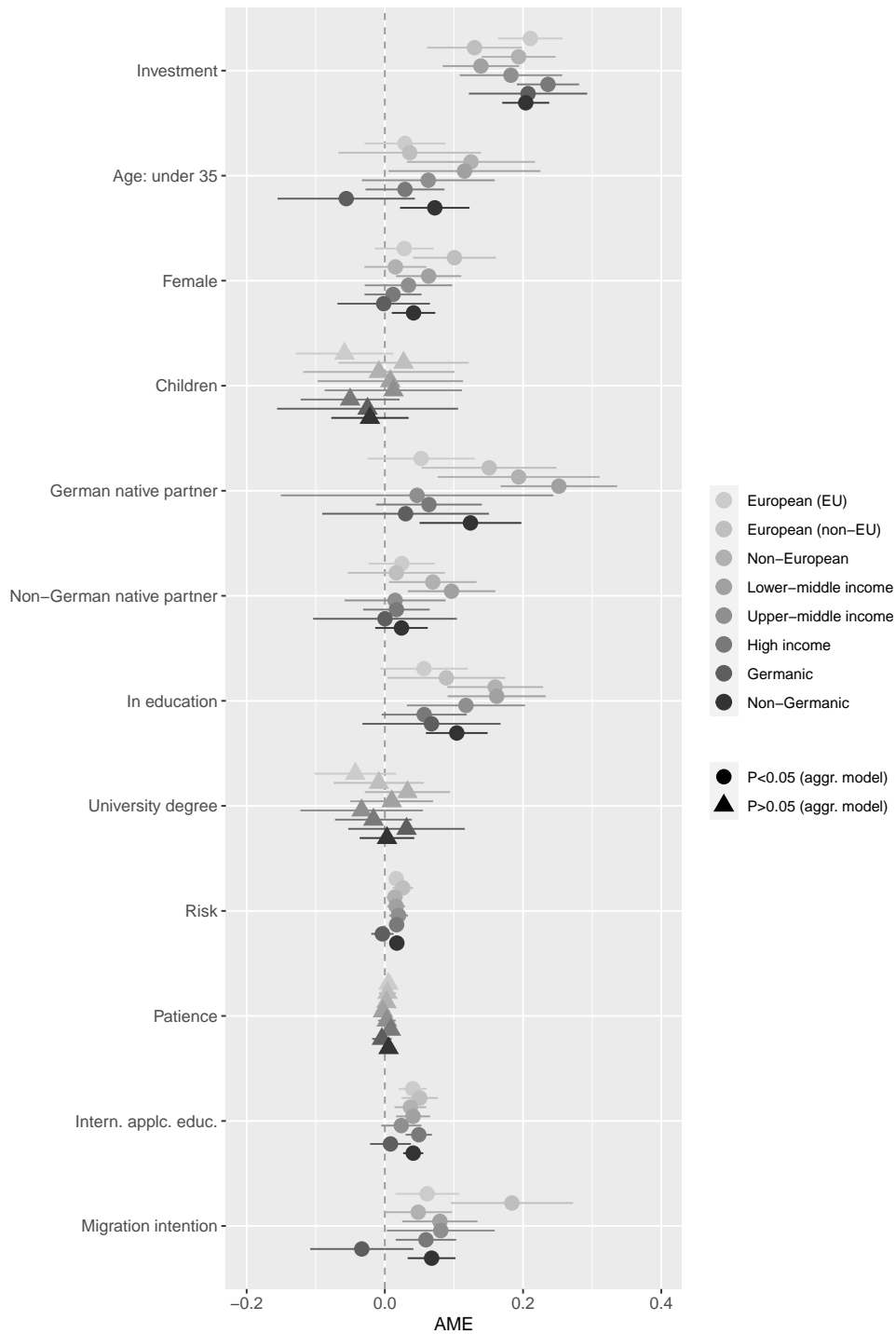
Notes: Dependent variable is professional use of German. Shapes according to the p-values of the aggregated model refers to Table 5.5 Column 3. For detailed estimation results see Table 5.A8 in the Appendix.

Part III

Migration aspirations and intentions

Chapter 6

Migration aspirations and intentions

Abstract.[59]   We carried out two multinational surveys to analyze aspirations and intentions to emigrate, and how these are linked to each other. One survey covered language course participants in 14 countries, and another students in 6 countries. We identify two groups that have been neglected in previous research on migration aspirations and intentions: those who intend to migrate permanently without aspirations to do so and those who intend to migrate temporarily, but report no aspiration to migrate permanently. Analyzing main motivations to emigrate shows that discrepancy among women is driven mainly by family, and among men by work and studies.

6.1 Introduction

The reasons why people leave their country of origin have been studied for many decades. A large body of migration literature evaluates immigrants in destination countries and offers important insights on migration choices (Abramitzky et al., 2012; Adserà and Pytliková, 2015; Beine et al., 2011; Borjas, 1987; McKenzie and Rapoport, 2010). There is, however, an important drawback: only those individuals who actually migrated are observed. To overcome this shortcoming, a growing strand of migration literature relies on survey data to study individuals at their origin before actual migration takes place in order to identify potential future migrants. Some studies focus on migration aspirations, which express a desire to migrate under ideal circumstances (e.g. Beine et al., 2020; Bertoli and Ruyssen, 2018; Docquier et al., 2020; Ruyssen and Salomone, 2018). Other studies use questions about migration intentions, which refer to considerations and plans within the actual situation subject to migration constraints (e.g. Chort, 2014; De Jong, 2000; Friebel et al., 2013; Ivlevs and King, 2012; Papapanagos and Sanfey, 2001; Uebelmesser, 2006; van Dalen and Henkens, 2012). There are also studies which combine both measures, e.g. Bertoli et al. (2020) and Manchin and Orazbayev (2018), where the latter does not specifically focus on international migration.[60]

Often studies make use of the Gallup World Poll (henceforth GWP) data, which is a repeated cross-sectional survey covering around 160 countries over 14 years and which allows for an analysis of representative populations. The GWP data comes, however, with two important limitations: First, the question on migration aspirations under ideal circumstances serves as a filter question for further questions regarding considerations and plans towards migration, such that the latter questions are only asked if the respondents indicate aspirations to migrate. Second, the questions on aspirations and intentions are limited to permanent migration in most waves of the GWP, and hence most works based on the GWP are limited to a permanent time-horizon.[61]

The first limitation creates a conditionality in responses, which frames aspirations to migrate as a necessary condition for migration considerations and plans. However, the desire to migrate on the one hand, and current consider-

---

[60] While this terminology is widely used, some studies refer to similar concepts, but use different terminology. We will subsume them under those two terms. For an extensive list of terms and questions employed in potential migration research, see Aslany et al. (2021) and Williams et al. (2018).

[61] For more discussion on this, see Section 6.2.3.

ations and plans of concrete steps towards migration on the other hand, are not necessarily formed in a consecutive order and intentions to migrate do not necessarily imply desire to do so. Cases where individuals consider or plan to migrate without having an underlying desire to do so are therefore not captured by the GWP. To the extent that those who intend to migrate with their partner, as so-called tied movers, are predominantly women and reply that they would ideally not like to emigrate, the GWP risks underestimating women's migration plans relative to men's.

The second limitation neglects potential migrants who only aspire or intend to migrate for a pre-defined time-span. Data on temporary migration is generally rare, despite evidence that repeat and return migrants make up a considerable share of overall migrants (Dustmann and Görlach, 2016). While there are studies on return migration, e.g. Dustmann and Kirchkamp (2002) which analyses the activity choices of returned migrants, and Dustmann (2003) which analyses how a change in wage differentials between origin and host country affects the optimal duration of stay, those studies use data of individuals after their realized emigration or return to their origin country.[62] As a consequence, migration analyses of potential future migrants who intend to stay only temporarily in their destination country are missing, to the best of our knowledge.

The aim of this study is to address the two above-mentioned limitations and to provide a comprehensive analysis of migration aspirations and intentions, including those potential future migrants who have often been overlooked. For this purpose, we conducted two multinational surveys: one among language course participants in 14 countries and one among university students in six countries. The first captures a self-selected subset of the population, many of whom can be expected to be relatively far in the migration process, while the second captures students in higher education. These novel datasets observe both migration aspirations (via the same question as the GWP) and migration intentions for all respondents, i.e. independent of their stated aspirations and separately for intentions to migrate temporarily and permanently. We find aspirations and current intentions to migrate to match perfectly for only 60.7 % of the language course sample, and only 48.7 % of the student sample. As far as the first limitation – the conditionality in responses – is concerned, we find that 2.9 % of the language course sample and 2.3 % of the student sample intends to migrate permanently despite a lack of aspirations to do so. Regarding the second limitation – the negligence of temporary potential migrants – we find

---

[62] For an extensive overview of this literature, see Dustmann and Görlach (2016).

that with 24.6 % of the language course sample, and 34.1 % of the student sample, a substantial share of respondents intends to migrate only temporarily.

Our framework allows us to identify several groups of potential future migrants and to contribute to the literature on migration aspirations and intentions in three ways. First, by observing both aspirations and intentions for all respondents, we can compare these two measures within individuals. Though we find that the former is an important determinant of the latter, their relationship is far from perfect, and even after controlling for aspirations, individual-level characteristics remain important in explaining intentions. Second, we are able to differentiate between intentions to migrate temporarily and permanently. This distinction allows us to shed light on the large group of those with intentions to migrate temporarily and to analyse possible differences in the determinants of intentions to migrate temporarily and permanently. Third, avoiding the conditionality in responses problem, we further observe those individuals who do not desire to migrate permanently, i.e. do not state aspirations to migrate permanently under ideal circumstances yet consider or plan to do so. Overall, with this more complete picture of potential future migrants, we can identify instances where aspirations and intentions differ and hence better understand the individual-level factors which potentially evoke intentions to move abroad despite the desire to stay under ideal circumstances or impede aspirations to migrate from being realized. With most studies relying on the GWP data, these groups have not received much attention so far in research of potential future migrants due to the conditionality problem and the focus on permanent migration.

We estimate individual-level determinants of aspirations and intentions to migrate, temporarily or permanently, via multinomial probit regression. For our analysis, we run our estimations for both datasets separately and find highly comparable patterns. We find that aspirations are an important determinant of intentions to migrate, yet the relationship is far from perfect, and after controlling for aspirations, individual-level characteristics remain important in explaining intentions to migrate both temporarily and permanently. Differences between aspirations and intentions to migrate permanently are primarily explained by family ties. Strong family ties at the origin, i.e. having children, can impede existing aspirations to migrate permanently from being translated into intentions. Family ties abroad, i.e. having a non-native partner, on the contrary, can lead to intentions to migrate permanently despite the absence of aspirations to do so. Intentions to migrate temporarily are primarily explained by educational plans. However, the relevance of determinants differs

171

by gender. The positive linkage between family ties abroad and intentions to migrate permanently is driven largely by women in both samples. The positive linkage between a university degree and intentions to migrate temporarily is robust only for men across both samples. Those findings are underlined by a descriptive analysis of motivations for potential migration in our samples. Among those with migration intentions, the share of respondents who state educational or professional reasons as their main motivation is consistently larger among men, and the share of those who state family and partner as their main motivation is consistently larger among women in both samples.

The remainder of the chapter is structured as follows. In Section 6.2, we review the literature on potential migration and introduce our migration choice model. Section 6.3 describes our datasets and provides some descriptive statistics. Section 6.4 shows the estimation strategy. Section 6.5 presents the results and Section 6.6 concludes.

6.2 Measuring potential migration

The literature on potential migration relies heavily on survey data and should in the best-case scenario convey migration aspirations and intentions reliably. Potential future migrants are identified based on differently framed survey questions.

6.2.1 Migration aspirations

Some recent studies utilize GWP data to determine potential migration (see Beine et al., 2020; Bertoli and Ruyssen, 2018; Bertoli et al., 2020; Docquier et al., 2020; Manchin and Orazbayev, 2018; Ruyssen and Salomone, 2018), making use of a measure we will subsequently call migration aspirations. The GWP asks respondents *"Ideally, if you had the opportunity, would you like to move permanently to another country, or would you prefer to continue living in [country in which the survey takes place]?"* and gives the response options *"Like to stay in [country in which the survey takes place]"* or *"Like to move to another country"*. This describes aspirations to migrate in a hypothetical, ideal-world scenario under the absence of any barriers and gives no statement about concrete considerations or plans.

6.2.2 Migration intentions

Some studies utilize migration intentions, which express considerations or plans to migrate in a real-world scenario, subject to constraints. Though there has been some debate on whether intentions can predict future behavior (Bertrand and Mullainathan, 2001; Manski, 1990), in the economic literature, intentions are frequently used to predict behavior (e.g. Falck et al., 2017; Juster, 1960; Okun, 1962; Tobin, 1959) and migration research is no exemption here. Under the 'best-case' hypothesis, i.e. when respondents state their intentions based on rational expectations, intentions are indeed the best predictor of future behavior (Manski, 1990). Burda et al. (1998) assume that intentions are a "monotonic function of the underlying driving variables which motivate migration", and studies frequently find migration intentions to be a strong predictor of subsequent behavior (e.g. Chort, 2014; Tjaden et al., 2019).

Several authors in economics (e.g. Friebel et al., 2013; Ivlevs and King, 2012; Papapanagos and Sanfey, 2001; Uebelmesser, 2006; van Dalen and Henkens, 2012) and neighboring social sciences (e.g. De Jong, 2000) determine potential migration, based on different questions which aim at capturing respondents' migration intentions. Often, the questions directly include the word 'intend' (Friebel et al., 2013) or 'consider' (Uebelmesser, 2006). Others refer to the likelihood or probability of migration (Ivlevs and King, 2012; Papapanagos and Sanfey, 2001). While most of the survey questions refer very generally to international migration, some use more specific questions on destinations (Friebel et al., 2013; Papapanagos and Sanfey, 2001) or explicitly refer to internal migration (Chort, 2014).

Similarly, questions asked about migration intentions also differ with respect to the time frame of their potential date of realization, ranging from no indicated time period (Papapanagos and Sanfey, 2001; Uebelmesser, 2006) to specific time periods of between six months (Friebel et al., 2013) up to ten years (van Dalen and Henkens, 2012). Furthermore, survey questions differ in whether they refer to general migration, like most of the previously mentioned studies do, or only to labour migration (Papapanagos and Sanfey, 2001; van Dalen and Henkens, 2012). In addition to heterogeneity in the survey questions used, the mentioned studies differ in their focus. While most of them are interested in emigration from a single country (Uebelmesser, 2006; van Dalen and Henkens, 2012), some additionally focus on specific destination countries (Papapanagos and Sanfey, 2001) and others on a specific determinant, e.g. xenophobic attacks (Friebel et al., 2013) or whether children of former migrants are more likely to

migrate themselves (Ivlevs and King, 2012).

6.2.3   Combining migration aspirations and intentions

In our survey, we ask about respondents' aspirations to migrate permanently in the same way as in the GWP questionnaire (see Section 6.2.1). In direct succession, and independent of their answer to the former, all respondents are asked about their migration intentions: *"Tick the statement that applies to your current situation"*. Out of five options, response options *"I would not move to another country under any circumstances"* and *"In principle, I would move to another country, but I have not thought about it in the last 12 months"* are classified as the respondent having no intentions to migrate, and the remaining response options *"I have been thinking about moving to another country in the last 12 months, but have no specific plans"*, *"I am planning a move to another country"*, and *"I already have a date for my planned move to another country"* are classified as the respondent having migration intentions.[63] Those with migration intentions are further divided by whether their intentions relate to temporary or permanent migration.[64] Combining those questions, we categorize potential migrants in six mutually exclusive combinations of aspirations and intentions, as shown in Figure 6.1.

According to a Roy-Borjas model of utility maximization in migration (Borjas, 1987), individuals who are able to migrate intend to migrate if their expected utility from relocating abroad, net of migration costs, exceeds their expected utility from staying. Then, both migration aspirations and migration intentions emerge as the result of a cost-benefit analysis of potential migration (Borjas, 1987). However, only migration intentions are subject to constraints and drivers according to the real-world situation of the individual, while migration aspirations are not. Such constraints, such as visa requirements, liquidity con-

---

[63]   This builds on measures of migration choice in multiple phases (e.g. De Jong, 2000), which in turn draw on 's (1991) *Theory of Planned Behavior* which differentiates between intentions, which do not include any concrete actions, and behavior including concrete actions.

[64]   Those respondents with migration intentions who state that they would most likely stay in their preferred destination country for more than 5 years or state that their return to [country in which the survey took place] after a temporary stay in their preferred destination country is unlikely are classified as having intentions to migrate permanently; the rest is classified as having intentions to migrate temporarily. Those who state no migration intentions are not asked those questions.

Figure 6.1: Combinations of aspirations and intentions.
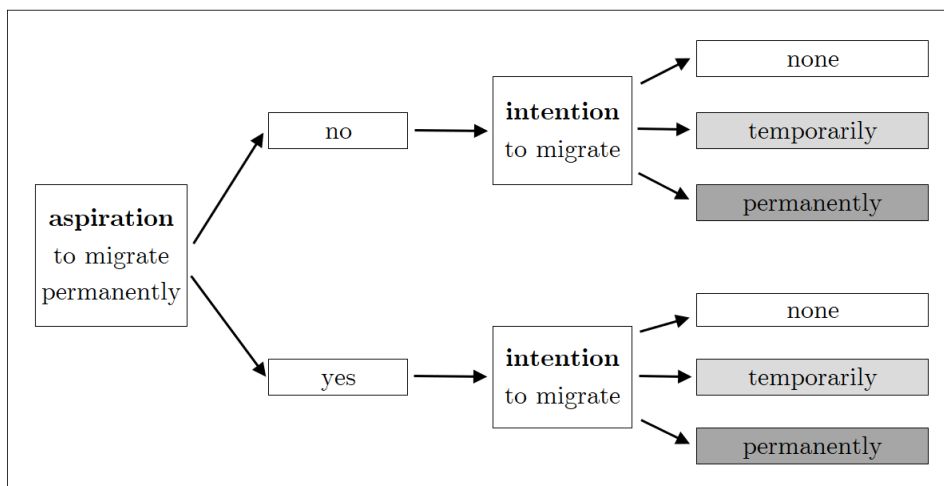
straints, and social ties at home, may prevent migration intentions even when one would ideally like to migrate. At the same time, migration intentions can also be evoked by migration drivers even when one would ideally not like to migrate, like a partner wanting to emigrate and career reasons. Different from Delogu et al. (2018) and Docquier et al. (2015), which assume that all migration aspirations would realize if visa restrictions were abolished, we consider factors explaining differences between aspirations and intentions in a broader sense. Whether migration aspirations result in actual consideration and plans then depends not only on the costs, but also on the constraints the individuals face when converting this desire into reality.[65]

Consequently, migration aspirations and intentions are not necessarily equivalent. Even so, some studies use the question on migration aspirations as a measure for migration intentions (Beine et al., 2020; Bertoli and Ruyssen, 2018). Others also include alternative questions on planning and preparation activities for robustness checks (Docquier et al., 2020). Ruyssen and Salomone (2018) use GWP data to combine the questions on migration aspirations and migration preparation in a Heckman selection model for those with aspira-

---

[65] This is closely related to the migration aspiration/ability model, which distinguishes between the desire to migrate and the ability to realize this aspiration (Carling and Schewel, 2018). We will not use that model, since we analyse individuals prior to actual migration.

tions. They find that perceived gender discrimination in the home country has a positive effect on migration aspirations but not on migration preparation.

Overall, migration aspirations and intentions can differ within individuals. These differences can have several reasons. First, the question on migration aspirations aims at a hypothetical scenario (*"Ideally, if you had the opportunity, would you like [....]"*), while migration intentions refer to the current situation under migration constraints and drivers.

Second, questions on aspirations and intentions may refer to different time horizons or, especially in the case of intentions, may not make clear the time horizon. As mentioned, the GWP question on migration intentions refers to current intentions, i.e. it asks about considerations within the last 12 months or fixed plans which are likely to be executed within the near future.[66] Aspirations, on the other hand, are not restricted to any timescale. Ideally wanting to migrate permanently yet not having current intentions to do so does not need to be a conflict, but rather is potentially just a question of looking for the optimal timing. Some individuals, for example, who would first like to finish their education or wait for their kids to leave home before migrating permanently, would state having migration aspirations despite not actually intending to migrate for years to come. Also, the question on migration aspirations is restricted to permanent migration.[67] Individuals who only aspire to migrate temporarily are inclined to answer negatively. At the same time, current intentions to migrate temporarily can still be consistent with long-term aspirations to migrate permanently. Individuals, for example, could want to migrate temporarily to one country to obtain an educational degree (for example to a country where tuition fees are comparably low) and only then migrate to a third country af-

---

[66] Henceforth, whenever we refer to 'intentions', we refer to this measure of current intentions to migrate.

[67] The GWP waves of 2009 to 2012 include an additional question which asks the same as above, only referring to 'temporarily' instead of 'permanently'. Still, most works utilize the GWP question referring to 'permanently' (Beine et al., 2020; Bertoli et al., 2020; Docquier et al., 2020; Manchin and Orazbayev, 2018) or combine the two questions without specifically differentiating between them (Bertoli and Ruyssen, 2018; Ruyssen and Salomone, 2018). The follow-up questions on considerations and preparations in the GWP do not specifically refer to a time horizon (*"To which country would you like to move?"* and *"Have you done any preparations for this move (For example applied for residency or visa, purchased a ticket, etc)?"*) or refer to the same time horizon as the initial question (*"Are you planning to move permanently to another country in the next 12 months or not?"*). Our measure of migration aspirations includes the word 'permanently'. Henceforth, whenever we refer to 'aspirations', we implicitly mean aspirations to migrate permanently.

176

ter graduation. Utilizing the 2009-2012 waves of GWP data, which include the question on migration aspirations regarding temporary international migration, Delogu et al. (2018) show that those intending to migrate temporarily are a considerable share of potential migrants.

Third, choices regarding migration aspirations and intentions are not necessarily made in a consecutive order, and one may not be conditional on the other. Aspirations to migrate are not necessarily followed by intentions to migrate, and intentions to migrate are not necessarily based on aspirations. Individuals could not aspire to migrate but still consider or plan to do so, e.g. due to their partner being offered a job opportunity abroad. Similarly, individuals who answer affirmatively to the question regarding their aspirations might decide against migration due to the constraints they face, like legal barriers or social ties. In fact, many individuals, especially from developing countries, would like to migrate under ideal circumstances yet will never do so. Aggregated GWP data from the 2018 wave, for example, shows that 25.0 % of surveyed Mexicans would like to migrate under ideal circumstances, yet only 310,000 Mexicans emigrated in 2019, which is a mere 0.2 % of the 127.6 million residents Mexico had in 2019. Further, some individuals might form intentions and only afterwards decide whether they want to stay abroad temporarily or permanently. Alternatively, some could form intentions to migrate only if they have the option to migrate temporarily (e.g. students' choice to do a semester or course abroad might be based on the option to stay temporarily, and they would not have formed those intentions if returning after their studies was not an option). As different orders of choice are possible, we do not model those choices as being consecutive, but rather as being simultaneous – resulting in our 6 combinations of aspirations and intentions shown in Figure 6.1. This is in stark contrast to the GWP data, which assumes a conditionality in choice. In the GWP, only those respondents who state positive migration aspirations are asked subsequent questions regarding their migration plans within the next 12 months (*"Are you planning to move permanently to another country in the next 12 months, or not?"*) and whether they have already prepared for this move (*"Have you done any preparation for this move (for example applied for residency or visa, purchased the ticket, etc.)?"*).[68] While those questions indeed indicate how far the respondents are in their respective migration decision-

---

[68]   Independent of this set of questions, there is a further measure on migration intentions in the GWP, which also includes the possibility of internal migration: *"In the next 12 months, are you likely or unlikely to move away from the city or area where you live?"*. See e.g. Dustmann and Okatenko (2014).

making process, they exclude all respondents who did not state aspirations to migrate permanently.

Allowing for differences between aspirations and intentions and extending the time horizon to temporary and permanent intentions, our survey design gives access to three groups of potential migrants which have not received much attention so far. By avoiding the conditionality in responses problem, we observe those who do not report aspirations to migrate permanently but nonetheless consider or plan to migrate permanently. By introducing a time horizon to the question on intentions, we further observe both those who do not report aspirations to migrate permanently but consider or plan to migrate temporarily; and those who do report aspirations to migrate permanently but again only consider or plan to migrate temporarily. We want to highlight those groups here because they can be large and thus are an important target of migration research and policy. At the same time, with most studies relying on GWP data, these groups have been largely neglected so far in migration research.

6.2.4 Determinants of aspirations and intentions

Sticking to 's (1987) and subsequent literature's perspective of utility maximization in migration, we investigate determinants of potential migration, i.e. how they influence the cost-benefit analysis of migration on the individual level.

Based on the *Human Capital Theory of Migration*, the net gain of migration decreases with age as it lowers the time to recoup the investment made by migration (Sjaastad, 1962). In addition, older individuals face higher migration costs in general as they lose specific human capital as well as their social and professional networks while having larger difficulties in adapting to a new language and a new environment (Belot and Ederveen, 2012).

According to this human capital theory, the net gain of migration also depends on education. Borjas (1987) showed that if skills are sufficiently transferable across countries, migrants from a country with narrower income differences who migrate to a country with wider income differences tend to come from the upper part of the skill distribution, while migrants from a country with wider income differences who migrate to a country with narrower income differences

tend to come from the lower part of the skill distribution.[69] Empirically, highly educated individuals have generally been found to be more likely to migrate (see, e.g., Borjas et al., 2019; Docquier et al., 2014; Grogger and Hanson, 2011). Thus, we control for whether respondents have a university degree and, additionally, whether they are still students.

Previous migration experience can reduce migration costs and constraints. It can be seen as a proxy for inherent mobility or indicate a learning effect. Individuals who have lived abroad before are generally more aware of their alternatives abroad, which lowers transaction costs. While migration can be evaluated ex-post positively as well as negatively, migration-related soft-skills and experience do not depend on such subjective evaluation. Either way, repeat migration constitutes a considerably large share of overall migration flows (e.g. DaVanzo, 1983). Having previously migrated has been shown to be a strong predictor of subsequent migration (Uebelmesser, 2006). Therefore, we test how having previously lived abroad is linked to migration aspirations and intentions. One concern here is that for those who have already emigrated from their country of birth (i.e. are foreign-born), the decision to migrate is substantially different from someone who contemplates emigrating from their country of birth. Leaving one's home country is a decision very different from leaving a host country, and for the foreign-born, emigration plans can constitute repeat migration as well as return migration to their home country. Since we cannot reasonably distinguish between both, we exclude foreign-born respondents (i.e. those who have been born in a country different from the country in which the survey took place) from the analysis.

When it comes to the role of family ties, it is clear that migration is not an individual decision but happens in the context of a social environment (Stark and Bloom, 1985). Family ties are bound to influence the migration decision, though the channels through which this happens can lead to different outcomes. Strong family ties at the source country – such as marriage, long-term relationship or children – could be a restricting factor regarding emigration. Family ties abroad, on the contrary, could work as an encouraging factor as they imply easier access to information about jobs and earning opportunities, as well as financial and emotional support. This all substantially lowers costs and risks related to migration (Manchin and Orazbayev, 2018). Also, a partner

---

[69] Poutvaara (2004, 2008) concluded that those with internationally applicable education, like STEM degrees or economics, can be expected to be mobile and more so when the international applicability of their education increases. Those with country-specific degrees, like law, instead would tend to stay in their country of education.

who wants to emigrate can evoke migration considerations and plans among individuals even if they have no family ties abroad, e.g. if joint emigration is necessary to maintain the relationship. Hence, generally, migration decisions are coordinated within households or families (Chort, 2014), yet that coordination seems to be more binding for women (Munk et al., 2017). Thus, we also control for gender in our estimations.

The uncertainty about potential returns and costs make migration a risky choice. Individuals who are more willing to take risks are expected to be more likely to consider migration and follow through with it. This has been shown in previous literature (Jaeger et al., 2010; Roca Paz and Uebelmesser, 2021). Lastly, the respondent's patience is included as a control, as it might be relevant for weighing costs and benefits of migration.

Table 6.A1 in the Appendix presents the definitions of the explanatory variables we derive from our data. Before discussing how we utilize those determinants within our estimation strategy (see Section 6.4), we have a look at our data in the next Section.

6.3 Data and descriptive statistics

We conducted two multinational surveys: one among language course participants at the Goethe Institute (henceforth GI)[70] in 14 countries and one among university students in six countries. The survey at the GI captures a self-selected subset of the population, many of whom can be expected to be relatively far in the migration process. The survey at universities allows us to assess the generalizability of our GI results for a population of young and highly skilled individuals.

---

[70] The GI is a German cultural association which aims to promote German culture and language around the globe. For that purpose, the GI is present in 98 countries, with a total of 157 institutes. It offers language services, i.e. language courses and standardized exams, and provides information about the German culture and society with events and libraries (Goethe Institute, 2019). See also Chapter 2 for further background information. In this chapter, we stick to the following convention: when referring to the association of the Goethe Institute, we use the abbreviation "GI". When talking about a specific branch of the GI in Germany or abroad, we refer to it as "institute".

6.3.1 Survey design

Survey at the Goethe Institute

We conducted a survey among language course participants at the GI. The survey consisted of a pen and paper questionnaire containing a wide range of questions on socio-demographic characteristics, education and labour market status, language skills as well as migration plans and previous migration experience. The survey took place between June and December 2018. In order to minimize potential language barriers, the questionnaires were translated into the main language of each country by professional translators and double-checked by at least one native speaker of each language. In India, the questionnaire was in English. As we aimed at a heterogeneous sample of countries, we identified groups of countries based on combinations of the following characteristics: geographic distance to Germany, linguistic distance to German, economic development and absence or presence of migration barriers vis-à-vis Germany. We conducted the survey in at least one large institute (in terms of course participants) for each group (see Table 6.1).

Table 6.1: Country characteristics and response rates. Panel A: GI survey.

Countries	Income (GNI/capita)	EU member	Participants	Response rate
Bosnia	Upper-middle	No	270	99 %
Czechia	High	Yes	481	82 %
Great Britain	High	Yes, but leaving	480	88 %
India	Lower-middle	No	830	72 %
Indonesia	Lower-middle	No	883	55 %
Italy	High	Yes	371	86 %
Japan	High	No	293	59 %
Mexico	Upper-middle	No	491	61 %
Netherlands	High	Yes	139	67 %
Poland	High	Yes	236	69 %
Romania	High	Yes	327	88 %
South Korea	High	No	470	63 %
Spain	High	Yes	611	83 %
Ukraine	Lower-middle	No	782	93 %

Response rates based on registered course participants, not actual attendance. High-income countries include countries which have a GNI per capita larger than $12,535 in current US-Dollars, as of 2020 (Czechia, Great Britain, Italy, Japan, Netherlands, Poland, Romania, South Korea, and Spain); middle-income countries (upper-middle and lower-middle) are countries which have a GNI per capita of $1,036 to $12,535 in current US-Dollars, as of 2020 (Bosnia, India, Indonesia, Mexico, and Ukraine). EU membership as of 2018.

In order to maximize the response rates, we took two measures. First, a member of the project team was present during the survey in all institutes in all

European countries. During one week of the course term, the pen and paper questionnaires were personally and directly distributed to all course participants present in the classroom. Participants then either filled in the questionnaire during the course break or after the course, or they took it home and returned it at a second course session within that week. For non-European countries, we sent the questionnaires by mail to the institutes, where the teachers distributed and collected the questionnaires, which were then sent back to Germany. In order to minimize errors in distributing the questionnaires, we prepared envelopes for each course containing the questionnaires, which were distributed to the respective teachers of the courses. Second, for each country, we raffled off one free language course at the survey institute in order to incentivize participation. Those measures resulted in response rates ranging from 67 to 99 % in European countries, and 59 to 72 % in non-European countries.[71] Table 6.1 gives an overview of the countries the survey was conducted in and their respective characteristics, as well as the number of participants and the response rates.

In total, 6,664 language course participants at institutes in 14 countries took part in the GI survey. Of those, 1,554 individuals had to be excluded from the analysis due to missing values in relevant variables or because they were foreign-born. This leaves us with a sample of 5,110 individuals. Descriptive statistics can be found in Table 6.A2 in the Appendix.

Survey among university students

Additionally, we conducted a survey among university students, which was designed similarly to the GI survey and contained the same questions on socio-demographic characteristics, education and labour market status, language

---

[71] In all European countries except the Netherlands, these numbers refer to actually distributed questionnaires. In the Netherlands and in non-European countries, the response rate is related to registered course participants, as we do not know the number of course participants who were present when the questionnaires were distributed. In the European institutes where a member of the project team was present, not all registered participants attended every lesson of their course, i.e. the number of registered participants is much larger than the number of present participants in many cases. Therefore, the response rate for non-European countries and the Netherlands, which is related to the number of registered course participants and not to the number of present course participants, gives a lower-bound. In Mexico, the response rate is related to the distributed questionnaires during the course inscription.

skills as well as migration plans and previous migration experience. For the survey, three European and three non-European countries were chosen, which were also part of the GI sample. The survey was conducted between April 1, 2019 and April 7, 2020 in all universities that agreed to participate, either at the university level or in selected faculties. Table 6.2 gives an overview of the participating universities in the respective countries, as well as the number of participants (for country characteristics, see Table 6.1).

Table 6.2: Countries, universities and participants. Panel B: Student survey.

Countries	Participating universities	Participants
Czechia	Masaryk University, University of Ostrava, University of Economics Prague	1,078
India	IIT Kanpur, Ashoka University	563
Indonesia	Institut Pertanian Bogor, Universitas Indonesia	
	Institut Teknologi Bandung, Politeknik Manufaktur Bandung, Universitas Padjadjaran	251
Italy	Universita Cattolica del Sacro Cuore	289
Mexico	El Colegio de Mexico, Centro de Investigacion y Docencia Economicas,	
	Universidad Nacional Autonoma de Mexico, Instituto Tecnologico Autonomo de Mexico	1,244
Spain	Universitat Autonoma de Barcelona, University of Barcelona, Carlos III University of Madrid	794

As the survey was conducted as an online survey, local university staff sent invitation e-mails including a link to the survey to students. Again, to minimize potential language barriers, the survey was offered in the local language and in English in all countries, apart from India, where the survey was only available in English. To incentivize participation, individuals could take part in a lottery, which was embedded in the questionnaire. The first prize in each survey was a cash payout of EUR 100, and there was also an opportunity to participate in two other lotteries with additional prizes, which depended on choices that respondents made. The largest single prize won among all participants was EUR 250.

Of the 4,219 students who participated in the survey, 977 individuals were excluded from the analysis due to missing values in relevant variables or because they were foreign-born. This leaves us with a sample of 3,242 students. Descriptive statistics can be found in Table 6.A2 in the Appendix.

6.3.2 Descriptive statistics

Both resulting datasets have limitations, and are not representative of a general population. Our study focuses on specific self-selected groups, as the surveyed individuals are either participants of language courses or university students.

Both groups are more likely to be better educated and more likely to aspire to migrate than the overall population, the latter especially applying to language course participants.

While individual characteristics in our GI and student samples are fairly comparable (see Table 6.A2 in the Appendix), there are some differences. As can be expected, occupational status and age differ between samples: Only 32.7 % of the GI sample are students compared to 100.0 % in the student sample, and the share of respondents in the age group of 18 to 24 years is only 38.7 % in the GI sample compared to 76.8 % in the student sample. Given the difference in age distribution, the share of respondents with children is higher in the GI sample than in the student sample, and the distribution over nationalities differs.[72] Both samples are especially comparable when it comes to the shares of migration aspirations and intentions among the younger age groups. Thus, in addition to running our main analysis on both samples separately, we will also restrict the GI sample to the younger age groups (under 35 years of age) in our subsequent analysis and compare the results.

Migration aspirations

A comparison of migration aspirations for each country based on data from the GWP and data from our GI and student samples illustrates this. In Figure 6.2, we see that in all countries the share of people that would ideally like to migrate permanently is much higher in the GI and student samples than in the representative GWP data: In the GI sample, we find the highest share of respondents who would ideally like to migrate permanently in Mexico with 90 % (compared to 25 % in the GWP data) and the lowest share in Czechia with 28 % (compared to 12 % in the GWP data). The largest discrepancy between the two groups can be observed for India (73 % in GI versus 6 % in GWP) and Indonesia (70 % in GI versus 2 % in GWP), while the smallest discrepancy can be observed for Italy (50 % in GI versus 36 % in GWP). In the student sample, the share of respondents with migration aspirations is higher than in the GWP in all countries as well. Yet, as expected, it is smaller than in the GI sample, with the exception of Czechia.

---

[72] Note that both the GI and the student samples have considerably high shares of respondents with migration experience (33.3 % and 34.6 %, respectively). This can be explained by how we defined migration experience – at least one prior stay abroad for at least three consecutive months.

Figure 6.2: Share of respondents with migration aspirations. GWP, GI and student sample.
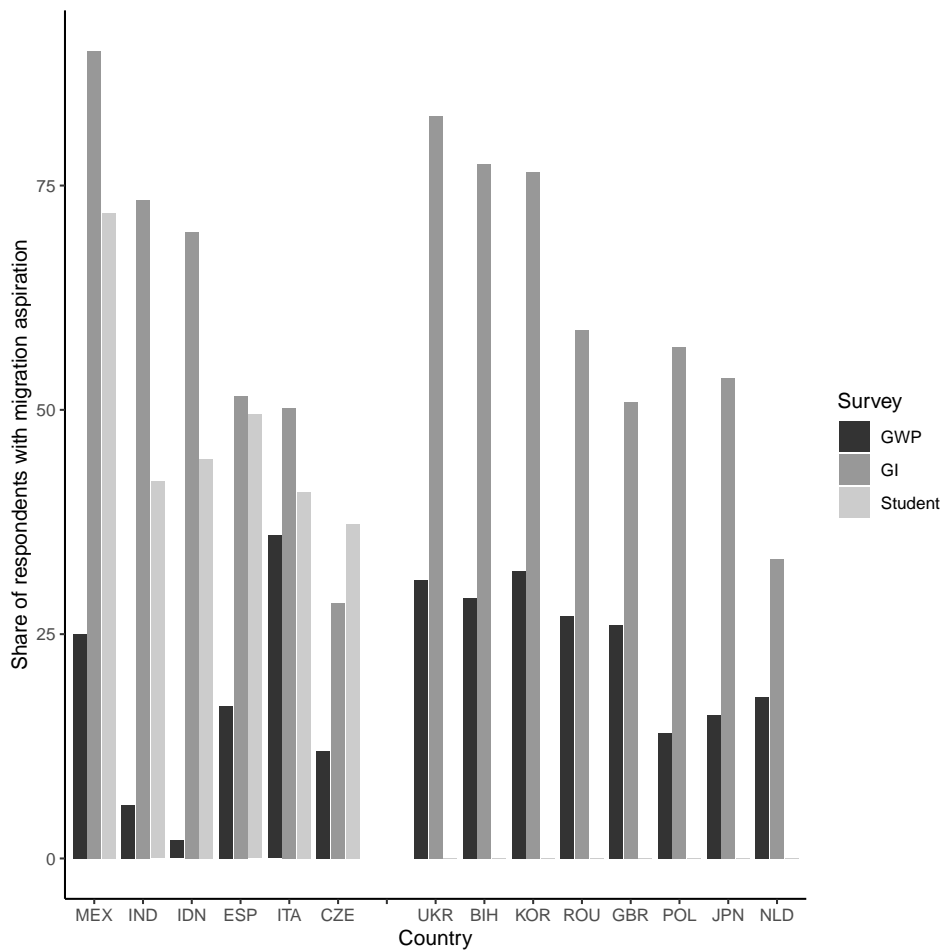
Note: Compares the shares of respondents with migration aspirations between the GWP data, and the GI and student samples. Data from the GWP refers to the 2018 wave, apart from data for Spain, Great Britain, Italy and the Netherlands that refers to the 2017 wave, as data for 2018 was not available.

The over-representation of individuals with migration aspirations in our GI sample relative to the GWP data can also be seen in Figure 6.3, where we plot the share of individuals with migration aspirations in the GWP data and in the GI sample by age groups. There are two observations: For all countries and age groups, the share of individuals with migration aspirations in the GI sample exceeds the corresponding share in the GWP data. At the same time, we find for most countries that the shares are most comparable for younger age groups and comparability deteriorates for middle-aged and older age groups. With respect to their stated migration aspirations, the younger individuals in the GI sample are thus closer to the general population of the same age group in the respective countries than the older individuals; among the older individuals in the GI sample, many more state migration aspirations than in the general population.

Due to the different age structure in the student sample and the GI sample, the comparison with the student sample is most appropriate when the GI sample and the GWP data are also restricted to the younger age group. Figure 6.4 shows that within the younger age groups, the share of respondents with migration aspirations in the student sample is most comparable to the share of respondents with migration aspirations in the GI sample. Still, it is smaller for all countries, with the exception of Czechia. Only in Italy, the share of student respondents with migration aspirations is lower than in the GWP data.

Migration intentions

With regard to migration intentions, for a direct comparison of the GI and student sample with the GWP data we would need to restrict the observations to only those with migration aspirations due to the problem of conditionality in responses (see Section 6.2.3). For this reason, we focus on the GI and student samples only. Compared to the GI sample, migration intentions are differently distributed in the student sample. Figure 6.5 shows for the younger age groups that the share of those without migration intentions is larger in the student sample than in the GI sample, and the share of those who intend to migrate permanently is consistently larger in the GI sample than in the student sample. Shares of those who intend to migrate temporarily are more comparable over both samples, and in Spain all shares are comparable.

Figure 6.3: Share of respondents with migration aspirations by age groups. GWP and GI sample.

Figure 6.4: Share of respondents with aspirations. GI sample and student sample, younger age groups.

Figure 6.5: Share of respondents with migration intentions. GI and student sample, younger age groups.

Note: Compares shares of migration intentions between the GI and the student samples, both restricted to the younger age groups (under 35 years of age), for the 6 countries which are observed in both samples. For the equivalent figure for the GI and the student samples including all age groups, see Figure 6.A1 in the Appendix.

189

A concern is that we do not observe actual migration but survey potential migrants in their origin country. While it is true that we do not know their future migration behavior, this approach allows us to collect information about migration aspirations and intentions from those who end up migrating as well as from those who stay, while post migration data only covers a selective subsample of those who actually migrated. Also, previous studies have shown that migration intentions are indeed a strong predictor of subsequent migration behavior (Chort, 2014; Tjaden et al., 2019). Since collecting data prior to potential migration can be cumbersome and cost-intensive, survey data on migration aspirations and intentions is scarce. To the best of our knowledge, the only globally representative data available on migration aspirations and intentions is supplied by the GWP. However, the GWP suffers from the conditionality in responses problem, as discussed in Section 6.2.3. Our dataset allows overcoming the conditionality in responses problem and additionally provides insights into intentions for different time horizons. Thus, we can identify certain migration patterns which the literature based on the GWP cannot, e.g. individuals who do not aspire to migrate permanently yet intend to migrate permanently or temporarily – groups of potential migrants which are of high interest for migration policy-makers and have not received much attention in the migration literature so far.

Joint distribution of migration aspirations and intentions

To observe in how far aspirations and intentions to migrate match in our data, we check their joint distribution in Table 6.3. Indeed, all six outcomes shown in Figure 6.1 occur in both samples.

In the GI sample (Panel A), we can make several observations: First, 38.6 % of respondents both aspire and intend to migrate permanently, and 22.1 % do neither aspire nor intend to migrate. This makes 60.7 % for whom aspirations and current intentions match perfectly, and leaves 39.3 % of respondents for whom they do not. Second, similar to Delogu et al. (2018), those intending to migrate temporarily are a substantial share of potential migrants that is overlooked in surveys asking only for intentions to migrate permanently. Indeed, we find that 24.6 % of respondents in the GI sample intend to migrate only temporarily. We will come back to this in Section 6.5.2. Third, we observe clear differences between aspirations and intentions to migrate permanently: 11.8 % of respondents aspire to migrate permanently yet have no current intentions to do so

Table 6.3: Joint distribution of aspirations and intentions.

Panel A: GI sample

	no intentions to migrate		intentions to migrate temporarily		intentions to migrate permanently		*Total*	
no aspirations to migrate permanently	1127	(22.1)	466	(9.1)	150	(2.9)	1743	(34.1)
aspirations to migrate permanently	602	(11.8)	790	(15.5)	1975	(38.6)	3367	(65.9)
Total	1729	(33.9)	1256	(24.6)	2125	(41.5)	5510	(100.0)

Panel B: Student sample

no aspirations to migrate permanently	1071	(33.0)	415	(12.8)	74	(2.3)	1560	(48.1)
aspirations to migrate permanently	483	(14.9)	609	(21.3)	509	(15.7)	1682	(51.9)
Total	1554	(47.9)	1105	(34.1)	583	(18.0)	3242	(100.0)

Note: Numbers of observation, percentages in parentheses. Percentages of total sample size in parentheses. Row and column *Total(s)* show row and column totals.

and 2.9 % of respondents intend to migrate permanently despite not desiring to do so.[73] We will come back to this in Section 6.5.3. While the distribution is somewhat different in the student sample (Panel B), all outcomes occur also in that sample. As expected, with 33.0 % a larger share of the student sample neither aspires nor intends to migrate permanently compared to the GI sample, and with 15.7 % a smaller share aspires and intends to migrate permanently. However, this still makes only 48.7 % of the student sample for whom aspirations and current intentions match perfectly. With 34.1 % we see a share of potential migrants who intend to migrate temporarily that is even larger than that in the GI sample. Lastly, 14.9 % aspire to migrate permanently yet have no current intentions to do so, and 2.3 % intend to migrate permanently without aspirations to do so.

Our framework allows us to identify three groups of potential migrants which have not been observed by studies relying on GWP data. By avoiding the condi-

---

[73] For comparison, Bertoli and Ruyssen (2018) analyse a subset of their observed GWP waves of 2007 and 2011. They find that of those who stated migration aspirations only 14.3 % also answered affirmatively to the question *"Are you planning to move to another country in the next 12 months, or not?"*, and only 42.7 % were to take concrete steps towards migration. Ruyssen and Salomone (2018) show that while 16.0 % of their women subsample of the GWP stated migration aspirations, only 4.0 % of those stated that they had done any preparation for this move.

tionality in responses problem, we observe those who have no desire to migrate, i.e. those who do not state aspirations to migrate permanently under ideal circumstances but who nonetheless consider or plan to migrate permanently. By introducing a time horizon to the question on intentions, we observe those who do not aspire to migrate permanently yet intend to migrate temporarily (9.1 % of respondents in the GI sample, and 12.8 % in the student sample) and those who aspire to migrate permanently yet only intend to migrate temporarily (15.5 % of respondents in the GI sample, and 21.3 % in the student sample).

Main motivations for migration intentions

Exploring motivations for potential migration can shed some light on the reasons why individuals intend to migrate permanently versus temporarily, or why they intend to migrate despite having no aspirations to do so. In our surveys, respondents who stated intentions to migrate were asked to name their preferred destination country for a potential move abroad and indicate the main reason for such a move (see Table 6.A1 in the Appendix for a detailed definition).

Table 6.4 shows how the shares of main motivations for intentions to move temporarily or permanently are distributed across educational reasons, professional reasons, family and partner, and other reasons among men and women in both samples. We find that main motivations for potential migration to the preferred destination differ between intentions to migrate temporarily and permanently, as well as between men and women. Temporary migration seems to be considered primarily for educational and professional reasons. The share of respondents who state either educational or professional reasons as the main motivation for potential migration is consistently larger among those who intend to migrate temporarily compared to those who intend to migrate permanently, for both genders and both samples. When it comes to gender differences, the share of respondents who state educational or professional reasons as their main motivation for either intentions to migrate temporarily or permanently is consistently larger among men compared to women. Family ties, on the contrary, are more strongly linked to intentions to migrate permanently than temporarily. The share of respondents who state family and partner as the main motivation for potential migration is consistently larger among those with intentions to migrate permanently than those with intentions to migrate temporarily. Also, family ties play a larger role among women, as the share of

Table 6.4: Main motivations for potential migration, by gender.

	women		men	
	intentions to migrate temporarily	intentions to migrate permanently	intentions to migrate temporarily	intentions to migrate permanently

Panel A: GI sample

	(n = 486)	(n = 762)	(n = 305)	(n = 522)
Educational reasons	43.2	37.9	39.0	34.9
Professional reasons	26.3	22.8	34.4	28.9
Family and partner	9.5	15.2	4.3	9.4
Other reasons	21.0	24.1	22.3	26.8

Panel B: Student sample

	(n = 629)	(n = 311)	(n = 469)	(n = 270)
Educational reasons	37.4	28.0	43.7	36.3
Professional reasons	29.6	27.0	27.5	28.1
Family and partner	5.4	10.3	4.1	4.5
Other reasons	27.6	34.7	24.7	31.1

Note: Column percentages. *n* show column totals. Numbers of observation differ from sample totals since only respondents who stated intentions to migrate answered the question on main motivations for potential migration and not all of those who intend to migrate indicated such main reason. Hence, in the GI sample we observe main reasons for potential migration for 1248 of 1929 women with intentions to migrate, and for 827 of 1349 men with intentions to migrate. In the student sample we observe main reasons for potential migration for 940 of 940 women with intentions to migrate, and for 739 of 741 men with intentions to migrate.

respondents who state family and partner as their main motivation for either intentions to migrate temporarily or permanently is consistently larger among women.

We explore potential differences between aspirations and intentions in more detail in Table 6.5, which shows how the shares of main motivations for intentions to move temporarily or permanently are distributed among those with and without aspirations to migrate permanently, separately for men and women in both samples. Again, we find gender differences in motivations for migration. For those without aspirations to migrate permanently, the share of respondents who state family and partner as the main motivation for intentions to migrate permanently is consistently larger among women. Contrarily, the share of respondents who state either educational or professional reasons is consistently larger for the men with intentions to migrate permanently despite having no aspirations to do so.

6.4 Estimation strategy

In Section 6.3 we showed how migration aspirations and intentions are related in our data and which main motivations respondents state for the latter. Now, we study individual-level determinants of aspirations and intentions, and how aspirations predict intentions. Then we explore potential differences between the two.

6.4.1 Relation between aspirations and intentions to migrate

In Section 6.5.1, we explore individual-level determinants of the probability of aspirations and intentions and how they relate. For this, we estimate the probability of having aspirations to migrate in a binary probit model via maximum likelihood method:

$$Pr(aspirations_i| \ X_i, C_i) = \alpha_1 + \beta_1 ' X_i + \gamma_1 ' C_i + \varepsilon_{1,i} \qquad (6.1)$$

Table 6.5: Main motivations for potential migration, by aspirations and intentions.

	no aspirations to migrate permanently		aspirations to migrate permanently	
	(1) intentions to migrate temporarily	(2) intentions to migrate permanently	(3) intentions to migrate temporarily	(4) intentions to migrate permanently
Panel A: GI sample, women				
	(n = 167)	*(n = 57)*	*(n = 319)*	*(n = 705)*
Educational reasons	32.3	19.3	48.9	39.4
Professional reasons	29.3	31.6	24.8	22.2
Family and partner	12.6	22.8	7.8	14.6
Other reasons	25.8	26.3	18.5	23.8
Panel A: GI sample, men				
	(n = 116)	*(n = 49)*	*(n = 189)*	*(n = 473)*
Educational reasons	44.0	20.4	36.0	36.4
Professional reasons	37.9	38.8	32.2	27.9
Family and partner	2.6	14.3	5.3	8.9
Other reasons	15.5	26.5	26.5	26.8
Panel B: Student sample, women				
	(n = 217)	*(n = 40)*	*(n = 412)*	*(n = 271)*
Educational reasons	37.3	25.0	37.4	28.4
Professional reasons	29.5	37.5	29.6	25.5
Family and partner	6.0	17.5	5.1	9.2
Other reasons	27.2	20.0	27.9	36.9
Panel B: Student sample, men				
	(n = 195)	*(n = 34)*	*(n = 274)*	*(n = 236)*
Educational reasons	44.6	44.1	43.1	35.2
Professional reasons	30.3	41.2	25.6	26.3
Family and partner	4.6	2.9	3.6	4.7
Other reasons	20.5	11.8	27.7	33.9

Note: Column percentages. *n* show column totals. Notes from Table 6.4) regarding sample size also apply here.

where $aspirations_i$ takes a value of 1 if respondent i states having aspirations to migrate and 0 otherwise. X_i represents a set of individual-specific explanatory variables of respondent i: gender, age, university degree, student, migration experience, partner and children, willingness to take risks and a measure of patience. C_i is a country dummy. $\varepsilon_{1,i}$ is an idiosyncratic error term. Table 6.A1 in the Appendix presents the definitions of all explanatory variables, and Table 6.A2 shows descriptive statistics.

Also, we estimate the probability of having intentions to migrate. Since we differentiate between time-horizons of migration intentions, we adapt a multinomial probit model and estimate it with the maximum likelihood method:

$$Pr(intentions_i \mid X_i, C_i) = \alpha_2 + \beta_2{}' X_i + \gamma_2{}' C_i + \varepsilon_{2,i} \qquad (6.2)$$

where $intentions_i$ can take any of the following outcomes: no migration intentions, intentions to migrate temporarily or intentions to migrate permanently. To investigate in how far aspirations to migrate are linked to the probability of intentions to migrate, we repeat Equation (6.2) and include respondent $i's$ aspirations to migrate as an explanatory variable:

$$Pr(intentions_i \mid X_i, C_i) = \alpha_3 + \zeta * aspirations_i + \beta_3{}' X_i + \gamma_3{}' C_i + \varepsilon_{3,i}$$
$$(6.3)$$

We hypothesize that ζ should be positive, as we expect intentions to migrate, both temporarily and permanently, to correlate with migration aspirations. Again, X_i is a set of the above listed individual-specific explanatory variables, C_i are controls, and $\varepsilon_{2,i}$ and $\varepsilon_{3,i}$ are idiosyncratic error terms.

In Section 6.5.2, we take a closer look at Equations (6.2) and (6.3) and exploit the differentiation between intentions to migrate temporarily and permanently in the multinomial probit model. This allows us to focus on the differences in determinants for a different time horizon.

6.4.2 Conditionality and differences

In Section 6.5.3, we explore whether determinants of intentions to migrate, either temporarily or permanently, differ between those with and without aspirations to migrate. We estimate Equation (6.2) conditional on the respondent's $aspirations_i$ in a multinomial probit model with the maximum likelihood

method:

$$Pr(intentions_i \mid aspirations_i, \ X_i, C_i) = \alpha_4 + \beta_4{}' X_i + \gamma_4{}' C_i + \varepsilon_{4,i} \quad (6.4)$$

with a special focus on those who have no underlying migration aspirations –
a group of potential migrants which has been largely neglected so far because
of the conditionality problem mentioned above. Results in the following are
shown for both the GI and student samples separately.

6.5   Results

6.5.1   Relation between aspirations and intentions to migrate

In this section, we analyse which individual-level factors determine the prob-
ability of migration aspirations and intentions and the extent to which the
first correlates with the latter. For this, we first adapt Equation (6.1) to ex-
plore individual-level determinants of the probability of aspirations. Column
(1) in Table 6.6[74] shows the results. Second, we adapt Equations (6.2) and
(6.3) to explore determinants of the probability of intentions. Columns (2)
and (3) in Table 6.6 show the results for intentions to migrate temporarily and
permanently, respectively, without controlling for aspirations (Equation 6.2),
and Columns (4) and (5) do so with aspirations as an additional explanatory
variable (Equation 6.3).[75]

We run all regressions separately for both samples to check how our findings
apply to different populations, and we find largely similar results. Both samples
show a strong and significant relation between aspirations and intentions. In
the GI sample, those with aspirations to migrate permanently are three times

---

[74]   Specification (1) estimates the probability of aspirations to migrate permanently via bino-
mial probit; reference category is 'no aspirations to migrate permanently' ($n = 1743$ (GI
sample), $n = 1560$ (student sample)). Specifications (2) and (4) estimate the probability
of intentions to migrate temporarily, and specifications (3) and (5) estimate the proba-
bility of intentions to migrate permanently via multinomial probit; reference category is
'no intentions to migrate' ($n = 1729$ (GI sample), $n = 1554$ (student sample)).
[75]   As discussed in Section 6.2.3 and shown in Figure 6.1 intentions to migrate temporarily
and intentions to migrate permanently are understood as mutually exclusive outcomes.

as likely to intend to migrate temporarily, and they are 23.1 times as likely to intend to migrate permanently compared to those without aspirations. In the student sample, magnitudes differ slightly, yet the direction is the same: those with aspirations are almost four times as likely to intend to migrate temporarily and 15.3 times as likely to intend to migrate permanently. While aspirations and intentions are strongly related and share explanatory determinants, their relationship is far from perfect and after controlling for aspirations, other determinants remain important in explaining intentions (compare Columns (2) and (3) with Columns (4) and (5) in Table 6.6). We discuss some of these factors in the following, with a focus on gender and family ties.

Women are more likely to state aspirations than men in both samples. This can be seen in Table 6.A3 in the Appendix, where we compare the joint distribution of aspirations and intentions by gender (as we did in Table 6.3 for the whole samples). 66.3 % of women in the GI sample state to have migration aspirations compared to 65.1 % of men. In the student sample, the difference is more clear. Here 53.7 % of women state aspirations to migrate compared to 49.7 % of men. Hence, the regression results in Table 6.6 show that being a women is positively associated with the likelihood of aspiration in both samples, though the result is only significant in the student sample. This is driven by women from middle-income countries, as we see when we compare results between middle- and high-income countries for both samples (see Table 6.A4 for high-income countries and Table 6.A5 for middle-income countries in the Appendix). Yet, this does not translate into intentions. As seen in the joint distribution of aspirations and intentions by gender in Table 6.A3, despite higher shares of aspirations to migrate permanently, women show lower shares of intentions to migrate permanently in both samples. Consequently, in the regression results in Table 6.6 we do not find gender to be positively associated with intentions to migrate in either sample. On the contrary, in the GI sample, women are significantly less likely to intend a permanent move abroad after controlling for their aspirations. This finding is in line with what Ruyssen and Salomone (2018) have shown with GWP data: women, especially in countries where gender discrimination is comparably high, are more likely to aspire to migrate, but they also face costs and obstacles in realizing these aspirations. Coulter et al. (2012), too, have shown that women are less likely to realize their migration aspirations.

Table 6.6: Aspirations and intentions.

Panel A: GI sample

	binomial probit	multinomial probit		multinomial probit	
	(1) aspirations to migrate permanently	(2) intentions to migrate temporarily	(3) intentions to migrate permanently	(4) intentions to migrate temporarily	(5) intentions to migrate permanently
	(n = 3367)	(n = 1256)	(n = 2125)	(n = 1256)	(n = 2125)
Migration aspiration				3.019***	23.153***
				(0.086)	(0.107)
Gender: woman	1.064	1.055	0.913	1.009	0.822**
	(0.040)	(0.081)	(0.073)	(0.084)	(0.084)
Age: under 18 years	1.257***	0.812	1.291*	0.749*	1.104
	(0.082)	(0.164)	(0.144)	(0.170)	(0.163)
Age: 25 to 34 years	1.049	1.032	1.408***	1.058	1.481***
	(0.066)	(0.132)	(0.122)	(0.137)	(0.139)
Age: 35 to 49 years	0.845**	0.518***	0.750*	0.551***	0.836
	(0.084)	(0.168)	(0.156)	(0.173)	(0.179)
Age: 50 to 64 years	0.567***	0.146***	0.331***	0.169***	0.430***
	(0.114)	(0.260)	(0.220)	(0.266)	(0.259)
Age: 65 years or above	0.347***	0.065***	0.099***	0.084***	0.191***
	(0.163)	(0.416)	(0.412)	(0.419)	(0.458)
University degree	0.849***	1.312**	0.903	1.412***	1.047
	(0.055)	(0.109)	(0.100)	(0.113)	(0.114)
Student	0.949	1.008	0.818*	1.018	0.830
	(0.061)	(0.119)	(0.109)	(0.124)	(0.124)
Migration experience	1.055	1.247**	1.291***	1.268**	1.316***
	(0.046)	(0.092)	(0.085)	(0.095)	(0.097)
Partner: native	1.027	0.930	1.089	0.919	1.083
	(0.054)	(0.107)	(0.101)	(0.111)	(0.115)
Partner: non-native	1.086	1.050	1.742***	1.089	1.919***
	(0.082)	(0.173)	(0.157)	(0.179)	(0.182)
Children	0.914	0.807	0.950	0.827	1.041
	(0.074)	(0.160)	(0.139)	(0.164)	(0.163)
Willingness to take risks	1.066***	1.093***	1.134***	1.075***	1.099***
	(0.010)	(0.020)	(0.018)	(0.020)	(0.020)
Patience	0.992	0.981	0.979	0.983	0.982
	(0.008)	(0.017)	(0.015)	(0.018)	(0.018)
Number of observations	5110	5110	5110	5110	5110
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	72.5	52.1	52.1	62.3	62.3
McFadden Pseudo R2	0.13	0.09	0.09	0.2	0.2

Panel B: Student sample

	(n = 1682)	(n = 1105)	(n = 583)	(n = 1105)	(n = 583)
Migration aspiration				3.836***	15.295***
				(0.091)	(0.143)
Gender: woman	1.156***	1.055	1.123	0.982	0.998
	(0.049)	(0.086)	(0.107)	(0.090)	(0.115)
Age: 25 to 34 years	0.979	1.195	0.994	1.215	1.030
	(0.070)	(0.122)	(0.151)	(0.127)	(0.164)
Age: 35 to 49 years	0.869	0.754	0.720	0.775	0.792
	(0.138)	(0.247)	(0.307)	(0.256)	(0.333)
University degree	0.934	1.331***	1.175	1.436***	1.337**
	(0.052)	(0.092)	(0.114)	(0.096)	(0.124)
Migration experience	1.254***	1.726***	1.803***	1.636***	1.632***
	(0.051)	(0.090)	(0.112)	(0.094)	(0.121)
Partner: native	0.846***	0.877	1.005	0.956	1.147
	(0.051)	(0.091)	(0.114)	(0.095)	(0.124)
Partner: non-native	1.235*	1.420	3.500***	1.456	3.609***
	(0.124)	(0.242)	(0.252)	(0.254)	(0.279)
Children	1.071	0.628*	0.457**	0.564**	0.364***
	(0.134)	(0.242)	(0.325)	(0.251)	(0.357)
Willingness to take risks	1.045***	1.126***	1.089***	1.112***	1.071**
	(0.011)	(0.021)	(0.025)	(0.022)	(0.027)
Patience	0.995	1.008	1.022	1.015	1.030
	(0.010)	(0.017)	(0.021)	(0.018)	(0.023)
Number of observations	3242	3242	3242	3242	3242
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	65.5	52.1	52.1	57.2	57.2
McFadden Pseudo R2	0.08	0.05	0.05	0.14	0.14

Note: Risk ratios, standard errors in parentheses. See further notes in footnote [74].
*p<0.1; **p<0.05; ***p<0.01

When it comes to social ties, the influence of ties at home should be regarded as well as the importance of networks abroad. Manchin and Orazbayev (2018), for example, find that networks abroad account for about 18 % of variation in international migration intentions, and social networks in the country of residence account for 2-4 %. While we do not focus on social networks in such a broad sense but only on family ties – i.e. partner and children – in our data, we find that family ties in the origin country as well as abroad play a role. While having a native partner at home is linked to a decreased likelihood of migration aspirations in the student sample, this does not translate into a significantly decreased likelihood of migration intentions for respondents with native partners. This only holds for women in the student sample (see Table 6.A6 which analyses only women), and no such result is found for men (see Table 6.A7 which analyses only men). The results regarding family ties abroad proxied by a non-native partner are, on the other hand, strong and robust. Being socially tied to a partner who has roots and contacts abroad makes it almost 2 times as likely to intend to migrate permanently, after controlling for underlying aspirations for respondents in the GI sample, and even 3.6 times as likely for those in the student sample. Subsampling by gender sheds more light on the importance of family ties (see Tables 6.A6 and 6.A7 for the respective subsamples). The positive linkage between having a non-native partner and an increased likelihood of intentions to move abroad permanently is large and robust for women in both samples – a finding in line with the theory that especially women are often so-called 'tied movers' who follow their partner (Coulter et al., 2012; Geist and McManus, 2011; Mincer, 1978). Our exploration of main motivations for potential migration in Table 6.4 underlines those results. In the GI sample, we find that for 15.2 % of the women who intend to migrate permanently the main motivation are family and partner, such as professional reasons or studies of a partner or having a partner or relatives who live in the destination country, compared to only 9.4 % of men. In the student sample it is the main motivation for intending a move to the preferred destination for 10.3 % of women with intentions to migrate permanently, and only 4.5 % of men.

We find a positive linkage between migration experience and intentions. However, this linkage remains robust after controlling for aspirations only among women (see Table 6.A6 in the Appendix) - a result in line with De Jong (2000), who also finds previous experience to be a determinant for future intentions only for women. This gender disparity could be rooted in the different roles family ties abroad play in migration decisions for men and women. Another possible explanation for this gender disparity are differences in risk preferences

between men and women. While an increase in willingness to take risks increases the likelihood of both aspirations and intentions for both genders, on average, in the GI sample, men have a higher willingness to take risks (with a mean of 0.66) than women (with a mean of 0.63). The same holds in the student sample, in which the average willingness to take risks is 0.54 for women and 0.57 for men.[76] Though those differences in means are not large, Welch two-sample t-tests reveal that both differences are strongly significant. 66.7% of men in the GI sample intend to migrate either temporarily or permanently, and 65.5% of women do. A simple OLS regression on this outcome shows that ceteris paribus the marginal effect of an increase of 0.03 in willingness to take risks (which is the difference between womens' and mens' average) can explain roughly 27% of the gender difference in migration intentions. Based on a similar calculation, in the student sample the differences in average willingness to take risks account for roughly 18% of the gender differences in intentions to migrate. In any case, previous migration is positively linked to the likelihood to intend migration for women only.

Lastly, the wide age distribution of the GI sample allows for a closer look at how age is linked to aspirations and intentions. As expected, from age 35 on, the likelihood of both aspirations and intentions to emigrate consistently decreases with age. In the student, sample age does not show much variation (with 76.8% of respondents being between 18 and 24 years old), and the age variable has little impact.

Overall, the influence of individual characteristics on migration aspirations and intentions differ more strongly between genders than between our two samples. In an attempt to make the GI and student samples even more comparable, we next restrict the GI sample to the younger age groups (under 35 years of age) and to the 6 countries which are surveyed in the student sample. Results (see Table 6.A8) are largely comparable between this GI sample that is restricted by age and countries, and the student sample.[77] The same holds for the GI sample restricted to the younger age groups but including all 14 countries (see Table 6.A9).

---

[76] Since the scales on which willingness to take risks differ between both samples (the scale ranges between 0 and 10 in the GI sample, and between 1 and 10 in the student sample; for details see the variables description in Table 6.A1), we utilize sample means normalized between 0 and 1.

[77] We do not restrict the student sample by age since only 5.1% of the student sample are 35 years or older.

6.5.2 Intentions to migrate temporarily and permanently

In Table 6.3, we have seen that 24.6 % of respondents in the GI sample and 34.1 % in the student sample state intentions to migrate temporarily. This is similar to Delogu et al. (2018), which shows with GWP data that individuals who intend to migrate temporarily are a considerable share of potential migrants. Thus, we will now take a closer look at those with intentions to migrate temporarily and how they differ from those with intentions to migrate permanently.

Table 6.6 allows us to compare how individual characteristics are linked to intentions to migrate temporarily versus permanently (see Columns (2) and (3) as well as Columns (4) and (5), respectively). We find that university graduates are 1.4 times as likely to intend a temporary move abroad in both the GI and student sample (after controlling for aspirations), while no such positive linkage is found between degree and aspirations in either sample. This finding might be driven by the pull of a more international labour market and better professional opportunities abroad, given that the human capital of highly educated people is more easily transferable. It might as well be due to individuals seeking further education abroad. Indeed, in our exploration of main motivations for potential migration in Table 6.4, we find that in the GI sample for 39.0 % of men who intend to migrate temporarily educational reasons such as studies abroad are the main motivation, and for 34.4 % professional reasons are. For women in the GI sample who intend to migrate it is 43.2 % and 26.3 %, respectively. In the student sample for 43.7 % of men (37.4 % of women) who intend to migrate temporarily educational reasons are the main motivation, and for 27.5 % (29.6 % of women) professional reasons are. As Docquier et al. (2014) argue, college-educated individuals do not necessarily show higher shares of aspirations to migrate, even though their actual emigration rates are much larger compared to those of the less educated. Here, too, gender differences occur. The positive association between degree and intentions to migrate temporarily is largely driven by men in both samples (see Table 6.A7 in the Appendix).

A closer look at the different age groups in the GI sample reveals some insight into a combined effect of education and age: The positive association between university degree and intentions to migrate temporarily is evident and robust only among the younger age groups (under 35 years of age). This can be seen in Table 6.A8, where the GI sample is split by age and respondents under the age of 35 years and aged 35 or older are analysed separately in the six countries

which are observed in both samples. As seen in Table 6.A9, this also holds for the GI sample when including all countries. Among those aged 35 years and older, no significant association between degree and intentions to migrate is found, neither in the GI sample restricted nor unrestricted on the country level (see Table 6.A8 and Table 6.A9, respectively). Differences in gender prevail also among the younger age groups. Splitting the respondents of the 14-country GI sample under 35 years of age by gender shows again that the positive association between university degree and intentions to migrate temporarily only occurs among men, while no such result is found for women (see Table 6.A10 in the Appendix). Overall, we find education – and with it career prospects – to be of larger importance regarding migration choices for men than for women, a finding supported by migration literature, e.g. Geist and McManus (2011), Mckinnish (2008), and Uebelmesser (2006). In our exploration of main motivations for potential migration in Table 6.4, we find that for 34.3 % of the men who intend to migrate temporarily professional reasons, such as expected higher income or a more interesting job, are the main motivation, compared to only 26.3 % of the women. Of the men who intend to migrate permanently in the GI sample 28.9 % state professional reasons as their main motivation, while only 22.8 % of the women do. In the student sample the differences are not as clear. Shares of men and women who intend to migrate and state professional reasons as their main motivation are fairly comparable here.

Family ties abroad, proxied by a non-native partner, significantly increase the likelihood of intention to migrate permanently. This result is driven by women in both samples – underlining the notion that especially women are 'tied movers' – while for men the association is positive but not significant (see Tables 6.A6 and 6.A7 in the Appendix).

Lastly, in the GI sample, we see that increasing age is negatively correlated with the intentions to migrate temporarily and permanently. Again, in the student sample we see no clear linkage between age and intentions due to the low variation in age distribution among students.

Overall, there are some factors which similarly predict intentions to migrate temporarily and permanently: Migration experience and a high willingness to take risks make both more likely, whereas increasing age makes both less likely. Nevertheless, there are important differences. Temporary migration seems to be considered primarily for educational and professional reasons. This is underlined by our exploration of main motivations for potential migration to the preferred destination country in Table 6.4. The share of respondents who state

educational reasons as the main motivation for potential migration is consistently larger among those who intend to migrate temporarily compared to those who intend to migrate permanently, for both genders and both samples. The share of respondents who state professional reasons as the main motivation is larger among those with intentions to migrate temporarily (34.4 % of men and 26.3 % of women) than those with intentions to migrate permanently (28.9 % of men and 22.8 % of women) in the GI sample. Among women in the student sample, the share of respondents who state professional reasons as the main motivation is larger among those with intentions to migrate temporarily (29.6 %) compared to intentions to migrate permanently (27.1 %). For men in the student sample, the shares are fairly close between intentions to migrate temporarily (27.5 %) and permanently (28.1 %). Family ties, on the contrary, are more strongly linked to intentions to migrate permanently than temporarily. In the GI sample, the share of respondents who state family and partner as the main motivation is larger among those with intentions to migrate permanently (9.4 % of men and 15.2 % of women) than those with intentions to migrate temporarily (4.3 % of men and 9.5 % of women). Among women in the student sample, the share of respondents who state family and partner as the main motivation is larger among those with intentions to migrate temporarily (10.3 %) compared to intentions to migrate temporarily (5.4 %). For men in the student sample, the shares are again fairly close between intentions to migrate permanently (4.5 %) and temporarily (4.1 %).

6.5.3   Conditionality and differences

In this section, we want to investigate which factors explain differences between aspirations and intentions in our data, and we focus on the relevance of the conditionality problem. For this, we estimate the multinomial probit models in Equation (6.4), conditional on the respondents' aspirations to migrate. Again, we analyse both our samples. Results are shown in Table 6.7.[78] Columns (1) and (2) show the probability of intentions to migrate temporarily and permanently for those without aspirations, and Columns (3) and (4) do this for those

---

[78]   In addition, we estimated a multinomial probit model with all 6 outcomes separately. Overall, results are comparable, but since the reference category here is always those with neither aspirations nor intentions, the model fit is inferior to our sample split in Table 6.7 and interpretation of results is not straightforward. Results are available from the authors upon request.

with aspirations, respectively. Comparing them reveals how the linkage between individual-level characteristics and migration intentions differs between those who aspire to migrate and those who do not. We see, for example, how a positive linkage between intentions to migrate and previous experience is only robust among those who aspire to migrate, and the same holds for willingness to take risks. Those and other differences in determinants of migration intentions help explain the differences between migration aspirations and intentions within our samples (as shown in Table 6.3). They will be discussed in detail in the following.

First, we concentrate on migration intentions of those respondents with no migration aspirations (Columns (1) and (2) of Table 6.7), as they have been largely neglected by the literature so far due to the conditionality in responses. As discussed before, highly educated individuals are likely to form intentions to migrate temporarily for career reasons. Consistent with these findings, in the GI sample those holding a university degree are 1.8 times as likely to intend to migrate temporarily, though without having aspirations to migrate permanently. Subsampling by gender reveals that this result is highly significant only for men in the GI sample (see Table 6.A13 for a subsample analysis for women and Table 6.A14 for men). As discussed in Section 6.5.2, the linkage between a degree and intentions to migrate temporarily is strongest among the younger age groups. This can be verified in Table 6.A15, where the GI sample is split by age and respondents under the age of 35 years and aged 35 or older are analysed separately in the six countries which are observed in both samples. When we observe only respondents aged 35 years and older, we find no association between degree and intentions to migrate temporarily. As can be seen in Table 6.A16, this also holds for the GI sample when including all countries. The importance of gender prevails, as the positive association between degree and intentions to migrate temporarily for those without aspirations among the younger respondents in the 14 country GI sample is only significant for men (see Table 6.A17 in the Appendix for a gender split.) In the student sample, the positive association between degree and intentions to migrate temporarily among those without aspirations is only significant for men. Consequently, educational and career reasons seem influential on the migration choice primarily for men.

Having a non-native partner increases the likelihood of intentions to migrate permanently among those who do not aspire to migrate, though the result is significant only in the student sample. Subsampling by gender reveals that this result is driven by women in the sample. Generally, women are often found to

support and follow their partner as 'tied movers' (Geist and McManus, 2011; Mincer, 1978), and our results suggest that they might do so even when they do not desire to migrate permanently (Coulter et al., 2012). In Table 6.5 we explored main motivations for potential migration to the preferred destination country by migration aspirations. In particular, we concentrate here on family and partner as a main motivation, which includes having a partner or relatives living in the destination country or professional reasons or studies of a partner. Family and partner are the main motivation for 22.8 % of the women in the GI sample who intend to migrate permanently despite having no aspiration to do so, while they are only for 14.3 % of the men. In the student sample they are the main motivation for 17.5 % of the women who intend to migrate permanently despite having no aspiration to do so compared to only 2.9 % of the men.

As far as migration experience is concerned, we see a relation in the student sample but not in the GI sample. In the former, respondents without aspirations but with previous experience are more likely to intend to move temporarily abroad than those without experience. This is in line with psychological literature (Ajzen, 1991), which concludes that one strong predictor of behavior is whether one has previously engaged in the same behavior, i.e. previous migration being a predictor of future migration (see e.g. DaVanzo, 1983). However, the result is significant only for women in the sample, suggesting that reducing the perceived riskiness of migration is especially important for them (De Jong, 2000).

Second, as shown in Table 6.3, 11.8 % of respondents in the GI sample, and 14.9 % in the student sample, have no current intentions to migrate despite their aspirations to migrate permanently. We utilize those respondents who aspire to migrate yet do not currently consider or plan migration as a reference group and compare to them those with intentions to migrate temporarily or permanently (see Columns (3) and (4) of Table 6.7). This reveals potential factors impeding migration.

When it comes to family ties, we find different results for our two samples. In the student sample, despite having migration aspirations, respondents with children are only 0.3 times as likely to intend to migrate temporarily and are 0.2 times as likely to intend to migrate permanently as those without children. Strong ties to their country of origin, in the form of children, can therefore increase migration costs and be a factor which impedes aspirations to migrate from being fulfilled. Having a non-native partner significantly increases the

206

likelihood of intentions to migrate permanently among those with aspirations in both samples. Subsampling by gender reveals again that this is mostly driven by the women in the samples. For men, on the other hand, education plays an important role. In the GI sample, among those with aspirations to migrate permanently, the students are only 0.7 times as likely to intend to migrate permanently as their non-student counterparts; and this result is driven by the men in the sample (see Table 6.A14 for the men subsample).

As discussed before, all older age groups (35 years and above) show a much lower likelihood of intentions to migrate, and this holds even for those with migration aspirations. As such, higher age can be a factor which impedes migration intentions, even for those who would ideally like to move abroad. This is evident in the GI sample results and most robust among respondents from high-income countries (see Table 6.A11 for a subsample analysis of high-income countries and Table 6.A12 for middle-income countries). In the student sample, interpretation of the results is not meaningful due to low variation in age among students.

Lastly, we saw in Section 6.5.1 in Table 6.6 how a higher willingness to take risks is positively associated with aspirations to migrate permanently in both samples. In Table 6.7[79] we also see how it is positively associated with the likelihood of intentions to move temporarily and permanently among those who have migration aspirations in both samples. Thus, it is associated with an increased likelihood to aspire emigration as well as an increased likelihood to consider or undergo the risky operation of emigration if the underlying aspirations exist.

---

[79] Specifications (1) and (2) estimate the probability of having intentions to migrate temporarily or permanently for those with no aspirations to migrate permanently; reference category is 'no intentions to migrate' ($n = 1127$ (GI sample), $n = 1071$ (student sample)). Specifications (3) and (4) estimate the probability of having intentions to migrate temporarily or permanently for those with aspirations to migrate permanently; reference category is 'no intentions to migrate' ($n = 602$ (GI sample), $n = 483$ (student sample)).

Table 6.7: Intentions by aspirations.

Panel A: GI sample

	no aspirations to migrate permanently		aspirations to migrate permanently	
	(1) intentions to migrate temporarily	(2) intentions to migrate permanently	(3) intentions to migrate temporarily	(4) intentions to migrate permanently
	(n = 466)	*(n = 150)*	*(n = 790)*	*(n = 1975)*
Gender: woman	0.964	0.711*	1.124	0.877
	(0.122)	(0.186)	(0.118)	(0.103)
Age: under 18 years	1.165	1.303	0.557***	0.941
	(0.261)	(0.413)	(0.223)	(0.189)
Age: 25 to 34 years	0.922	1.323	1.285	1.721***
	(0.205)	(0.338)	(0.191)	(0.171)
Age: 35 to 49 years	0.442***	1.388	0.648*	0.764
	(0.254)	(0.391)	(0.254)	(0.228)
Age: 50 to 64 years	0.163***	0.214**	0.195***	0.461**
	(0.372)	(0.653)	(0.409)	(0.327)
Age: 65 years or above	0.090***	0.171**	0.093***	0.157***
	(0.537)	(0.856)	(0.724)	(0.575)
University degree	1.805***	1.161	1.165	0.935
	(0.172)	(0.281)	(0.154)	(0.136)
Student	1.207	1.100	0.838	0.716**
	(0.192)	(0.324)	(0.165)	(0.145)
Migration experience	1.106	1.213	1.468***	1.449***
	(0.139)	(0.214)	(0.137)	(0.123)
Partner: native	0.840	0.932	1.020	1.204
	(0.160)	(0.251)	(0.165)	(0.149)
Partner: non-native	1.154	1.662	1.069	2.057***
	(0.247)	(0.368)	(0.282)	(0.251)
Children	0.864	1.006	0.834	1.074
	(0.225)	(0.304)	(0.263)	(0.225)
Willingness to take risks	1.045	1.092*	1.094***	1.110***
	(0.030)	(0.047)	(0.028)	(0.024)
Patience	0.985	0.954	0.983	0.985
	(0.026)	(0.041)	(0.024)	(0.021)
Number of observations	1743	1743	3367	3367
Country FE	✓	✓	✓	✓
Correctly predicted values	64.9	64.9	61.3	61.3
McFadden Pseudo R2	0.08	0.08	0.08	0.08

Panel B: Student sample

	(n = 415)	*(n = 74)*	*(n = 690)*	*(n = 509)*
Gender: woman	0.909	1.048	1.075	1.016
	(0.129)	(0.268)	(0.129)	(0.138)
Age: 25 to 34 years	1.093	0.909	1.430*	1.173
	(0.179)	(0.379)	(0.193)	(0.206)
Age: 35 to 49 years	0.471*	1.448	1.181	0.844
	(0.421)	(0.626)	(0.362)	(0.413)
University degree	1.246	1.451	1.607***	1.368**
	(0.135)	(0.276)	(0.141)	(0.150)
Migration experience	1.945***	0.822	1.449***	1.733***
	(0.134)	(0.310)	(0.136)	(0.146)
Partner: native	0.986	1.671*	0.910	1.028
	(0.134)	(0.284)	(0.139)	(0.149)
Partner: non-native	1.177	4.647***	1.892	4.256***
	(0.354)	(0.561)	(0.418)	(0.411)
Children	0.882	0.724	0.335***	0.231***
	(0.364)	(0.735)	(0.357)	(0.432)
Willingness to take risks	1.130***	0.931	1.098***	1.093***
	(0.031)	(0.060)	(0.031)	(0.033)
Patience	1.002	0.907*	1.037	1.068**
	(0.026)	(0.053)	(0.026)	(0.027)
Number of observations	1560	1560	1682	1682
Country FE	✓	✓	✓	✓
Correctly predicted values	69.2	69.2	46.2	46.2
McFadden Pseudo R2	0.06	0.06	0.05	0.05

Note: Risk ratios, standard errors in parentheses. See further notes in footnote [79].
*p<0.1; **p<0.05; ***p<0.01

In summary, individual-level characteristics help explain the differences between migration aspirations and intentions, which we observe in Table 6.3. Overall, our results suggest that for women, the main driver of these differences are family ties. For them, having a non-native partner increases the likelihood of intentions to migrate permanently despite the absence of aspirations. Having children, on the other hand, impedes existing aspirations to migrate from being translated into intentions to move, temporarily or permanently. Though these results are only significant within the student sample, they are in line with former literature on women being predominantly influenced in their migration decisions by social ties (De Jong, 2000; Mincer, 1978; Nakosteen and Zimmer, 1980; Uebelmesser, 2006). For men, we find education and career to be the most influential in their migration-decision and to explain differences between aspirations and intentions. Among men in both samples, holding a university degree is positively associated with intentions to migrate temporarily, even despite the absence of migration aspirations. In the GI sample, where we can differentiate between students and non-students, men who are studying are also less likely to intend a permanent move abroad, despite aspirations to do so. These findings, too, go hand in hand with existing literature describing professional reasons as the main drivers of migration decisions for men (Geist and McManus, 2011; Mckinnish, 2008; Munk et al., 2017; Nakosteen and Zimmer, 1980).

6.6   Conclusion

We combine different survey questions on migration aspirations and migration intentions, and separately consider intentions to migrate temporarily and permanently, to identify potential future migrants which have not received much attention in the literature. For this, we conduct two multinational surveys – one among language course participants in 14 countries and one among university students in six countries. We use multinomial probit estimations to analyse individual-level determinants of migration aspirations and intentions. Our results contribute to the literature in three ways:

First, though comparing aspirations and intentions to migrate within individuals shows that the former is an important determinant of the latter, we find that after controlling for aspirations, individual-level characteristics remain important in explaining intentions.

Second, as we introduce a clear time horizon to our measure of intentions, we are able to differentiate between intentions to migrate temporarily and permanently. This allows us to shed light on the substantial group of respondents who intend to migrate temporarily and to highlight differences between this group and the group of those with intentions to migrate permanently. Our results suggest that a temporary move abroad is considered or planned primarily for educational and career reasons, while intentions for a permanent move abroad are mostly explained by family ties. We also find that the relevance of determinants differs by gender. The positive linkage between educational status and intentions to migrate temporarily is robust only for men across both samples. The linkage between family ties and intentions to migrate permanently is driven largely by women in both samples. A short exploration of motivations for potential migration sheds some light on these gender differences. The share of respondents who state educational or professional reasons as their main motivation for migration intentions is consistently larger among men, and the share of those who state family and partner as their main motivation is consistently larger among women in both samples.

Third, our framework allows for aspirations and intentions to migrate permanently to differ. This sheds light on respondents who do not aspire to migrate permanently but intend to do so and those who aspire to migrate permanently but do not intend this. The former is another group of potential migrants which has been largely neglected so far, as they are not included in the GWP data due to the conditionality in responses. Exploiting these differences, our results suggest that strong family ties at the origin, i.e. having children, can impede a desire to migrate permanently from being translated into intentions. Family ties abroad, i.e. having a non-native partner, on the other hand, can induce intentions to migrate permanently despite the absence of aspirations to do so. Having a university degree can induce intentions to migrate temporarily also among those without aspirations. Again, we find that family ties are especially important for women with regard to their migration choice and can explain differences in their aspirations and intentions, while education and career reasons do so for men.

Our analysis shows that migration choices are made at the intersection of desires under ideal circumstances and considerations and plans given real circumstances. While pure desires to leave one's country permanently might originate in pull factors abroad, push factors at the country of origin may result in considerations and plans to emigrate even if one would ideally like to stay. A better understanding of why aspirations and intentions sometimes differ could

210

reveal such push and pull factors and thus be of great value for the design of targeted policy interventions. However, closer evaluation of these differences requires large-scale observations of migration aspirations and intentions prior to actual migration. To date, the GWP is the only globally representative survey available on migration aspirations and intentions, yet it suffers from the conditionality in responses and lacks a clear differentiation of potential future migration into temporary and permanent moves. Resolving this would allow observation of all migration patterns which have received little attention in the literature on potential future migrants so far and which we have shown in our two samples to play an essential role for migration research. Our analysis provides a first step in this direction.

Appendices to Chapter 6

6.7 Appendix A

Figure 6.A1: Share of respondents with migration intentions. GI and student sample.
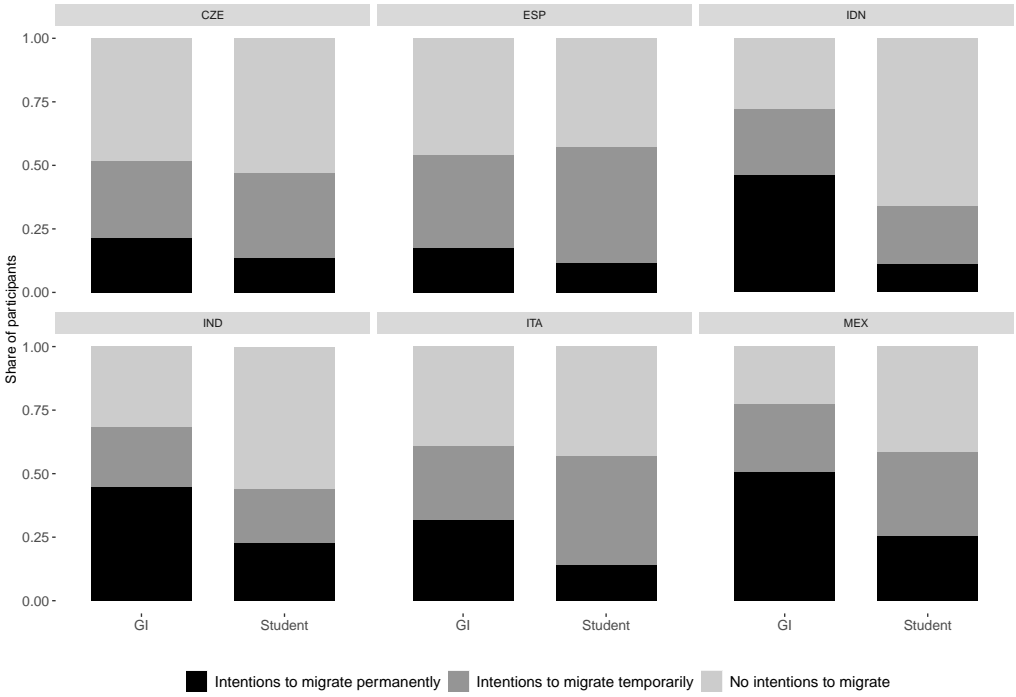
Note: Compares the shares of migration intentions between the GI and the student samples (including all age groups), for the 6 countries which are observed in both samples.

Table 6.A1: Variables description.

Variable	Type	Description
Migration aspirations	Binary	Indicates respondent's aspirations to migrate permanently: *"Ideally, if you had the opportunity, would you like to move permanently to another country or would you prefer to continue living in [country in which the survey took place]?"* - *"Like to permanently move to another country"* (migration aspirations) - *"Like to stay in [country in which the survey took place]."* (no migration aspirations) Reference category is 'no migration aspirations'.
Migration intentions	Categorical	Indicates respondent's intentions to migrate: *"Tick the statement that applies to your current situation"* - *"I would not move to another country under any circumstances"* (no intentions to migrate) - *"In principle, I would move to another country, but I have not thought about it in the last 12 months"* (no intentions to migrate) - *"I have been thinking about moving to another country in the last 12 months, but have no specific plans."* (intentions to migrate) - *"I am planning a move to another country."* (intentions to migrate) - *"I already have a date for my planned move to another country."* (intentions to migrate) Those respondents with migration intentions are further asked for their preferred destination country and their preferred length of stay (*"How long would you most likely stay in [preferred destination country]"?"*) and likelihood of return (*"How likely is it that you will return to [country in which the survey took place] after a temporary stay in [preferred destination country]?"*). Those who state that they would most likely stay in their preferred destination country for more than 5 years or state that their return to [country in which the survey took place] after a temporary stay in their preferred destination country is unlikely are classified as having permanent migration intentions; the rest is classified as having temporary migration intentions. Those who state no migration intentions are not asked those questions. Reference category is 'no migration intentions'.
Gender: woman	Binary	Indicates respondent's gender. Takes a value of 1 if respondent indicated to be a woman, and a value of 0 if respondent indicated to be a man, or *"No answer/prefer not to say"*, or if response is missing.
Gender: n/a	Binary	Indicates respondent's gender. Takes a value of 1 if respondent indicated *"No answer/prefer not to say"*, or if response is missing, and a value of 0 if respondent indicated to be either a man or a women. This dummy is not reported in the result tables.
Age	Numerical (1-6)	Indicates respondent's age group according to the ranges: *under 18, 18 to 24, 25 to 34, 35 to 49, 50 to 64, 65 and older.* Reference category is 18 to 24 years for whole samples and samples restricted to younger age groups and 35 to 49 years for samples restricted to older age groups.
University degree	Binary	Indicates whether respondent has a university degree or not. Reference category is 'no university degree'.
Student	Binary	Indicates whether respondent is a student or not. Reference category is 'no student'.
Migration experience	Binary	Indicates whether respondent has stayed abroad for at least three consecutive months in the past. Reference category is 'no migration experience'.

Table 6.A1: Variables description (continued).

Variable	Type	Description
Partner: native	Binary	Indicates whether respondent is in a long-term relationship with or married to a partner whose native language is consistent with the one(s) spoken in the country in which the survey took place. Reference category is 'no native partner'.
Partner: non-native	Binary	Indicates whether respondent is in a long-term relationship with or married to a partner whose native language is different from the one(s) spoken in the country in which the survey took place. Reference category is 'no non-native partner'.
Children	Binary	Indicates whether respondent has any children. Reference category is 'no children'.
Willingness to take risks	Numerical (0-10 / 1-10)	Measures respondents' willingness to take risks (*"Would you describe yourself as someone who tries to avoid risks (risk-averse) or as someone who is willing to take risks (risk-prone)?"*) on a 11-point scale from 0 for *"risk-averse"* to 10 for *"risk-prone"* in the GI sample; and on a 10-point scale from 1 for *"risk-averse"* to 10 for *"risk-prone"* in the student sample.
Patience	Numerical (0-10 / 1-10)	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 *"very impatient"* to 10 for *"very patient"* in the GI sample; and on a 10-point scale from 1 for *"very impatient"* to 10 for *"very patient"* in the student sample.
Main motivation for potential migration	Categorical	Indicates respondents' main reason for a potential move to their preferred destination country. Educational reasons include study/education/PhD. Professional reasons include work experience/(unpaid) traineeship, own higher income, more interesting job, poor job prospects in origin country, transfer by employer and other own professional reasons. Family and partner include professional reasons/studies of partner, partner lives in the destination country, other family/partner related reasons, friends/relatives live in destination country (South Korea only). Other reasons include interest in the country and culture, adventure, environmental reasons, higher quality of life, and all other reasons. Those who state no migration intentions are not asked this question.

Table 6.A2: Descriptive statistics.

Panel A: GI sample

	no intentions to migrate	intentions to migrate temporarily	intentions to migrate permanently	*Total*
Gender: woman	58.6	59.5	55.6	57.6
Gender: man	38.9	37.6	41.3	39.5
Age: under 18 years	11.0	11.2	17.4	13.7
Age: 18 to 24 years	36.1	43.9	37.8	38.7
Age: 25 to 34 years	19.6	29.7	29.0	26.0
Age: 35 to 49 years	17.2	12.3	12.8	14.2
Age: 50 to 64 years	10.2	2.3	2.6	5.1
Age: 65 years or over	6.0	0.6	0.4	2.4
University degree	62.5	63.5	56.4	60.2
Student	31.1	36.3	31.9	32.7
Migration experience	33.4	36.8	31.2	33.3
Partner: native	33.8	25.6	24.9	28.1
Partner: non-native	6.6	7.5	9.3	8.0
Children	19.9	8.1	11.2	13.4
Willingness to take risks	6.0	6.5	6.7	6.4
(0=risk averse ... 10=risk prone)				
Patience	6.2	6.1	6.3	6.2
(0=very impatient ... 10=very patient)				
Number of observations	1729	1256	2125	5110

Panel B: Student sample

Gender: woman	52.8	56.9	53.3	54.3
Gender: man	46.9	42.6	46.3	45.3
Age: under 18 years	0.3	0.2	0.2	0.2
Age: 18 to 24 years	77.4	75.9	76.7	76.8
Age: 25 to 34 years	16.0	20.0	19.0	17.9
Age: 35 to 49 years	5.0	3.5	3.6	4.3
Age: 50 to 64 years	1.3	0.4	0.5	0.8
Age: 65 years or over	0.1	0.0	0.0	0.0
University degree	45.6	54.1	50.4	49.4
Migration experience	26.2	43.5	40.1	34.6
Partner: native	39.0	35.3	32.8	36.6
Partner: non-native	2.2	4.2	7.4	3.8
Children	6.6	3.5	2.9	4.9
Willingness to take risks	5.7	6.2	6.2	6.0
(1=risk averse ... 10=risk prone)				
Patience	5.9	5.8	5.9	5.9
(1=very impatient ... 10=very patient)				
Number of observations	1554	1105	583	3242

Note: Shares of observations; except for willingness to take risks and patience which show means.

Table 6.A3: Joint distribution of aspirations and intentions, by gender.

Panel A: GI sample, women

	no intentions to migrate		intentions to migrate temporarily		intentions to migrate permanently		*Total*	
no aspirations to migrate permanently	651	(22.1)	254	(9.0)	75	(2.6)	990	(33.7)
aspirations to migrate permanently	363	(12.3)	483	(16.4)	1107	(37.6)	1953	(66.3)
Total	1014	(34.4)	747	(25.4)	1182	(40.2)	2943	(100.0)

Panel A: GI sample, men

	no intentions to migrate		intentions to migrate temporarily		intentions to migrate permanently		*Total*	
no aspirations to migrate permanently	447	(22.1)	187	(9.3)	71	(3.5)	705	(34.9)
aspirations to migrate permanently	225	(11.1)	285	(14.1)	806	(39.9)	1316	(65.1)
Total	672	(33.2)	472	(23.4)	877	(43.4)	2021	(100.0)

Panel B: Student sample, women

	no intentions to migrate		intentions to migrate temporarily		intentions to migrate permanently		*Total*	
no aspirations to migrate permanently	559	(31.7)	217	(12.3)	40	(2.3)	816	(46.3)
aspirations to migrate permanently	262	(14.9)	412	(23.4)	271	(15.4)	945	(53.7)
Total	821	(46.6)	629	(35.7)	311	(17.7)	1761	(100.0)

Panel B: Student sample, men

	no intentions to migrate		intentions to migrate temporarily		intentions to migrate permanently		*Total*	
no aspirations to migrate permanently	509	(34.6)	197	(13.4)	34	(2.3)	740	(50.3)
aspirations to migrate permanently	220	(15.0)	274	(18.6)	236	(16.1)	730	(49.7)
Total	729	(49.6)	471	(32.0)	270	(18.4)	1470	(100.0)

Note: Numbers of observation, percentages in parentheses. Row and column *Total(s)* show row and column totals; percentages of total sample size in parentheses.

Table 6.A4: Aspirations and intentions. High-income countries.

Panel A: GI sample

	binomial probit	multinomial probit		multinomial probit	
	(1) aspirations to migrate permanently	(2) intentions to migrate temporarily	(3) intentions to migrate permanently	(4) intentions to migrate temporarily	(5) intentions to migrate permanently
	(n = 1296)	(n = 648)	(n = 730)	(n = 648)	(n = 730)
Migration aspiration				3.538***	20.390***
				(0.121)	(0.148)
Gender: woman	1.001	0.989	0.692***	0.935	0.616***
	(0.056)	(0.113)	(0.110)	(0.117)	(0.126)
Age: under 18 years	1.231	0.497**	0.878	0.419***	0.651
	(0.149)	(0.295)	(0.270)	(0.307)	(0.304)
Age: 25 to 34 years	0.959	0.931	1.323	0.980	1.520**
	(0.095)	(0.185)	(0.183)	(0.192)	(0.208)
Age: 35 to 49 years	0.746***	0.422***	0.638**	0.467***	0.805
	(0.111)	(0.216)	(0.219)	(0.224)	(0.250)
Age: 50 to 64 years	0.472***	0.115***	0.205***	0.141***	0.312***
	(0.140)	(0.300)	(0.295)	(0.308)	(0.339)
University degree	0.905	1.186	0.906	1.251	0.962
	(0.081)	(0.156)	(0.159)	(0.162)	(0.179)
Student	0.979	0.745*	0.841	0.740*	0.817
	(0.090)	(0.175)	(0.171)	(0.183)	(0.194)
Migration experience	1.011	1.139	1.150	1.162	1.187
	(0.058)	(0.114)	(0.113)	(0.118)	(0.129)
Partner: native	1.002	0.908	0.991	0.899	0.989
	(0.069)	(0.135)	(0.138)	(0.139)	(0.159)
Partner: non-native	0.998	1.237	1.806***	1.341	2.248***
	(0.108)	(0.214)	(0.215)	(0.223)	(0.246)
Children	0.958	0.865	1.020	0.860	1.023
	(0.091)	(0.190)	(0.182)	(0.195)	(0.211)
Willingness to take risks	1.107***	1.149***	1.200***	1.116***	1.131***
	(0.013)	(0.027)	(0.026)	(0.028)	(0.029)
Patience	0.991	0.956**	0.977	0.956*	0.978
	(0.012)	(0.023)	(0.023)	(0.024)	(0.026)
Number of observations	2433	2433	2433	2433	2433
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	67.4	53	53	61.7	61.7
McFadden Pseudo R2	0.12	0.11	0.11	0.22	0.22

Panel B: Student sample

	(n = 674)	(n = 626)	(n = 204)	(n = 626)	(n = 204)
Migration aspiration				4.151***	15.348***
				(0.127)	(0.206)
Gender: woman	1.126*	1.081	1.112	1.018	1.012
	(0.069)	(0.120)	(0.174)	(0.126)	(0.188)
Age: 25 to 34 years	0.940	1.009	0.842	1.030	0.883
	(0.105)	(0.180)	(0.259)	(0.187)	(0.282)
Age: 35 to 49 years	1.024	0.194***	0.719	0.172***	0.662
	(0.243)	(0.550)	(0.616)	(0.570)	(0.683)
University degree	0.838**	1.220	1.013	1.386**	1.251
	(0.071)	(0.123)	(0.179)	(0.130)	(0.194)
Migration experience	1.378***	2.019***	2.386***	1.856***	2.027***
	(0.070)	(0.121)	(0.174)	(0.127)	(0.189)
Partner: native	0.751***	0.998	0.935	1.165	1.247
	(0.070)	(0.121)	(0.178)	(0.128)	(0.193)
Partner: non-native	1.309*	1.624	3.704***	1.602	3.594***
	(0.152)	(0.298)	(0.329)	(0.316)	(0.367)
Children	0.949	0.755	0.323*	0.734	0.237*
	(0.229)	(0.419)	(0.669)	(0.436)	(0.755)
Willingness to take risks	1.091***	1.175***	1.111**	1.136***	1.043
	(0.017)	(0.029)	(0.042)	(0.031)	(0.045)
Patience	1.007	1.008	1.053	1.011	1.048
	(0.014)	(0.025)	(0.036)	(0.026)	(0.038)
Number of observations	1594	1594	1594	1594	1594
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	63.9	54.5	54.5	60.5	60.5
McFadden Pseudo R2	0.06	0.06	0.06	0.14	0.14

Note: Risk ratios, standard errors in parentheses. Specification (1) estimates the probability of aspirations to migrate permanently via binomial probit; reference category is 'no aspirations to migrate permanently' (n = 1137 (GI sample), n = 920 (student sample)). Specifications (2) and (4) estimate the probability of intentions to migrate temporarily, and specifications (3) and (5) the probability of intentions to migrate permanently via multinomial probit; reference category is 'no intentions to migrate' (n = 1055 (GI sample), n = 764 (student sample)). High-income countries include countries which have a GNI per capita larger than \$12,535 in current US-Dollars, as of 2020 (Czechia, Great Britain, Italy, Japan, Netherlands, Poland, Romania, South Korea, and Spain); middle-income countries are countries which have a GNI per capita of \$1,036 to \$12,535 in current US-Dollars, as of 2020 (Bosnia, India, Indonesia, Mexico, and Ukraine). All EU member states in the sample are high-income countries, and all middle-income countries are not EU member states.
*p<0.1; **p<0.05; ***p<0.01

Table 6.A5: Aspirations and intentions. Middle-income countries.

Panel A: GI sample

	binomial probit	multinomial probit		multinomial probit	
	(1) aspirations to migrate permanently	(2) intentions to migrate temporarily	(3) intentions to migrate permanently	(4) intentions to migrate temporarily	(5) intentions to migrate permanently
	(n = 2071)	(n = 608)	(n = 1395)	(n = 608)	(n = 1395)
Migration aspiration				2.513***	26.610***
				(0.123)	(0.159)
Gender: woman	1.115*	1.143	1.116	1.109	1.025
	(0.058)	(0.119)	(0.101)	(0.122)	(0.114)
Age: under 18 years	1.277**	1.077	1.570***	1.042	1.415*
	(0.099)	(0.205)	(0.174)	(0.210)	(0.198)
Age: 25 to 34 years	1.146	1.070	1.398**	1.041	1.326
	(0.096)	(0.197)	(0.171)	(0.205)	(0.194)
Age: 35 to 49 years	1.060	0.647	0.893	0.625	0.803
	(0.142)	(0.307)	(0.245)	(0.311)	(0.280)
Age: 50 to 64 years	0.994	0.158**	0.790	0.158**	0.740
	(0.256)	(0.800)	(0.404)	(0.806)	(0.486)
University degree	0.807***	1.469**	0.930	1.627***	1.145
	(0.075)	(0.154)	(0.131)	(0.159)	(0.151)
Student	0.919	1.306	0.836	1.344*	0.865
	(0.081)	(0.166)	(0.143)	(0.172)	(0.163)
Migration experience	1.104	1.428**	1.491***	1.456**	1.510***
	(0.077)	(0.162)	(0.138)	(0.167)	(0.156)
Partner: native	1.075	0.980	1.200	0.963	1.189
	(0.088)	(0.184)	(0.155)	(0.190)	(0.175)
Partner: non-native	1.232	0.783	1.590*	0.749	1.486
	(0.132)	(0.300)	(0.237)	(0.309)	(0.271)
Children	0.764**	0.708	0.839	0.782	1.130
	(0.134)	(0.312)	(0.234)	(0.314)	(0.271)
Willingness to take risks	1.019	1.026	1.076***	1.028	1.077***
	(0.015)	(0.030)	(0.025)	(0.030)	(0.028)
Patience	0.996	1.016	0.991	1.018	0.993
	(0.012)	(0.026)	(0.022)	(0.026)	(0.024)
Number of observations	2677	2677	2677	2677	2677
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	77.3	52	52	62.9	62.9
McFadden Pseudo R2	0.04	0.04	0.04	0.16	0.16

Panel B: Student sample

	(n = 1008)	(n = 479)	(n = 379)	(n = 479)	(n = 379)
Migration aspiration				3.541***	15.622***
				(0.133)	(0.201)
Gender: woman	1.195***	1.012	1.125	0.929	0.972
	(0.069)	(0.126)	(0.136)	(0.131)	(0.148)
Age: 25 to 34 years	0.937	1.461**	1.066	1.537**	1.165
	(0.098)	(0.175)	(0.192)	(0.183)	(0.208)
Age: 35 to 49 years	0.774	1.358	0.812	1.525	1.006
	(0.172)	(0.305)	(0.364)	(0.316)	(0.396)
University degree	1.051	1.338**	1.251	1.351**	1.311*
	(0.078)	(0.144)	(0.151)	(0.150)	(0.165)
Migration experience	1.083	1.389**	1.411**	1.388**	1.382**
	(0.078)	(0.140)	(0.151)	(0.146)	(0.164)
Partner: native	0.991	0.737**	1.046	0.737**	1.034
	(0.078)	(0.143)	(0.152)	(0.148)	(0.165)
Partner: non-native	1.048	1.112	3.050***	1.224	3.693***
	(0.214)	(0.434)	(0.399)	(0.448)	(0.443)
Children	1.085	0.618	0.503*	0.553*	0.418**
	(0.168)	(0.303)	(0.376)	(0.314)	(0.410)
Willingness to take risks	1.003	1.063**	1.065**	1.071**	1.079**
	(0.016)	(0.030)	(0.032)	(0.031)	(0.035)
Patience	0.987	1.007	1.006	1.015	1.017
	(0.014)	(0.025)	(0.027)	(0.026)	(0.029)
Number of observations	1648	1648	1648	1648	1648
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	67.1	50.2	50.2	54.3	54.3
McFadden Pseudo R2	0.07	0.04	0.04	0.12	0.12

Note: Risk ratios, standard errors in parentheses. Specification (1) estimates the probability of aspirations to migrate permanently via binomial probit; reference category is 'no aspirations to migrate permanently' (n = 606 (GI sample), n = 640 (student sample)). Specifications (2) and (4) estimate the probability of intentions to migrate temporarily, and specifications (3) and (5) estimate the probability of intentions to migrate permanently via multinomial probit; reference category is 'no intentions to migrate' (n = 674 (GI sample), n = 790 (student sample)). High-income countries include countries which have a GNI per capita larger than $12,535 in current US-Dollars, as of 2020 (Czechia, Great Britain, Italy, Japan, Netherlands, Poland, Romania, South Korea, and Spain); middle-income countries are countries which have a GNI per capita of $1,036 to $12,535 in current US-Dollars, as of 2020 (Bosnia, India, Indonesia, Mexico, and Ukraine). All EU member states in the sample are high-income countries, and all middle-income countries are not EU member states.
*p<0.1; **p<0.05; ***p<0.01

Table 6.A6: Aspirations and intentions. Women.

Panel A: GI sample

	binomial probit	multinomial probit		multinomial probit	
	(1) aspirations to migrate permanently	(2) intentions to migrate temporarily	(3) intentions to migrate permanently	(4) intentions to migrate temporarily	(5) intentions to migrate permanently
	(n = 1953)	(n = 747)	(n = 1182)	(n = 747)	(n = 1182)
Migration aspiration				3.054***	23.125***
				(0.113)	(0.147)
Age: under 18 years	1.387***	0.822	1.764***	0.742	1.464*
	(0.117)	(0.232)	(0.204)	(0.240)	(0.227)
Age: 25 to 34 years	1.093	0.992	1.512**	0.996	1.538**
	(0.089)	(0.172)	(0.163)	(0.180)	(0.184)
Age: 35 to 49 years	0.804**	0.567***	0.755	0.610**	0.876
	(0.110)	(0.214)	(0.206)	(0.221)	(0.235)
Age: 50 to 64 years	0.540***	0.110***	0.333***	0.128***	0.440**
	(0.155)	(0.388)	(0.302)	(0.396)	(0.353)
Age: 65 years or above	0.375***	0.111***	0.107***	0.137***	0.190**
	(0.229)	(0.520)	(0.643)	(0.528)	(0.696)
University degree	0.838**	1.229	0.896	1.335**	1.027
	(0.072)	(0.141)	(0.133)	(0.146)	(0.150)
Student	0.975	1.137	0.965	1.157	0.967
	(0.083)	(0.158)	(0.149)	(0.165)	(0.167)
Migration experience	1.050	1.475***	1.474***	1.510***	1.546***
	(0.061)	(0.120)	(0.114)	(0.124)	(0.129)
Partner: native	0.960	0.936	0.912	0.941	0.931
	(0.071)	(0.137)	(0.133)	(0.141)	(0.151)
Partner: non-native	1.116	1.081	1.867***	1.136	2.050***
	(0.106)	(0.218)	(0.202)	(0.227)	(0.233)
Children	0.990	0.854	1.110	0.863	1.161
	(0.099)	(0.208)	(0.187)	(0.212)	(0.218)
Willingness to take risks	1.083***	1.080***	1.150***	1.057**	1.102***
	(0.013)	(0.025)	(0.024)	(0.026)	(0.027)
Patience	1.001	0.991	0.991	0.988	0.988
	(0.011)	(0.022)	(0.020)	(0.023)	(0.023)
Number of observations	2943	2943	2943	2943	2943
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	73.3	52.2	52.2	61.4	61.4
McFadden Pseudo R2	0.15	0.10	0.10	0.20	0.20

Panel B: Student sample

	(n = 945)	(n = 629)	(n = 311)	(n = 629)	(n = 311)
Migration aspiration				4.609***	14.187***
				(0.126)	(0.198)
Age: 25 to 34 years	0.915	1.260	0.946	1.357*	1.053
	(0.095)	(0.167)	(0.209)	(0.176)	(0.225)
Age: 35 to 49 years	0.767	0.591	0.502	0.643	0.606
	(0.188)	(0.346)	(0.432)	(0.357)	(0.458)
University degree	0.876*	1.155	1.148	1.300**	1.349*
	(0.070)	(0.123)	(0.156)	(0.130)	(0.168)
Migration experience	1.311***	2.063***	2.060***	1.927***	1.868***
	(0.069)	(0.120)	(0.151)	(0.126)	(0.163)
Partner: native	0.791***	0.959	0.800	1.086	0.958
	(0.068)	(0.119)	(0.155)	(0.126)	(0.166)
Partner: non-native	1.131	1.834**	4.063***	2.100**	4.911***
	(0.153)	(0.308)	(0.318)	(0.328)	(0.354)
Children	1.145	0.513**	0.591	0.429**	0.452*
	(0.180)	(0.332)	(0.411)	(0.349)	(0.449)
Willingness to take risks	1.068***	1.111***	1.115***	1.085***	1.071*
	(0.016)	(0.029)	(0.036)	(0.030)	(0.038)
Patience	0.984	1.023	0.989	1.036	1.006
	(0.013)	(0.023)	(0.029)	(0.024)	(0.031)
Number of observations	1761	1761	1761	1761	1761
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	66.8	50.9	50.9	57.2	57.2
McFadden Pseudo R2	0.10	0.06	0.06	0.14	0.14

Note: Risk ratios, standard errors in parentheses. Specification (1) estimates the probability of aspirations to migrate permanently via binomial probit; reference category is 'no aspirations to migrate permanently' (n = 990 (GI sample), n = 816 (student sample)). Specifications (2) and (4) estimate the probability of intentions to migrate temporarily, and specifications (3) and (5) estimate the probability of intentions to migrate permanently via multinomial probit; reference category is 'no intentions to migrate' (n = 1014 (GI sample), n = 821 (student sample)). Respondents who gave no answer regarding their gender or indicated *"No answer/prefer not to say"* are excluded from the gender subsamples.
*p<0.1; **p<0.05; ***p<0.01

Table 6.A7: Aspirations and intentions. Men.

Panel A: GI sample

	binomial probit	multinomial probit		multinomial probit	
	(1) aspirations to migrate permanently	(2) intentions to migrate temporarily	(3) intentions to migrate permanently	(4) intentions to migrate temporarily	(5) intentions to migrate permanently
	(n = 1316)	(n = 472)	(n = 877)	(n = 472)	(n = 877)
Migration aspiration				2.899***	23.192***
				(0.140)	(0.164)
Age: under 18 years	1.153	1.055	1.025	0.986	0.885
	(0.122)	(0.252)	(0.217)	(0.259)	(0.251)
Age: 25 to 34 years	1.038	1.229	1.317	1.272	1.387
	(0.104)	(0.217)	(0.193)	(0.224)	(0.222)
Age: 35 to 49 years	0.956	0.427***	0.697	0.433***	0.668
	(0.135)	(0.288)	(0.250)	(0.296)	(0.291)
Age: 50 to 64 years	0.564***	0.155***	0.248***	0.177***	0.311***
	(0.180)	(0.399)	(0.349)	(0.407)	(0.413)
Age: 65 years or above	0.359***	0.041***	0.071***	0.051***	0.102***
	(0.244)	(0.691)	(0.603)	(0.698)	(0.688)
University degree	0.823**	1.452**	0.912	1.560**	1.110
	(0.089)	(0.186)	(0.163)	(0.192)	(0.187)
Student	0.932	0.914	0.672**	0.901	0.670**
	(0.092)	(0.191)	(0.168)	(0.199)	(0.193)
Migration experience	1.073	1.036	1.138	1.050	1.124
	(0.072)	(0.151)	(0.135)	(0.157)	(0.156)
Partner: native	1.149	0.971	1.455**	0.940	1.406*
	(0.087)	(0.184)	(0.164)	(0.190)	(0.189)
Partner: non-native	1.044	1.006	1.570*	1.018	1.732*
	(0.140)	(0.304)	(0.271)	(0.314)	(0.315)
Children	0.804*	0.771	0.856	0.822	1.072
	(0.119)	(0.271)	(0.225)	(0.276)	(0.265)
Willingness to take risks	1.039**	1.110***	1.114***	1.102***	1.101***
	(0.015)	(0.033)	(0.028)	(0.034)	(0.032)
Patience	0.981	0.960	0.951**	0.968	0.960
	(0.013)	(0.028)	(0.025)	(0.029)	(0.029)
Number of observations	2021	2021	2021	2021	2021
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	71.6	53.1	53.1	64.5	64.5
McFadden Pseudo R2	0.11	0.10	0.10	0.22	0.22

Panel B: Student sample

	(n = 730)	(n = 471)	(n = 270)	(n = 471)	(n = 270)
Migration aspiration				3.139***	16.427***
				(0.134)	(0.209)
Age: 25 to 34 years	1.018	1.106	0.999	1.088	0.967
	(0.104)	(0.185)	(0.224)	(0.190)	(0.245)
Age: 35 to 49 years	1.010	0.930	1.083	0.913	1.069
	(0.208)	(0.376)	(0.454)	(0.385)	(0.502)
University degree	1.004	1.558***	1.197	1.610***	1.291
	(0.078)	(0.142)	(0.170)	(0.146)	(0.188)
Migration experience	1.178**	1.331**	1.466**	1.278*	1.297
	(0.079)	(0.140)	(0.171)	(0.145)	(0.188)
Partner: native	0.931	0.770*	1.362*	0.811	1.462**
	(0.080)	(0.147)	(0.174)	(0.151)	(0.192)
Partner: non-native	1.471*	0.857	2.565**	0.752	1.995
	(0.220)	(0.416)	(0.435)	(0.426)	(0.475)
Children	0.942	0.823	0.289**	0.789	0.260**
	(0.204)	(0.366)	(0.552)	(0.375)	(0.600)
Willingness to take risks	1.018	1.141***	1.069*	1.139***	1.082**
	(0.016)	(0.031)	(0.036)	(0.032)	(0.040)
Patience	1.006	0.981	1.062*	0.982	1.062*
	(0.015)	(0.027)	(0.032)	(0.028)	(0.035)
Number of observations	1470	1470	1470	1470	1470
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	63.7	53.8	53.8	59.3	59.3
McFadden Pseudo R2	0.06	0.05	0.05	0.14	0.14

Note: Risk ratios, standard errors in parentheses. Specification (1) estimates the probability of aspirations to migrate permanently via binomial probit; reference category is 'no aspirations to migrate permanently' (n = 705 (GI sample), n = 740 (student sample)). Specifications (2) and (4) estimate the probability of intentions to migrate temporarily, and specifications (3) and (5) estimate the probability of intentions to migrate permanently via multinomial probit; reference category is 'no intentions to migrate' (n = 672 (GI sample), n = 729 (student sample)). Respondents who gave no answer regarding their gender or indicated *"No answer/prefer not to say"* are excluded from the gender subsamples.
*p<0.1; **p<0.05; ***p<0.01

Table 6.A8: Aspirations and intentions. GI sample, 6 countries. Younger age groups (under 35 years of age) and older age groups (35 years and above).

GI sample, younger age groups (under 35 years of age)

	binomial probit	multinomial probit		multinomial probit	
	(1) aspirations to migrate permanently	(2) intentions to migrate temporarily	(3) intentions to migrate permanently	(4) intentions to migrate temporarily	(5) intentions to migrate permanently
	(n = 1680)	(n = 717)	(n = 977)	(n = 717)	(n = 977)
Migration aspiration				2.366***	20.462***
				(0.116)	(0.159)
Gender: woman	1.095	1.056	0.999	1.024	0.917
	(0.058)	(0.112)	(0.105)	(0.114)	(0.117)
Age: under 18 years	1.270**	0.780	1.208	0.728	1.006
	(0.103)	(0.204)	(0.184)	(0.208)	(0.207)
Age: 25 to 34 years	1.081	1.092	1.350*	1.093	1.320
	(0.091)	(0.173)	(0.169)	(0.177)	(0.188)
University degree	0.882*	1.604***	1.011	1.687***	1.135
	(0.075)	(0.143)	(0.137)	(0.146)	(0.153)
Student	0.999	1.163	0.896	1.161	0.879
	(0.081)	(0.154)	(0.149)	(0.159)	(0.165)
Migration experience	1.105	1.577***	1.375**	1.547***	1.332*
	(0.075)	(0.140)	(0.141)	(0.143)	(0.156)
Partner: native	1.012	1.121	1.245	1.116	1.242
	(0.086)	(0.162)	(0.164)	(0.166)	(0.182)
Partner: non-native	1.054	1.037	1.796**	1.098	2.028**
	(0.132)	(0.271)	(0.262)	(0.282)	(0.301)
Children	0.717*	0.565	0.644	0.620	0.933
	(0.189)	(0.374)	(0.348)	(0.379)	(0.403)
Willingness to take risks	1.059***	1.069**	1.113***	1.057*	1.085***
	(0.015)	(0.028)	(0.026)	(0.029)	(0.029)
Patience	0.982	0.987	0.972	0.992	0.982
	(0.012)	(0.024)	(0.022)	(0.024)	(0.025)
Number of observations	2443	2443	2443	2443	2443
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	71.4	46.8	46.8	56.8	56.8
McFadden Pseudo R2	0.09	0.04	0.04	0.14	0.14

GI sample, older age groups (35 years and above)

	(n = 193)	(n = 98)	(n = 127)	(n = 98)	(n = 127)
Migration aspiration				7.872***	19.646***
				(0.332)	(0.337)
Gender: woman	0.788*	1.061	0.616*	1.149	0.707
	(0.134)	(0.269)	(0.255)	(0.284)	(0.288)
Age: 50 to 64 years	0.838	0.387**	0.564*	0.363**	0.546
	(0.167)	(0.373)	(0.320)	(0.398)	(0.370)
Age: 65 years or above	0.488***	0.193***	0.196***	0.243**	0.281*
	(0.252)	(0.572)	(0.590)	(0.584)	(0.657)
University degree	0.785	1.408	0.779	1.576	1.023
	(0.203)	(0.436)	(0.387)	(0.469)	(0.462)
Migration experience	0.805	1.208	0.856	1.383	0.987
	(0.134)	(0.271)	(0.257)	(0.289)	(0.296)
Partner: native	1.161	0.980	0.941	0.856	0.827
	(0.175)	(0.333)	(0.330)	(0.355)	(0.375)
Partner: non-native	1.088	1.189	1.322	1.174	1.305
	(0.214)	(0.425)	(0.411)	(0.449)	(0.473)
Children	0.898	0.396***	0.738	0.388***	0.686
	(0.152)	(0.306)	(0.287)	(0.325)	(0.326)
Willingness to take risks	1.060*	0.949	1.273***	0.951	1.252***
	(0.034)	(0.068)	(0.071)	(0.072)	(0.078)
Patience	0.937**	0.858***	0.806***	0.860**	0.814***
	(0.029)	(0.058)	(0.055)	(0.061)	(0.063)
Number of observations	478	478	478	478	478
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	73.2	61.1	61.1	68.8	68.8
McFadden Pseudo R2	0.21	0.15	0.15	0.26	0.26

Note: GI sample restricted to younger age groups (under 35 years of age) or older age groups (35 years and above) and to the 6 countries which are also observed in the student sample. Risk ratios, standard errors in parentheses. Specification (1) estimates the probability of aspirations to migrate permanently via binomial probit; reference category is 'no aspirations to migrate permanently' ($n = 763$ (GI sample, younger age groups), $n = 285$ (GI sample, older age groups)). Specifications (2) and (4) estimate the probability of intentions to migrate temporarily, and specifications (3) and (5) estimate the probability of intentions to migrate permanently via multinomial probit; reference category is 'no intentions to migrate' ($n = 749$ (GI sample, younger age groups), $n = 253$ (GI sample, older age groups)).

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

221

Table 6.A9: Aspirations and intentions. GI sample, 14 countries. Younger age groups (under 35 years of age) and older age groups (35 years and above).

GI sample, younger age groups (under 35 years of age)

	binomial probit	multinomial probit		multinomial probit	
	(1) aspirations to migrate permanently	(2) intentions to migrate temporarily	(3) intentions to migrate permanently	(4) intentions to migrate temporarily	(5) intentions to migrate permanently
	(n = 2845)	(n = 1065)	(n = 1788)	(n = 1065)	(n = 1788)
Migration aspiration				2.628***	23.622***
				(0.094)	(0.125)
Gender: woman	1.106**	1.057	0.963	0.998	0.848*
	(0.046)	(0.091)	(0.082)	(0.093)	(0.094)
Age: under 18 years	1.259***	0.839	1.355**	0.788	1.180
	(0.083)	(0.167)	(0.146)	(0.171)	(0.166)
Age: 25 to 34 years	1.046	0.993	1.345**	1.011	1.404**
	(0.068)	(0.135)	(0.125)	(0.139)	(0.142)
University degree	0.856***	1.301**	0.903	1.375***	1.019
	(0.060)	(0.117)	(0.108)	(0.121)	(0.123)
Student	0.940	1.020	0.829*	1.035	0.844
	(0.061)	(0.121)	(0.110)	(0.125)	(0.125)
Migration experience	1.165***	1.326***	1.462***	1.309**	1.411***
	(0.055)	(0.108)	(0.101)	(0.111)	(0.114)
Partner: native	0.943	0.944	1.084	0.980	1.176
	(0.063)	(0.126)	(0.119)	(0.130)	(0.136)
Partner: non-native	1.059	1.155	1.667**	1.197	1.784**
	(0.103)	(0.215)	(0.202)	(0.223)	(0.230)
Children	0.805	0.736	0.768	0.778	0.969
	(0.136)	(0.293)	(0.254)	(0.298)	(0.298)
Willingness to take risks	1.060***	1.088***	1.121***	1.075***	1.093***
	(0.011)	(0.022)	(0.020)	(0.023)	(0.023)
Patience	0.993	0.993	0.994	0.996	0.999
	(0.010)	(0.019)	(0.017)	(0.019)	(0.020)
Number of observations	4005	4005	4005	4005	4005
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	73.1	50	50	59.9	59.9
McFadden Pseudo R2	0.08	0.06	0.06	0.17	0.17

GI sample, older age groups (35 years and above)

	(n = 522)	(n = 191)	(n = 337)	(n = 191)	(n = 337)
Migration aspiration				6.960***	26.717***
				(0.218)	(0.228)
Gender: woman	0.913	1.085	0.688**	1.103	0.682*
	(0.087)	(0.190)	(0.169)	(0.202)	(0.199)
Age: 50 to 64 years	0.685***	0.289***	0.481***	0.317***	0.532**
	(0.106)	(0.249)	(0.208)	(0.262)	(0.250)
Age: 65 years or above	0.424***	0.126***	0.149***	0.173***	0.236***
	(0.160)	(0.412)	(0.412)	(0.422)	(0.462)
University degree	0.737**	1.245	0.791	1.472	1.114
	(0.145)	(0.331)	(0.277)	(0.349)	(0.334)
Migration experience	0.849*	1.263	1.028	1.471*	1.268
	(0.086)	(0.188)	(0.169)	(0.201)	(0.200)
Partner: native	1.277**	0.796	1.027	0.617**	0.758
	(0.109)	(0.225)	(0.216)	(0.241)	(0.255)
Partner: non-native	1.209	0.720	1.889**	0.673	1.902*
	(0.148)	(0.325)	(0.286)	(0.340)	(0.335)
Children	0.882	0.829	0.950	0.903	1.015
	(0.096)	(0.209)	(0.185)	(0.223)	(0.218)
Willingness to take risks	1.092***	1.095**	1.189***	1.058	1.132***
	(0.020)	(0.045)	(0.040)	(0.048)	(0.047)
Patience	0.987	0.924*	0.909***	0.913**	0.897**
	(0.018)	(0.040)	(0.036)	(0.043)	(0.042)
Number of observations	1105	1105	1105	1105	1105
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	70.8	62.8	62.8	71.4	71.4
McFadden Pseudo R2	0.18	0.17	0.17	0.30	0.30

Note: GI sample restricted to younger age groups (under 35 years of age) or older age groups (35 years and above). Risk ratios, standard errors in parentheses. Specification (1) estimates the probability of aspirations to migrate permanently via binomial probit; reference category is 'no aspirations to migrate permanently' ($n = 1160$ (GI sample, younger age groups), $n = 583$ (GI sample, older age groups)). Specifications (2) and (4) estimate the probability of intentions to migrate temporarily, and specifications (3) and (5) estimate the probability of intentions to migrate permanently via multinomial probit; reference category is 'no intentions to migrate' ($n = 1152$ (GI sample, younger age groups), $n = 577$ (GI sample, older age groups)).
*p<0.1; **p<0.05; ***p<0.01

Table 6.A10: Aspirations and intentions. GI sample, 14 countries. Younger age groups (under 35 years of age), by gender.

GI sample, younger age groups (under 35 years of age), women

	binomial probit	multinomial probit		multinomial probit	
	(1) aspirations to migrate permanently	(2) intentions to migrate temporarily	(3) intentions to migrate permanently	(4) intentions to migrate temporarily	(5) intentions to migrate permanently
	(n = 1660)	(n = 643)	(n = 1003)	(n = 634)	(n = 1003)
Migration aspiration				2.637***	25.423***
				(0.124)	(0.175)
Age: under 18 years	1.389***	0.836	1.879***	0.771	1.597**
	(0.119)	(0.236)	(0.207)	(0.242)	(0.232)
Age: 25 to 34 years	1.074	0.944	1.414**	0.945	1.441*
	(0.091)	(0.176)	(0.167)	(0.183)	(0.188)
University degree	0.843**	1.199	0.909	1.284	1.031
	(0.079)	(0.151)	(0.143)	(0.157)	(0.162)
Student	0.957	1.153	0.975	1.173	0.989
	(0.084)	(0.161)	(0.150)	(0.166)	(0.168)
Migration experience	1.160**	1.651***	1.657***	1.632***	1.625***
	(0.074)	(0.140)	(0.135)	(0.144)	(0.151)
Partner: native	0.867*	0.936	0.862	0.981	0.965
	(0.083)	(0.159)	(0.156)	(0.164)	(0.177)
Partner: non-native	1.103	1.244	1.903**	1.314	2.070**
	(0.132)	(0.274)	(0.258)	(0.288)	(0.296)
Children	0.859	0.769	0.785	0.799	0.907
	(0.180)	(0.366)	(0.330)	(0.376)	(0.386)
Willingness to take risks	1.079***	1.076**	1.138***	1.059*	1.095***
	(0.015)	(0.029)	(0.027)	(0.030)	(0.030)
Patience	1.001	1.000	1.005	0.998	1.004
	(0.013)	(0.025)	(0.023)	(0.025)	(0.026)
Number of observations	2315	2315	2315	2315	2315
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	74.3	50	50	59	59
McFadden Pseudo R2	0.10	0.06	0.06	0.17	0.17

GI sample, younger age groups (under 35 years of age), men

	(n = 1103)	(n = 402)	(n = 731)	(n = 402)	(n = 731)
Migration aspiration				2.528***	22.848***
				(0.153)	(0.190)
Age: under 18 years	1.136	1.079	1.008	1.028	0.881
	(0.124)	(0.256)	(0.222)	(0.262)	(0.255)
Age: 25 to 34 years	1.029	1.198	1.258	1.238	1.307
	(0.106)	(0.221)	(0.198)	(0.227)	(0.227)
University degree	0.854*	1.482**	0.909	1.543**	1.030
	(0.095)	(0.197)	(0.175)	(0.202)	(0.200)
Student	0.925	0.926	0.673**	0.916	0.661**
	(0.094)	(0.193)	(0.170)	(0.201)	(0.195)
Migration experience	1.180*	1.027	1.318*	1.027	1.287
	(0.086)	(0.179)	(0.161)	(0.185)	(0.185)
Partner: native	1.072	0.981	1.581**	0.999	1.691**
	(0.103)	(0.220)	(0.197)	(0.226)	(0.226)
Partner: non-native	0.995	1.082	1.446	1.122	1.566
	(0.173)	(0.371)	(0.350)	(0.380)	(0.398)
Children	0.730	0.748	0.855	0.797	1.233
	(0.215)	(0.514)	(0.436)	(0.521)	(0.500)
Willingness to take risks	1.032*	1.101***	1.097***	1.097**	1.089**
	(0.018)	(0.037)	(0.032)	(0.038)	(0.036)
Patience	0.985	0.982	0.971	0.991	0.981
	(0.015)	(0.031)	(0.028)	(0.032)	(0.032)
Number of observations	1576	1576	1576	1576	1576
Country FE	✓	✓	✓	✓	✓
Correctly predicted values	71.6	50.7	50.7	62.6	62.6
McFadden Pseudo R2	0.07	0.07	0.07	0.18	0.18

Note: GI sample restricted to younger age groups (under 35 years of age) or older age groups (35 years and above). Risk ratios, standard errors in parentheses. Specification (1) estimates the probability of aspirations to migrate permanently via binomial probit; reference category is 'no aspirations to migrate permanently' (n = 665 (GI sample, women), n = 473 (GI sample, men)). Specifications (2) and (4) estimate the probability of intentions to migrate temporarily, and specifications (3) and (5) estimate the probability of intentions to migrate permanently via multinomial probit; reference category is 'no intentions to migrate' (n = 678 (GI sample, women), n = 443 (GI sample, men)). Respondents who gave no answer regarding their gender or indicated *"No answer/prefer not to say"* are excluded from the gender subsamples.

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

Table 6.A11: Intentions by aspirations. High-income countries.

Panel A: GI sample

	no aspirations to migrate permanently		aspirations to migrate permanently	
	(1) intentions to migrate temporarily	(2) intentions to migrate permanently	(3) intentions to migrate temporarily	(4) intentions to migrate permanently
	(n = 280)	*(n = 86)*	*(n = 368)*	*(n = 644)*
Gender: woman	0.919	0.516***	1.122	0.686**
	(0.158)	(0.242)	(0.183)	(0.163)
Age: under 18 years	1.037	0.507	0.206***	0.500**
	(0.439)	(0.857)	(0.422)	(0.345)
Age: 25 to 34 years	0.894	1.082	1.229	1.929**
	(0.268)	(0.452)	(0.284)	(0.261)
Age: 35 to 49 years	0.351***	1.142	0.646	0.813
	(0.315)	(0.503)	(0.348)	(0.328)
Age: 50 to 64 years	0.143***	0.121***	0.174***	0.417**
	(0.412)	(0.777)	(0.496)	(0.433)
Age: 65 years or above	0.074***	0.121**	0.051***	0.122***
	(0.567)	(0.910)	(0.887)	(0.701)
University degree	1.707**	1.063	0.896	0.773
	(0.228)	(0.373)	(0.239)	(0.223)
Student	0.863	0.934	0.589**	0.709
	(0.263)	(0.453)	(0.255)	(0.231)
Migration experience	1.047	1.250	1.286	1.234
	(0.163)	(0.255)	(0.183)	(0.167)
Partner: native	0.863	0.882	0.905	1.074
	(0.186)	(0.307)	(0.228)	(0.214)
Partner: non-native	1.460	1.493	1.238	2.610***
	(0.284)	(0.437)	(0.391)	(0.361)
Children	0.883	1.123	0.827	0.916
	(0.258)	(0.360)	(0.331)	(0.294)
Willingness to take risks	1.105***	1.115*	1.132***	1.145***
	(0.038)	(0.061)	(0.042)	(0.037)
Patience	0.955	0.963	0.960	0.982
	(0.033)	(0.053)	(0.036)	(0.033)
Number of observations	1137	1137	1296	1296
Country FE	✓	✓	✓	✓
Correctly predicted values	69.6	69.6	57.6	57.6
McFadden Pseudo R2	0.11	0.11	0.10	0.10

Panel B: Student sample

	(n = 281)	*(n = 39)*	*(n = 345)*	*(n = 165)*
Gender: woman	0.877	0.954	1.428*	1.297
	(0.159)	(0.361)	(0.212)	(0.248)
Age: 25 to 34 years	1.197	0.973	0.728	0.736
	(0.226)	(0.538)	(0.352)	(0.399)
Age: 35 to 49 years	0.160**	/†	0.201*	1.044
	(0.799)		(0.924)	(0.873)
University degree	1.134	1.202	1.979***	1.504
	(0.166)	(0.365)	(0.223)	(0.264)
Migration experience	2.053***	0.904	1.633**	2.422***
	(0.161)	(0.391)	(0.213)	(0.252)
Partner: native	1.166	1.030	1.202	1.364
	(0.160)	(0.355)	(0.223)	(0.260)
Partner: non-native	1.423	2.050	2.180	5.213***
	(0.409)	(0.824)	(0.573)	(0.581)
Children	0.849	/†	0.478	0.110**
	(0.503)		(0.845)	(1.022)
Willingness to take risks	1.157***	0.854*	1.139**	1.102
	(0.039)	(0.088)	(0.052)	(0.061)
Patience	0.994	0.861**	1.060	1.147***
	(0.034)	(0.075)	(0.043)	(0.051)
Number of observations	920	920	674	674
Country FE	✓	✓	✓	✓
Correctly predicted values	65.1	65.1	53.9	53.9
McFadden Pseudo R2	0.07	0.07	0.07	0.07

Note: Risk ratios, standard errors in parentheses. Specifications (1) and (2) estimate the probability of having intentions to migrate temporarily or permanently for those with no aspirations to migrate permanently; reference category is 'no intentions to migrate' (*n* = 771 (GI sample), *n* = 600 (student sample)). Specifications (3) and (4) estimate the probability of having intentions to migrate temporarily or permanently for those with aspirations to migrate permanently; reference category is 'no intentions to migrate' (*n* = 284 (GI sample), *n* = 164 (student sample)). High-income countries include countries which have a GNI per capita larger than $12,535 in current US-Dollars, as of 2020 (Czechia, Great Britain, Italy, Japan, Netherlands, Poland, Romania, South Korea, and Spain); middle-income countries are countries which have a GNI per capita of $1,036 to $12,535 in current US-Dollars, as of 2020 (Bosnia, India, Indonesia, Mexico, and Ukraine). All EU member states in the sample are high-income countries, and all middle-income countries are not EU member states. †Since none of the respondents who have no migration aspirations but intentions to migrate permanently (*n* = 39) are between 35 and 49 years of age or have children, the coefficients are not meaningful. *p<0.1; **p<0.05; ***p<0.01

Table 6.A12: Intentions by aspirations. Middle-income countries.

Panel A: GI sample

	no aspirations to migrate permanently		aspirations to migrate permanently	
	(1) intentions to migrate temporarily	(2) intentions to migrate permanently	(3) intentions to migrate temporarily	(4) intentions to migrate permanently
	(n = 186)	*(n = 64)*	*(n = 422)*	*(n = 1331)*
Gender: woman	1.059	1.100	1.133	1.028
	(0.199)	(0.295)	(0.158)	(0.134)
Age: under 18 years	1.331	1.999	0.879	1.269
	(0.341)	(0.529)	(0.271)	(0.228)
Age: 25 to 34 years	0.742	1.595	1.283	1.503*
	(0.352)	(0.520)	(0.265)	(0.235)
Age: 35 to 49 years	0.650	1.416	0.529	0.660
	(0.519)	(0.691)	(0.406)	(0.337)
Age: 50 to 64 years	/†	1.621	0.190*	0.574
		(1.250)	(0.901)	(0.580)
University degree	2.192***	1.225	1.433*	1.071
	(0.270)	(0.430)	(0.203)	(0.175)
Student	1.897**	1.304	1.065	0.730*
	(0.292)	(0.469)	(0.219)	(0.190)
Migration experience	1.279	1.080	1.652**	1.707***
	(0.288)	(0.419)	(0.218)	(0.194)
Partner: native	0.783	1.069	1.131	1.340
	(0.339)	(0.445)	(0.245)	(0.214)
Partner: non-native	0.529	2.282	0.918	1.621
	(0.553)	(0.669)	(0.409)	(0.348)
Children	0.891	0.683	0.948	1.441
	(0.496)	(0.600)	(0.456)	(0.370)
Willingness to take risks	0.944	1.057	1.074*	1.102***
	(0.051)	(0.078)	(0.039)	(0.033)
Patience	1.051	0.945	1.007	0.992
	(0.045)	(0.065)	(0.034)	(0.029)
Number of observations	606	606	2071	2017
Country FE	✓	✓	✓	✓
Correctly predicted values	59.9	59.9	64.3	64.3
McFadden Pseudo R2	0.06	0.06	0.05	0.05

Panel B: Student sample

	(n = 134)	*(n = 35)*	*(n = 345)*	*(n = 344)*
Gender: woman	0.946	1.113	0.926	0.920
	(0.227)	(0.408)	(0.166)	(0.166)
Age: 25 to 34 years	1.015	0.805	2.103***	1.453
	(0.318)	(0.571)	(0.237)	(0.243)
Age: 35 to 49 years	0.976	2.079	2.162*	0.954
	(0.545)	(0.772)	(0.418)	(0.480)
University degree	1.390	1.655	1.294	1.230
	(0.246)	(0.448)	(0.190)	(0.185)
Migration experience	1.688**	0.819	1.324	1.444**
	(0.254)	(0.534)	(0.183)	(0.184)
Partner: native	0.652	3.278***	0.739	0.881
	(0.267)	(0.445)	(0.185)	(0.185)
Partner: non-native	0.662	14.024***	1.633	3.508**
	(0.818)	(0.835)	(0.630)	(0.594)
Children	1.080	1.224	0.334***	0.281***
	(0.540)	(0.802)	(0.404)	(0.478)
Willingness to take risks	1.074	1.019	1.066	1.090**
	(0.050)	(0.090)	(0.040)	(0.040)
Patience	1.016	0.950	1.023	1.030
	(0.044)	(0.077)	(0.033)	(0.033)
Number of observations	640	640	1008	1008
Country FE	✓	✓	✓	✓
Correctly predicted values	73.6	73.6	41.6	41.6
Correctly predicted values	69.2	69.2	46.2	46.2
McFadden Pseudo R2	0.06	0.06	0.05	0.05

Note: Risk ratios, standard errors in parentheses. Specifications (1) and (2) estimate the probability of having intentions to migrate temporarily or permanently for those with no aspirations to migrate permanently; reference category is 'no intentions to migrate' (*n* = 356 (GI sample), *n* = 471 (student sample)). Specifications (3) and (4) estimate the probability of having intentions to migrate temporarily or permanently for those with aspirations to migrate permanently; reference category is 'no intentions to migrate' (*n* = 318 (GI sample), *n* = 319 (student sample)). High-income countries include countries which have a GNI per capita larger than \$12,535 in current US-Dollars, as of 2020 (Czechia, Great Britain, Italy, Japan, Netherlands, Poland, Romania, South Korea, and Spain); middle-income countries are countries which have a GNI per capita of \$1,036 to \$12,535 in current US-Dollars, as of 2020 (Bosnia, India, Indonesia, Mexico, and Ukraine). All EU member states in the sample are high-income countries, and all middle-income countries are not EU member states. †Since none of the respondents who have no migration aspirations but intentions to migrate temporarily (*n* = 186) are between 50 and 64 years of age, the coefficient is not meaningful. *p<0.1; **p<0.05; ***p<0.01

Table 6.A13: Intentions by aspirations. Women.

Panel A: GI sample

	no aspirations to migrate permanently		aspirations to migrate permanently	
	(1) intentions to migrate temporarily	(2) intentions to migrate permanently	(3) intentions to migrate temporarily	(4) intentions to migrate permanently
	(n = 264)	*(n = 75)*	*(n = 483)*	*(n = 1107)*
Age: under 18 years	1.014	3.637**	0.621	1.228
	(0.395)	(0.601)	(0.302)	(0.255)
Age: 25 to 34 years	0.790	0.906	1.378	2.079***
	(0.285)	(0.529)	(0.246)	(0.224)
Age: 35 to 49 years	0.397***	1.296	0.895	0.905
	(0.342)	(0.575)	(0.321)	(0.297)
Age: 50 to 64 years	0.078***	0.284	0.202***	0.513
	(0.613)	(0.950)	(0.556)	(0.428)
Age: 65 years or above	0.105***	0.247	0.182**	0.153**
	(0.706)	(1.240)	(0.865)	(0.876)
University degree	1.457*	1.905	1.232	0.917
	(0.229)	(0.414)	(0.196)	(0.175)
Student	1.220	1.590	0.995	0.832
	(0.264)	(0.483)	(0.215)	(0.193)
Migration experience	1.323	1.366	1.802***	1.746***
	(0.188)	(0.320)	(0.177)	(0.162)
Partner: native	0.762	1.002	1.057	0.995
	(0.208)	(0.352)	(0.209)	(0.195)
Partner: non-native	1.308	2.176	1.010	1.997**
	(0.323)	(0.495)	(0.345)	(0.311)
Children	1.172	1.055	0.624	1.003
	(0.297)	(0.430)	(0.338)	(0.292)
Willingness to take risks	1.051	1.112	1.063*	1.105***
	(0.039)	(0.066)	(0.037)	(0.032)
Patience	0.962	0.902*	1.024	1.020
	(0.034)	(0.056)	(0.031)	(0.028)
Number of observations	990	990	1953	1953
Country FE	✓	✓	✓	✓
Correctly predicted values	66.4	66.4	59.4	59.4
McFadden Pseudo R2	0.09	0.09	0.08	0.08

Panel B: Student sample

	(n = 217)	*(n = 40)*	*(n = 412)*	*(n = 271)*
Age: 25 to 34 years	1.374	0.851	1.425	1.138
	(0.249)	(0.521)	(0.264)	(0.284)
Age: 35 to 49 years	0.276*	0.236	1.181	1.058
	(0.665)	(1.224)	(0.503)	(0.560)
University degree	0.988	1.329	1.569**	1.430*
	(0.191)	(0.382)	(0.186)	(0.204)
Migration experience	2.585***	1.071	1.544**	1.831***
	(0.182)	(0.410)	(0.181)	(0.196)
Partner: native	1.309	1.048	0.928	0.863
	(0.182)	(0.380)	(0.180)	(0.199)
Partner: non-native	1.725	6.299***	3.330*	7.154***
	(0.439)	(0.643)	(0.636)	(0.629)
Children	0.492	2.161	0.302**	0.234***
	(0.570)	(0.919)	(0.475)	(0.555)
Willingness to take risks	1.088*	0.888	1.100**	1.127**
	(0.044)	(0.086)	(0.043)	(0.047)
Patience	0.995	0.894	1.080**	1.058
	(0.037)	(0.075)	(0.034)	(0.037)
Number of observations	816	816	945	945
Country FE	✓	✓	✓	✓
Correctly predicted values	68	68	47.6	47.6
McFadden Pseudo R2	0.08	0.08	0.05	0.05

Note: Risk ratios, standard errors in parentheses. Specifications (1) and (2) estimate the probability of having intentions to migrate temporarily or permanently for those with no aspirations to migrate permanently; reference category is 'no intentions to migrate' (*n* = 651 (GI sample), *n* = 559 (student sample)). Specifications (3) and (4) estimate the probability of having intentions to migrate temporarily or permanently for those with aspirations to migrate permanently; reference category is 'no intentions to migrate' (*n* = 363 (GI sample), *n* = 262 (student sample)). Respondents who gave no answer regarding their gender or indicated *"No answer/prefer not to say"* are excluded from the gender subsamples.
*p<0.1; **p<0.05; ***p<0.01

Table 6.A14: Intentions by aspirations. Men.

Panel A: GI sample

	no aspirations to migrate permanently		aspirations to migrate permanently	
	(1) intentions to migrate temporarily	(2) intentions to migrate permanently	(3) intentions to migrate temporarily	(4) intentions to migrate permanently
	(n = 187)	*(n = 71)*	*(n = 285)*	*(n = 806)*
Age: under 18 years	1.989*	0.516	0.604	0.746
	(0.391)	(0.637)	(0.360)	(0.306)
Age: 25 to 34 years	1.107	2.314*	1.375	1.295
	(0.323)	(0.459)	(0.320)	(0.279)
Age: 35 to 49 years	0.325***	1.861	0.466*	0.534*
	(0.430)	(0.571)	(0.444)	(0.381)
Age: 50 to 64 years	0.209***	0.214*	0.159***	0.309**
	(0.538)	(0.925)	(0.688)	(0.554)
Age: 65 years or above	0.069***	/†	0.025***	0.076***
	(0.863)		(1.319)	(0.872)
University degree	2.729***	0.669	1.096	1.025
	(0.302)	(0.411)	(0.265)	(0.230)
Student	1.059	0.765	0.786	0.600**
	(0.311)	(0.450)	(0.270)	(0.233)
Migration experience	1.009	1.178	1.106	1.170
	(0.230)	(0.315)	(0.231)	(0.202)
Partner: native	0.996	0.910	0.950	1.615*
	(0.271)	(0.382)	(0.290)	(0.251)
Partner: non-native	1.036	1.208	1.083	2.147*
	(0.413)	(0.596)	(0.529)	(0.459)
Children	0.681	0.945	1.119	1.271
	(0.374)	(0.463)	(0.467)	(0.387)
Willingness to take risks	1.054	1.073	1.148***	1.128***
	(0.051)	(0.074)	(0.047)	(0.038)
Patience	1.006	1.015	0.923*	0.926**
	(0.044)	(0.064)	(0.042)	(0.036)
Number of observations	705	705	1316	1316
Country FE	✓	✓	✓	✓
Correctly predicted values	65.4	65.4	64.1	64.1
McFadden Pseudo R2	0.13	0.13	0.10	0.10

Panel B: Student sample

	(n = 197)	*(n = 34)*	*(n = 274)*	*(n = 236)*
Age: 25 to 34 years	0.887	0.859	1.416	1.144
	(0.267)	(0.590)	(0.292)	(0.306)
Age: 35 to 49 years	0.729	5.510**	1.154	0.629
	(0.605)	(0.836)	(0.541)	(0.621)
University degree	1.469*	1.410	1.681**	1.290
	(0.198)	(0.424)	(0.220)	(0.226)
Migration experience	1.343	0.586	1.284	1.550*
	(0.206)	(0.509)	(0.215)	(0.226)
Partner: native	0.707	3.109***	0.854	1.264
	(0.214)	(0.429)	(0.225)	(0.233)
Partner: non-native	0.591	/††	0.992	2.405
	(0.634)		(0.615)	(0.593)
Children	1.492	0.164	0.390*	0.234**
	(0.516)	(1.245)	(0.557)	(0.699)
Willingness to take risks	1.174***	0.990	1.105**	1.082
	(0.044)	(0.091)	(0.048)	(0.048)
Patience	0.992	0.951	0.974	1.088**
	(0.039)	(0.078)	(0.041)	(0.043)
Number of observations	740	740	730	730
Country FE	✓	✓	✓	✓
Correctly predicted values	69.6	69.6	47.3	47.3
McFadden Pseudo R2	0.08	0.08	0.06	0.06

Note: Risk ratios, standard errors in parentheses. Specifications (1) and (2) estimate the probability of having intentions to migrate temporarily or permanently for those with no aspirations to migrate permanently; reference category is 'no intentions to migrate' ($n = 447$ (GI sample), $n = 509$ (student sample)). Specifications (3) and (4) estimate the probability of having intentions to migrate temporarily or permanently for those with aspirations to migrate permanently; reference category is 'no intentions to migrate' ($n = 225$ (GI sample), $n = 220$ (student sample)). Respondents who gave no answer regarding their gender or indicated *"No answer/prefer not to say"* are excluded from the gender subsamples. †Since none of the respondents who have no migration aspirations but intentions to migrate permanently ($n = 71$) are 65 years or above, the coefficient is not meaningful. ††Since none of the respondents who have no migration aspirations but intentions to migrate permanently ($n = 34$) have a non-native partner, the coefficient is not meaningful.
*p<0.1; **p<0.05; ***p<0.01

Table 6.A15: Intentions by aspirations. GI sample, 6 countries. Younger age groups (under 35 years of age) and older age groups (35 years and above).

GI sample, younger age groups (under 35 years of age)

	no aspirations to migrate permanently		aspirations to migrate permanently	
	(1) intentions to migrate temporarily	(2) intentions to migrate permanently	(3) intentions to migrate temporarily	(4) intentions to migrate permanently
	(n = 273)	(n = 62)	(n = 444)	(n = 915)
Gender: woman	0.908	0.943	1.138	0.969
	(0.169)	(0.297)	(0.157)	(0.138)
Age: under 18 years	1.191	1.021	0.524**	0.903
	(0.311)	(0.492)	(0.282)	(0.240)
Age: 25 to 34 years	1.045	1.393	1.253	1.422
	(0.263)	(0.479)	(0.254)	(0.235)
University degree	2.018***	0.944	1.478*	1.093
	(0.222)	(0.402)	(0.199)	(0.181)
Student	1.492	1.118	0.927	0.744
	(0.244)	(0.431)	(0.216)	(0.196)
Migration experience	1.276	1.024	1.970***	1.655**
	(0.209)	(0.391)	(0.209)	(0.198)
Partner: native	1.049	0.824	1.351	1.508*
	(0.236)	(0.448)	(0.252)	(0.236)
Partner: non-native	1.026	1.823	1.270	2.497*
	(0.368)	(0.605)	(0.509)	(0.468)
Children	0.690	0.289	0.795	1.671
	(0.452)	(1.099)	(0.869)	(0.770)
Willingness to take risks	0.995	1.093	1.110***	1.109***
	(0.043)	(0.077)	(0.040)	(0.035)
Patience	0.992	0.967	0.995	0.987
	(0.036)	(0.064)	(0.033)	(0.030)
Number of observations	763	763	1680	1680
Country FE	✓	✓	✓	✓
Correctly predicted values	56.7	56.7	56.8	56.8
McFadden Pseudo R2	0.03	0.03	0.05	0.05

GI sample, older age groups (35 years and above)

	(n = 47)	(n = 26)	(n = 51)	(n = 101)
Gender: woman	0.887	1.357	1.945	0.510
	(0.375)	(0.483)	(0.519)	(0.431)
Age: 50 to 64 years	0.616	0.241*	0.151***	0.703
	(0.484)	(0.810)	(0.710)	(0.501)
Age: 65 years or above	0.183**	0.159*	0.421	0.432
	(0.811)	(1.108)	(1.070)	(0.943)
University degree	1.241	/†	1.149	0.387
	(0.590)		(0.926)	(0.749)
Migration experience	1.208	1.701	1.598	0.650
	(0.379)	(0.521)	(0.523)	(0.444)
Partner: native	0.794	0.516	0.757	0.867
	(0.471)	(0.633)	(0.668)	(0.595)
Partner: non-native	2.196	0.932	0.378	0.836
	(0.587)	(0.727)	(0.798)	(0.703)
Children	0.521	1.310	0.224***	0.473
	(0.425)	(0.571)	(0.551)	(0.459)
Willingness to take risks	1.152	1.210	0.780**	1.172
	(0.107)	(0.131)	(0.123)	(0.110)
Patience	0.807**	0.805**	0.900	0.866*
	(0.084)	(0.107)	(0.102)	(0.086)
Number of observations	285	285	193	193
Country FE	✓	✓	✓	✓
Correctly predicted values	76.1	76.1	66.8	66.8
McFadden Pseudo R2	0.15	0.15	0.21	0.21

Note: GI sample restricted to younger age groups (under 35 years of age) or older age groups (35 years and above), and to the 6 countries which are also observed in the student sample. Risk ratios, standard errors in parentheses. Specifications (1) and (2) estimate the probability of having intentions to migrate temporarily or permanently for those with no aspirations to migrate permanently; reference category is 'no intentions to migrate' (n = 428 (GI sample, younger age groups), n = 212 (GI sample, older age groups)). Specifications (3) and (4) estimate the probability of having intentions to migrate temporarily or permanently for those with aspirations to migrate permanently; reference category is 'no intentions to migrate' (n = 321 (GI sample, younger age groups), n = 41 (GI sample, older age groups)). Since only 1 of 478 respondents in the GI older-age-group sample is a student, we control for it, but do not report the coefficient. †Since all respondents in the older-age-group sample who have no migration aspirations but who have intentions to migrate permanently (n = 26) possess a university degree, the coefficient is not meaningful.
*p<0.1; **p<0.05; ***p<0.01

Table 6.A16: Intentions by aspirations. GI sample, 14 countries. Younger age groups (under 35 years of age) and older age groups (35 years and above).

GI sample, younger age groups (under 35 years of age)

	no aspirations to migrate permanently		aspirations to migrate permanently	
	(1)	(2)	(3)	(4)
	intentions to migrate temporarily	intentions to migrate permanently	intentions to migrate temporarily	intentions to migrate permanently
	(n = 386)	(n = 106)	(n = 679)	(n = 1682)
Gender: woman	0.968	0.723	1.075	0.899
	(0.138)	(0.225)	(0.129)	(0.112)
Age: under 18 years	1.187	1.140	0.600**	1.050
	(0.265)	(0.418)	(0.227)	(0.193)
Age: 25 to 34 years	0.913	1.529	1.201	1.549**
	(0.210)	(0.353)	(0.195)	(0.175)
University degree	1.708***	0.819	1.192	0.979
	(0.186)	(0.319)	(0.163)	(0.146)
Student	1.199	0.993	0.890	0.759*
	(0.194)	(0.330)	(0.167)	(0.147)
Migration experience	1.060	1.123	1.552***	1.626***
	(0.165)	(0.274)	(0.159)	(0.145)
Partner: native	0.905	1.067	1.139	1.324
	(0.186)	(0.307)	(0.194)	(0.178)
Partner: non-native	1.053	1.762	1.462	2.207**
	(0.307)	(0.529)	(0.372)	(0.344)
Children	0.923	0.594	0.771	1.197
	(0.370)	(0.601)	(0.576)	(0.500)
Willingness to take risks	1.018	1.093	1.126***	1.118***
	(0.033)	(0.056)	(0.031)	(0.027)
Patience	0.996	0.994	0.996	0.999
	(0.029)	(0.049)	(0.027)	(0.024)
Number of observations	1160	1160	2845	2845
Country FE	✓	✓	✓	✓
Correctly predicted values	58.5	58.5	60.9	60.9
McFadden Pseudo R2	0.04	0.04	0.07	0.07

GI sample, older age groups (35 years and above)

	(n = 80)	(n = 44)	(n = 111)	(n = 293)
Gender: woman	0.994	0.738	1.481	0.673
	(0.280)	(0.357)	(0.326)	(0.273)
Age: 50 to 64 years	0.409**	0.182***	0.286***	0.708
	(0.359)	(0.605)	(0.409)	(0.326)
Age: 65 years or above	0.226***	0.118***	0.153**	0.266**
	(0.541)	(0.823)	(0.732)	(0.589)
University degree	1.736	/†	0.955	0.551
	(0.509)		(0.541)	(0.451)
Migration experience	1.282	1.592	1.756*	1.232
	(0.280)	(0.379)	(0.313)	(0.267)
Partner: native	0.653	0.670	0.518*	0.770
	(0.340)	(0.476)	(0.386)	(0.351)
Partner: non-native	1.107	1.345	0.349*	1.914
	(0.444)	(0.575)	(0.553)	(0.465)
Children	0.892	1.329	0.933	0.899
	(0.319)	(0.416)	(0.339)	(0.290)
Willingness to take risks	1.163**	1.058	0.960	1.114*
	(0.073)	(0.092)	(0.071)	(0.060)
Patience	0.920	0.854*	0.927	0.919
	(0.061)	(0.081)	(0.064)	(0.055)
Number of observations	583	583	522	522
Country FE	✓	✓	✓	✓
Correctly predicted values	79.1	79.1	64.8	64.8
McFadden Pseudo R2	0.16	0.16	0.20	0.20

Note: GI sample restricted to younger age groups (under 35 years of age) or older age groups (35 years and above). Risk ratios, standard errors in parentheses. Specifications (1) and (2) estimate the probability of having intentions to migrate temporarily or permanently for those with no aspirations to migrate permanently; reference category is 'no intentions to migrate' ($n = 668$ (GI sample, younger age groups), $n = 459$ (GI sample, older age groups)). Specifications (3) and (4) estimate the probability of having intentions to migrate temporarily or permanently for those with aspirations to migrate permanently; reference category is 'no intentions to migrate' ($n = 484$ (GI sample, younger age groups), $n = 118$ (GI sample, older age groups)). Since only 3 of 1105 respondents in the GI older-age-groups-sample are students, we control for it but do not report the coefficient. †Since all respondents who have no migration aspirations but who have intentions to migrate permanently ($n = 44$) possess a university degree, the coefficient is not meaningful.
*p<0.1; **p<0.05; ***p<0.01

Table 6.A17: Intentions by aspirations. GI sample, 14 countries. Younger age groups (under 35 years of age), by gender.

GI sample, younger age groups (under 35 years of age), women

	no aspirations to migrate permanently		aspirations to migrate permanently	
	(1)	(2)	(3)	(4)
	intentions to migrate temporarily	intentions to migrate permanently	intentions to migrate temporarily	intentions to migrate permanently
	(n = 219)	*(n = 51)*	*(n = 415)*	*(n = 952)*
Age: under 18 years	1.022	3.896**	0.662	1.362
	(0.399)	(0.634)	(0.308)	(0.262)
Age: 25 to 34 years	0.809	0.941	1.292	1.882***
	(0.291)	(0.576)	(0.254)	(0.231)
University degree	1.348	1.471	1.257	0.992
	(0.247)	(0.478)	(0.209)	(0.188)
Student	1.196	1.520	1.069	0.891
	(0.266)	(0.510)	(0.220)	(0.197)
Migration experience	1.280	1.262	2.175***	2.050***
	(0.221)	(0.424)	(0.206)	(0.191)
Partner: native	0.871	1.277	1.069	0.966
	(0.238)	(0.453)	(0.244)	(0.229)
Partner: non-native	1.301	3.645*	1.390	2.144*
	(0.416)	(0.742)	(0.453)	(0.420)
Children	0.959	0.345	0.694	1.003
	(0.486)	(1.102)	(0.665)	(0.580)
Willingness to take risks	1.019	1.078	1.094**	1.118***
	(0.044)	(0.080)	(0.041)	(0.036)
Patience	0.953	0.895	1.047	1.043
	(0.039)	(0.071)	(0.034)	(0.030)
Number of observations	655	655	1160	1160
Country FE	✓	✓	✓	✓
Correctly predicted values	60.3	60.3	59.4	59.4
McFadden Pseudo R2	0.06	0.06	0.07	0.07

GI sample, younger age groups (under 35 years of age), men

	(n = 158)	*(n = 52)*	*(n = 244)*	*(n = 679)*
Age: under 18 years	2.006*	0.460	0.663	0.799
	(0.399)	(0.641)	(0.367)	(0.313)
Age: 25 to 34 years	1.099	2.674**	1.379	1.189
	(0.335)	(0.491)	(0.327)	(0.286)
University degree	2.732***	0.581	1.097	0.984
	(0.315)	(0.455)	(0.281)	(0.246)
Student	1.041	0.665	0.840	0.618**
	(0.317)	(0.464)	(0.274)	(0.237)
Migration experience	0.995	1.236	0.972	1.293
	(0.274)	(0.393)	(0.268)	(0.236)
Partner: native	1.000	0.917	0.996	1.947**
	(0.322)	(0.464)	(0.349)	(0.304)
Partner: non-native	0.908	0.823	1.606	2.440
	(0.498)	(0.876)	(0.704)	(0.651)
Children	0.871	0.769	1.251	2.375
	(0.623)	(0.819)	(1.210)	(1.073)
Willingness to take risks	1.035	1.165*	1.180***	1.118**
	(0.057)	(0.088)	(0.053)	(0.044)
Patience	1.053	1.080	0.926*	0.930*
	(0.050)	(0.077)	(0.046)	(0.040)
Number of observations	473	473	1103	1103
Country FE	✓	✓	✓	✓
Correctly predicted values	59.2	59.2	63.5	63.5
McFadden Pseudo R2	0.09	0.09	0.09	0.09

Note: GI sample restricted to younger age groups (under 35 years of age) or older age groups (35 years and above). Risk ratios, standard errors in parentheses. Specifications (1) and (2) estimate the probability of having intentions to migrate temporarily or permanently for those with no aspirations to migrate permanently; reference category is 'no intentions to migrate' ($n = 385$ (GI sample, women), $n = 263$ (GI sample, men)). Specifications (3) and (4) estimate the probability of having intentions to migrate temporarily or permanently for those with aspirations to migrate permanently; reference category is 'no intentions to migrate' ($n = 293$ (GI sample, women), $n = 180$ (GI sample, women)). Respondents who gave no answer regarding their gender or indicated *"No answer/prefer not to say"* are excluded from the gender subsamples.

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

Chapter 7

General Conclusion

With Chapters 2 to 6 this thesis contributes to three strands of literature. While Part I deals with the effects of foreign language acquisition on the country level, Part II switches to a micro perspective and focuses on motives behind individual decisions to acquire foreign language skills. Also from an individual perspective, Part III contributes to a better understanding of potential migrants in their country of origin before actual migration takes place.

In this final chapter, we try to derive policy implications from our main findings and discuss limitations of our research which open avenues for further research. Since the results focus mainly on Germany, policy implications are explicitly outlined for Germany.

7.1 Policy Implications

In Part I, we have analysed how the presence of German language-learning opportunities abroad affects migration to Germany and German exports. We find in both chapters a significant and positive correlation and give a causal pathway from language-learning opportunities to increased migration flows and exports, respectively. While compulsory language learning is barely within the reach of German policy-markers, language-learning opportunities for adults in the form the Goethe-Institut (GI) is already an essential part of the German foreign cultural policy and their presence can be further extended.

In combination with Jaschke and Keita (2021) who provide evidence for the effects of institutes on positive (self-)selection of migrants related to education and integration outcomes after arrival, we support the strategy of the German government in the context of the new Skilled Immigration Act effective since March 2020. The increasing importance of integrating migrants can be seen in a broad National Integration Action Plan where the GI plays a major role in the pre-integration of migrants (Beauftragte der Bundesregierung für Migration und Flüchtlinge und Integration, 2020). The GI's task is to prepare migrants for their stay in Germany, this includes language courses.

Chapter 4 shows the export-promoting effect of language-learning opportunities. The results emphasize the importance of establishing new institutes where they are more effective. The strength of the effect differs in three aspects. First, it is more effective to open a first institute in a country than further institutes in countries where the GI is already present. Second, the effect is stronger in countries where trade relationships are not yet very profound and that are more distant in an economic, linguistic, and geographical sense. Third, the potential host country of an institute should demand differentiated goods that need a certain degree of communication when traded.

While policy implications in Part I are quite straightforward, this is not the case for Part II and III as they combine existing ideas in a more conceptual way. But in terms of policy implications both have one aspect in common: they show the importance of family-related and gender-related aspects in migration. While language learners in a tied-movers context are more likely to generate spillover effects from consumption motive to professional use, family ties in the origin and the prospective destination country are the main explanation for the mismatch between aspirations and intentions. Both cases show, that governments should not forget family ties when designing migration and integration policies as they play a decisive role in realizing migration decisions and leading to a successful integration by reaping the full potential of a family that decides to migrate.

7.2 Limitations and Future Research

Even though our research contributes to closing research gaps, limitations of our analysis remain and show the need for further research in order to provide robustness and replicability of our results.

Part I analyses the effect of language-learning opportunities on migration and trade. While it establishes a positive connection between them, these results are limited by the available data which focus only on the German language and only on one supplier of language services, the GI. While the GI has the advantage to provide language courses to a large number of countries worldwide, the costs for courses are above average in comparison to similar language courses and might be used only by a selective type of language learners. Other ways of accessing language courses (e.g. at universities and private language schools) cannot be covered by our data. The recent rise in online courses, fueled even more by the COVID-19 pandemic, may have brought about a fundamental shift in the way languages are learned. Digital learning, or the combination of online and face-to-face courses, is becoming increasingly popular and provides access to language courses to a broader group of interested people whose access may also has been previously limited. This is another reason why our results should be replicated with other data on other forms of language acquisition.

While in the context of migration there is barely an alternative language to the main spoken language of the host country, for trade the English language serves often as Lingua Franca. Since our analysis cannot look at this aspect, further research is needed to investigate the connection to the English proficiency of the trade partners. In particular, it is interesting to know under which circumstances or for which type of goods English language skills are a substitute or perhaps even a complement for communication in trade. This allows specifying the effect of language skills on trade and its policy implications.

In the context of migration, research on the individual effects of language skills acquired already in the origin country is largely missing. On the one hand Chapter 3 shows that the presence of the GI increases migration flows to Germany and Jaschke and Keita (2021) provide evidence that the presence of the institutes also increases immigrants language skills and leads to a selection of higher qualified migrants. On the other hands insights on the effect of pre-migration language skills compared to those acquired in the destination country are missing. They would help to design more efficient language courses and legal requirements for migrants in order to improve their economic and social integration into the host country. A potential way of generating representative data on this topic would be to include a module on pre-migration language learning in the IAB-SOEP Migration Sample. Specific questions on the GI would even allow an evaluation of its role in the pre-integration process of future migrants.

Part II and III present both rather novel ideas of combining different existing concepts. What they also have in common, is to some extent missing representativeness for the overall population of interest. While the analysis of motives behind the decision to acquire language skills in Chapter 5 focus solely on the German language, the analysis of potential migrants tries to overcome the focus on Germany by additionally applying the data of university students where a clear focus on Germany is not observable. This is fine for policy implications in the German context but needs further applications of the ideas in order to confirm our results. But also beyond the focus on Germany and the German language, both surveys are selective in terms of education, as they over-represent high-skilled individuals. For many countries, however, medium-skilled individuals are also of interest as demographic change generates needs for these skills. In Germany, the recently implemented Skilled Workers Immigration Act with a special focus on non-academic skilled workers makes this clear (Bundesregierung, 2019). A greater representation in the analysis of potential migration of this group would help to better understand their migration decisions process and to adjust policies to make Germany more attractive as a country of destination.

This thesis can be seen as a first step of establishing the presented ideas, but there is a need for more replications with more representative data of the total population in destination countries for Chapter 6 and of the population of language learners for Chapter 5. After having established our results for potential migration among a broader population, already established results with migration intentions and aspirations (e.g. Bertoli and Ruyssen, 2018; Bertoli et al., 2020; Docquier et al., 2020; Ruyssen and Salomone, 2018) might be replicated in the light of our more precisely distinguished concepts of potential migration and a closer focus on the expected time-frame of the potential migration. Chapter 5 and 6 cannot provide causal inference. While for Chapter 5 identifications strategies to find causal evidence have to be developed, for Chapter 6 replication studies may use the identification strategies of the original studies.

Bibliography

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson (2012): Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration. *American Economic Review* 102(5), 1832–1856.

Adserà, Alícia and Mariola Pytliková (2015): The Role of Language in Shaping International Migration. *The Economic Journal* 125(586), 49–81.

Ajzen, Icek (1991): The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes* 50(2), 179–211.

Aldashev, Alisher, Johannes Gernandt, and Stephan L. Thomsen (2009): Language usage, participation, employment and earnings: Evidence for foreigners in West Germany with multiple sources of selection. *Labour Economics* 16(3), 330–341.

Allansson, Marie, Erik Melander, and Lotta Themnér (2017): Organized violence, 1989–2016. *Journal of Peace Research* 54(4), 574–587.

Alstadsaeter, A. (2011): Measuring the Consumption Value of Higher Education. *CESifo Economic Studies* 57(3), 458–479.

Anderson, James E. and Eric Van Wincoop (2003): Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review* 93(1), 170–192.

Aparicio Fenoll, Ainhoa and Zoë Kuehn (2016): "Does foreign language proficiency foster migration of young individuals within the European Union?" In: *The Economics of Language Policy*. Ed. by Bengt-Arne Wickstroem and Michele Gazzola. Cambridge, MA: MIT Press, 331–355.

Aparicio Fenoll, Ainoa and Zoë Kuehn (2019): Immigrants move where their skills are scarce: Evidence from English proficiency. *Labour Economics* 61, 101748.

Arcidiacono, Peter (2004): Ability Sorting and the Returns to College Major. *Journal of Econometrics* 121(1-2), 343–375.

Aslany, Maryam, Jørgen Carling, Mathilde Bålsrud Mjelva, and Tone Sommerfelt (2021): Systematic Review of Determinants of Migration Aspirations.

QuantMig Project Deliverable 2.2. University of Southampton, Southampton, 1–65.

Auswärtiges Amt (1970): Leitsätze für die auswärtige Kulturpolitik. Bonn.

Auswärtiges Amt and Goethe-Institut (2004): Rahmenvertrag. Munich and Berlin.

Auswärtiges Amt (2015): Deutsch als Fremdsprache weltweit. Datenerhebung 2015. Berlin.

Barbieri, Katherine and Omar M.G. Keshk (2016): Correlates of War Project Trade Data Set Codebook, Version 4.0.

Barbieri, Katherine, Omar M.G. Keshk, and Brian M. Pollins (2009): Trading data: Evaluating our assumptions and coding rules. *Conflict Management and Peace Science* 26(5), 471–491.

Beauftragte der Bundesregierung für Migration und Flüchtlinge und Integration (2020): Nationaler Aktionsplan Integration: Bericht Phase I – Vor der Zuwanderung: Erwartungen steuern – Orientierung geben.

Becker, Gary S. (1964): Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. New York: National Bureau of Economic Research.

Beine, Michel, Frédéric Docquier, and Çaglar Özden (2011): Diasporas. *Journal of Development Economics* 95(1), 30–41.

Beine, Michel, Joël Machado, and Ilse Ruyssen (2020): Do Potential Migrants Internalize Migrant Rights in OECD Host Societies? *Canadian Journal of Economics* 53(4), 1429–1456.

Belot, Michèle and Sjef Ederveen (2011): Cultural barriers in migration between OECD countries. *Journal of Population Economics* 25(3), 1077–1105.

Belot, Michèle and Sjef Ederveen (2012): Cultural Barriers in Migration Between OECD Countries. *Journal of Population Economics* 25(3), 1077–1105.

Bertoli, Simone and Jesús Fernández-Huertas Moraga (2013): Multilateral resistance to migration. *Journal of Development Economics*. Migration and Development 102, 79–100.

Bertoli, Simone and Ilse Ruyssen (2018): Networks and Migrants' Intended Destination. *Journal of Economic Geography* 18, 705–728.

Bertoli, Simone, Jesús Fernández-Huertas Moraga, and Francesc Ortega (2011): Immigration Policies and the Ecuadorian Exodus. *The World Bank Economic Review* 25(1), 57–76.

Bertoli, Simone, Herbert Brücker, and Jesús Fernández-Huertas Moraga (2016): The European crisis and migration to Germany. *Regional Science and Urban Economics* 60, 61–72.

Bertoli, Simone, Frédéric Docquier, Hillel Rapoport, and Ilse Ruyssen (2020): Weather Shocks and Migration Intentions in Western Africa: Insights from a Multilevel Analysis. CESifo Working Papers 8064. Munich.

Bertrand, Marianne and Sendhil Mullainathan (2001): Do People Mean What They Say? Implications for Subjective Survey Data. *The American Economic Review* 91(2), 67–72.

Beverelli, Cosimo and Gianluca Orefice (2019): Migration deflection: The role of Preferential Trade Agreements. *Regional Science and Urban Economics* 79, 103469.

Bleakley, Hoyt and Aimee Chin (2010): Age at Arrival, English Proficiency, and Social Assimilation among US Immigrants. *American Economic Journal: Applied Economics* 2(1), 165–192.

Borjas, George J. (1987): Self-Selection and the Earnings of Immigrants. *American Economic Review* 77(4), 531–553.

Borjas, George J., Ilpo Kauppinen, and Panu Poutvaara (2019): Self-Selection of Emigrants: Theory and Evidence on Stochastic Dominance in Observable and Unobservable Characteristics. *The Economic Journal* 129(617), 143–171.

Bratu, Cristina, Matz Dahlberg, Mattias Engdahl, and Till Nikolka (2020): Spillover effects of stricter immigration policies. *Journal of Public Economics* 190, 104239.

Bredtmann, Julia, Klaus Nowotny, and Sebastian Otten (2020): Linguistic distance, networks and migrants' regional location choice. *Labour Economics* 65, 101863.

Brynjolfsson, Erik, Xiang Hui, and Meng Liu (2019): Does Machine Translation Affect International Trade? Evidence from a Large Digital Platform. *Management Science*.

Budría, Santiago, Alberto Colino, and Carlos Martínez de Ibarreta (2019): The impact of host language proficiency on employment outcomes among immigrants in Spain. *Empirica* 46(4), 625–652.

Bundesamt für Statistik (2012): Ausgewählte Indikatoren im regionalen Vergleich, 2012. URL: https://www.bfs.admin.ch/bfs/de/home/statistiken/regionalstatistik/regionale-portraets-kennzahlen/kantone.assetdetail.2420204.html (visited on 11/30/2021).

– (2016): Statistik der ausländischen Wohnbevölkerung (1991-2009). URL: https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/erhebungen/petra.assetdetail.6805.html (visited on 11/30/2021).

– (2017): Einwanderung der ständigen Wohnbevölkerung nach Kanton, Staatsangehörigkeit (Kategorie), Herkunftsstaat, Geschlecht und Altersklasse. URL: https://https://www.bfs.admin.ch/bfs/de/home/statistiken/

`bevoelkerung / migration – integration / internationale – wanderung .`
`html` (visited on 11/30/2021).

Bundesamtes für Migration und Flüchtlinge (2008): Migrationsbericht des Bundesamtes für Migration und Flüchtlinge im Auftrag der Bundesregierung 2007. Nürnberg.

Bundesregierung (2018): Eckpunkte zur Fachkräfteeinwanderung aus Drittstaaten. Berlin.

– (2019): Entwurf eines Fachkräfteeinwanderungsgesetzes. Drucksache 19/8285.

Burda, Michael C., Wolfgang Härdle, Marlene Müller, and Axel Werwatz (1998): Semiparametric Analysis of German East-West Migration Intentions: Facts and Theory. *Journal of Applied Econometrics* 13(5), 525–541.

Canton, Erik and Frank de Jong (2005): The demand for higher education in The Netherlands, 1950–1999. *Economics of Education Review* 24(6), 651–663.

Carling, Jørgen and Kerilyn Schewel (2018): Revisiting Aspiration and Ability in International Migration. *Journal of Ethnic and Migration Studies* 44(6), 945–963.

Chiswick, Barry R. and Paul W. Miller (1995): The Endogeneity between Language and Earnings: International Analyses. *Journal of Labor Economics* 13(2), 246–288.

– (1998): English Language Fluency Among Immigrants in the United States. *Research in Labor Economics* 17.

– (2015): "Chapter 5 - International Migration and the Economics of Language". In: *Handbook of the Economics of International Migration*. Ed. by Barry R. Chiswick Miller and Paul W. Vol. 1. Handbook of the Economics of International Migration. North-Holland, 211–269.

Chort, Isabelle (2014): Mexican Migrants to the US: What Do Unrealized Migration Intentions Tell Us About Gender Inequalities? *World Development* 59, 535–552.

Coase, Ronald H. (1937): The Nature of the Firm. *Economica* 4(16), 386–405.

Comtrade, UN (2019): United Nations Commodity Trade Statistics Database. (Visited on 11/30/2021).

Coulter, Rory, Maarten van Ham, and Peteke Feijten (2012): Partner (Dis)Agreement on Moving Desires and the Subsequent Moving Behaviour of Couples. *Population, Space and Place* 18(1), 16–30.

DaVanzo, Julie (1983): Repeat Migration in the United States: Who Moves Back and Who Moves On? *The Review of Economics and Statistics* 65(4), 552–559.

De Jong, Gordon F. (2000): Expectations, Gender, and Norms in Migration Decision-Making. *Population Studies* 54(3), 307–319.

Delogu, Marco, Frédéric Docquier, and Joël Machado (2018): Globalizing Labor and the World Economy: The Role of Human Capital. *Journal of Economic Growth* 23(2), 223–258.

Destatis (2016): Qualitätsbericht Wanderungen.

Destatis (2018a): Ausländer: Deutschland, Stichtag, Geschlecht/Altersjahre/ Familienstand, Ländergruppierungen/Staatsangehörigkeit - Tabelle 12521- 0002. URL: `https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/_inhalt.html` (visited on 11/30/2021).

– (2018b): Aussenhandel. URL: `https://www.destatis.de/DE/Themen/Wirtschaft/Aussenhandel/_inhalt.html` (visited on 11/30/2021).

Disdier, Anne-Célia, Silvio Tai, Lionel Fontagné, and Thierry Mayer (2010): Bilateral Trade of Cultural Goods. *Review of World Economics* 145(4), 575–595.

Docquier, Frédéric, Giovanni Peri, and Ilse Ruyssen (2014): The Cross-Country Determinants of Potential and Actual Migration. *International Migration Review* 48, 37–99.

Docquier, Frédéric, Joël Machado, and Khalid Sekkat (2015): Efficiency Gains from Liberalizing Labor Mobility. *The Scandinavian Journal of Economics* 117(2), 303–346.

Docquier, Frédéric, Aysit Tansel, and Riccardo Turati (2020): Do Emigrants Self-Select Along Cultural Traits? Evidence from the MENA Countries. *International Migration Review* 54(2), 388–422.

Dustmann, Christian (2003): Return Migration, Wage Differentials, and the Optimal Migration Duration. *European Economic Review* 47(2), 353–369.

Dustmann, Christian and Francesca Fabbri (2003): Language Proficiency and Labour Market Performance of Immigrants in the UK. *The Economic Journal* 113(489), 695–717.

Dustmann, Christian and Joseph-Simon Görlach (2016): The Economics of Temporary Migrations. *Journal of Economic Literature* 54(1), 98–136.

Dustmann, Christian and Oliver Kirchkamp (2002): The Optimal Migration Duration and Activity Choice after Re-Migration. *Journal of Development Economics* 67(2), 351–372.

Dustmann, Christian and Anna Okatenko (2014): Out-Migration, Wealth Constraints, and the Quality of Local Amenities. *Journal of Development Economics* 110, 52–63.

Dustmann, Christian and Arthur Van Soest (2001): Language Fluency and Earnings: Estimation with Misclassified Language Indicators. *The Review of Economics and Statistics* 83(4), 663–674.

– (2002): Language and the Earnings of Immigrants. *Industrial and Labor Relations Review* 55(3), 473–492.

Egger, Peter H. and Mario Larch (2008): Interdependent Preferential Trade Agreement Memberships: An Empirical Analysis. *Journal of International Economics* 76(2), 384–399.

Egger, Peter H. and Andrea Lassmann (2012): The Language Effect in International Trade: A Meta-Analysis. *Economics Letters* 116(2), 221–224.

– (2015): The Causal Impact of Common Native Language on International Trade: Evidence from a Spatial Regression Discontinuity Design. *The Economic Journal* 125(584), 699–745.

Egger, Peter H. and Farid Toubal (2016): "Common Spoken Languages and International Trade". In: *The Palgrave Handbook of Economics and Language*. Ed. by Victor Ginsburgh and Shlomo Weber. London: Palgrave Macmillan UK, 263–289.

– (2018): Native Language and Acquired Language as Determinants of Product-Level Trade. *The World Economy* 41(7), 1833–1846.

Eugster, Beatrix, Rafael Lalive, Andreas Steinhauer, and Josef Zweimüller (2017): Culture, Work Attitudes, and Job Search: Evidence from the Swiss Language Border. *Journal of the European Economic Association* 15(5), 1056–1100.

Eurostat (2021): Pupils by education level and modern foreign language studied - absolute numbers and % of pupils by language studied - EDUC_UOE_LANG01. URL: `https://ec.europa.eu/eurostat/databrowser/view/EDUC_UOE_LANG01__custom_1382891/default/table` (visited on 11/27/2021).

Falck, Oliver, Robert Gold, and Stephan Heblich (2017): Lifting the Iron Curtain: School-Age Education and Entrepreneurial Intentions. *Journal of Economic Geography* 17(5), 1111–1148.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer (2015): The Next Generation of the Penn World Table. *American Economic Review* 105(10), 3150–3182.

Felbermayr, Gabriel, Volker Grossmann, and Wilhelm Kohler (2015): "Migration, International Trade, and Capital Formation: Cause or Effect?" In: *Handbook of the Economics of International Migration*. Ed. by Barry R. Chiswick and Paul W. Miller. Vol. 1. North-Holland, 913–1025.

Felbermayr, Gabriel J. and Farid Toubal (2010): Cultural Proximity and Trade. *European Economic Review* 54(2), 279–293.

Fidrmuc, Jan and Jarko Fidrmuc (2016): Foreign Languages and Trade: Evidence from a Natural Experiment. *Empirical Economics* 50(1), 31–49.

Földvári, Péter (2006): "The Economic Impact of the European Integration on the Netherlands. A Quantitative Analysis of Foreign Trade and Foreign Direct Investments". PhD Thesis. Utrecht University.

Fouquin, Michel and Jules Hugot (2016): Two Centuries of Bilateral Trade and Gravity Data: 1827-2014. Tech. rep. CEPII research center.

Freedom House (2018): Freedom in the World: Data and Resources. URL: https://freedomhouse.org/ (visited on 11/30/2021).

Friebel, Guido, Juan Miguel Gallego, and Mariapia Mendola (2013): Xenophobic Attacks, Migration Intentions, and Networks: Evidence from the South of Africa. *Journal of Population Economics* 26(2), 555–591.

Fry, Richard and B. Lindsay Lowell (2003): The Value of Bilingualism in the U.S. Labor Market. *Industrial and Labor Relations Review* 57(1), 128–140.

Geist, Claudia and Patricia A. McManus (2011): Different Reasons, Different Results: Implications of Migration by Gender and Family Status. *Demography* 49(1), 197–217.

Ghosh, Sucharita, Donald Lien, and Steven Yamarik (2017): Does the Confucius Institute Network Impact Cultural Distance? A Panel Data Analysis of Cross-Border Flows in and out of China. *Asian Economic Journal* 31(3), 299–323.

Ginsburgh, Victor, Jacques Melitz, and Farid Toubal (2017): Foreign Language Learning and Trade. *Review of International Economics* 25(2), 320–361.

Ginsburgh, Victor A. and Juan Prieto Rodriguez (2011): Returns to Foreign Languages of Native Workers in the European Union. *ILR Review* 64(3), 599–618.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Havard Strand (2002): Armed Conflict 1946-2001: A New Dataset. *Journal of Peace Research* 39(5), 615–637.

Goethe-Institut (2010): Rechenschaftsbericht 2009/10. Munich.

– (2011): Jahrbuch 2010/11. Munich.

– (2013): Rechenschaftsbericht 2012/2013. Munich.

– (2014): Jahrbuch 2013/14. Munich.

– (2015): Jahrbuch 2014/15. Munich.

– (2018): Pressemappe Jahrespressekonferenz 2018. Munich.

– (2020): Jahrbuch 2019/20. Munich.

Goethe Institute (2019): Annual Report 2018/19. Tech. rep. Goethe-Institute e.V. Munich.

Goethe-Institute and Auswärtiges Amt (2004): Rahmenvetrag zwischen Goethe-Institut und der Bundesrepublik Deutschland. Berlin.

Grogger, Jeffrey and Gordon H. Hanson (2011): Income Maximization and the Selection and Sorting of International Migrants. *Journal of Development Economics* 95(1), 42–57.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2009): Cultural Biases in Economic Exchange? *Quarterly Journal of Economics* 124(3), 1095–1131.

Head, Keith and Thierry Mayer (2014): "Gravity Equations: Workhorse,Toolkit, and Cookbook". In: *Handbook of International Economics*. Ed. by Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff. Vol. 4. Handbook of International Economics. Elsevier, 131–195.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein (2008): Estimating Trade Flows: Trading Partners and Trading Volumes. *The Quarterly Journal of Economics* 123(2), 441–487.

Hutchinson, William K. (2005): "Linguistic Distance" as a Determinant of Bilateral Trade. *Southern Economic Journal* 72(1), 1–15.

Isphording, Ingo and Sebastian Otten (2013a): The Costs of Babylon—Linguistic Distance in Applied Economics. *Review of International Economics* 21(2), 354–369.

Isphording, Ingo E. and Sebastian Otten (2013b): The Costs of Babylon—Linguistic Distance in Applied Economics. *Review of International Economics* 21(2), 354–369.

Ivlevs, Artjoms and Roswitha M. King (2012): Family Migration Capital and Migration Intentions. *Journal of Family and Economic Issues* 33(1), 118–129.

Jaeger, David A., Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, and Holger Bonin (2010): Direct Evidence on Risk Attitudes and Migration. *The Review of Economics and Statistics* 92(3), 684–689.

Jaschke, Philipp and Sekou Keita (2021): Say it like Goethe: Language learning facilities abroad and the self-selection of immigrants. *Journal of Development Economics* 149, 102597.

Johnston, Lauren A., Stephen L. Morgan, and Yuesheng Wang (2015): The Gravity of China's African Export Promise. *The World Economy* 38(6), 913–934.

Juster, F. Thomas (1960): Prediction and Consumer Buying Intentions. *The American Economic Review* 50(2), 604–617.

Kaitinnis, Anna (2018): "Das Goethe-Institut als Mittlerorganisation der AKP. Eine Verortung im Rahmen innenpolitischer, außenpolitischer und außenkul-

turpolitischer Entwicklungen". In: *Botschafter der Demokratie*. Wiesbaden: Springer Fachmedien Wiesbaden, 45–80.

Kodde, David A. and Jozef M. M. Ritzen (1984): Integrating Consumption and Investment Motives in a Neoclassical Model of Demand for Education. *Kyklos* 37(4), 598–608.

Lang, Kevin and Erez Siniver (2009): The Return to English in a Non-English Speaking Country: Russian Immigrants and Native Israelis in Israel. *The B.E. Journal of Economic Analysis & Policy* 9(1).

Lazear, Edward (1977): Education: Consumption or Production? *Journal of Political Economy* 85(3), 569–597.

Lien, Donald and Melody Lo (2017): Economic Impacts of Cultural Institutes. *The Quarterly Review of Economics and Finance* 64, 12–21.

Lien, Donald and Chang Hoon Oh (2014): Determinants of the Confucius Institute Establishment. *The Quarterly Review of Economics and Finance* 54(3), 437–441.

Lien, Donald, Chang Hoon Oh, and W. Travis Selmier (2012): Confucius Institute Effects on China's Trade and FDI: Isn't It Delightful When Folks Afar Study Hanyu? *International Review of Economics & Finance* 21(1), 147–155.

Lien, Donald, Melody Lo, and David Bojanic (2019): Asymmetric Effects of Cultural Institutes on Trade and Foreign Direct Investment. *The World Economy* 42(5), 1520–1553.

Lohmann, Johannes (2011): Do language barriers affect trade? *Economics Letters* 110(2), 159–162.

Manchin, Miriam and Sultan Orazbayev (2018): Social Networks and the Intention to Migrate. *World Development* 109, 360–374.

Manski, Charles F. (1990): The Use of Intentions Data to Predict Behavior: A Best-Case Analysis. *Journal of the American Statistical Association* 85(412), 934–940.

Mayda, Anna Maria (2010): International migration: a panel data analysis of the determinants of bilateral flows. *Journal of Population Economics* 23(4), 1249–1274.

Mayer, Thierry and Soledad Zignago (2011): Notes on CEPII's distances measures: The GeoDist database. *CEPII Working Paper* (2011-25).

McFadden, Daniel (1974): "Conditional logit analysis of qualitative choice behavior". In: *Frontiers in Econometrics*. Ed. by Paul Zarembka. New York: Academic Press.

McKenzie, David and Hillel Rapoport (2010): Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks. *The Review of Economics and Statistics* 92(4), 811–821.

243

Mckinnish, Terra (2008): Spousal Mobility and Earnings. *Demography* 45(4), 829–849.

Melitz, Jacques (2008): Language and foreign trade. *European Economic Review* 52(4), 667–699.

Melitz, Jacques and Farid Toubal (2014): Native Language, Spoken Language, Translation and Trade. *Journal of International Economics* 93(2), 351–363.

Mincer, Jacob (1978): Family Migration Decisions. *Journal of Political Economy* 86(5), 749–773.

Munk, Martin D., Till Nikolka, and Panu Poutvaara (2017): International Family Migration and the Dual-Earner Model. CESifo Working Paper 6377. Munich.

Nakosteen, Robert A. and Michael Zimmer (1980): Migration and Income: The Question of Self-Selection. *Southern Economic Journal* 46(3), 840–851.

Netzwerk Deutsch (2010): Statistische Erhebungen 2010: Die deutsche Sprache in der Welt. Berlin et al.

Nye, Joseph S. (2004): Soft Power: The means to Success in World Politics. New York: Public Affairs.

Okun, Arthur M. (1962): The Predictive Value of Surveys of Business Intentions. *The American Economic Review* 52(2), 218–225.

Oosterbeek, Hessel and Hans van Ophem (2000): Schooling choices: Preferences, discount rates, and rates of return. *Empirical Economics* 25(1), 15–34.

Oosterbeek, Hessel and Dinand Webbink (1995): Enrolment in higher education in The Netherlands. *De Economist* 143(3), 367–380.

Ortega, Francesc and Giovanni Peri (2013): The effect of income and immigration policies on international migration. *Migration Studies* 1(1), 47–74.

Papapanagos, Harry and Peter Sanfey (2001): Intention to Emigrate in Transition Countries: The Case of Albania. *Journal of Population Economics* 14(3), 491–504.

Pedersen, Peder J., Mariola Pytlikova, and Nina Smith (2008): Selection and network effects—Migration flows into OECD countries 1990–2000. *European Economic Review* 52(7), 1160–1186.

Pesaran, M. Hashem (2006): Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica* 74(4), 967–1012.

Poutvaara, Panu (2004): Educating Europe: Should Public Education Be Financed with Graduate Taxes or Income-Contingent Loans? *CESifo Economic Studies* 50(4), 663–684.

– (2008): Public and Private Education in an Integrated Europe: Studying to Migrate and Teaching to Stay? *Scandinavian Journal of Economics* 110(3), 591–608.

244

Rauch, James E (1999): Networks versus Markets in International Trade. *Journal of international Economics* 48(1), 7–35.

Rauch, James E. and Vitor Trindade (2002): Ethnic Chinese Networks in International Trade. *The Review of Economics and Statistics* 84(1), 116–130.

Roca Paz, Roberto and Silke Uebelmesser (2021): Risk Attitudes and Migration Decisions. *Journal of Regional Science* 1, 1–36.

Ruyssen, Ilse and Sara Salomone (2018): Female Migration: A Way out of Discrimination? *Journal of Development Economics* 130, 224–241.

Saiz, Albert and Elena Zoido (2005): Listening to What the World Says: Bilingualism and Earnings in the United States. *Review of Economics and Statistics* 87(3), 523–538.

Santos Silva, J. M. C. Santos and Silvana Tenreyro (2006): The Log of Gravity. *The Review of Economics and Statistics* 88(4), 641–658.

Schaafsma, Joseph (1976): The Consumption and Investment Aspects of the Demand for Education. *The Journal of Human Resources* 11(2), 233.

Schneider, Gerald and Julia Schiller (2000): Goethe ist nicht überall : eine empirische Analyse der Standortentscheidungen in der Auswärtigen Kulturpolitik. *Zeitschrift für Internationale Beziehungen* 7(1), 5–32.

Schödel, Günther (2007): "Runderlass Des Ministerialdirigenten Schödel". In: *Akten zur Auswärtigen Politik der Bundesrepublik Deutschland 1976*. München: R. Oldenbourg Verlag, 1162–1167.

Schultz, Theodore W. (1960): Capital Formation by Education. *Journal of Political Economy* 68(6), 571–583.

Selmier, W. Travis and Chang Hoon Oh (2013): The Power of Major Trade Languages in Trade and Foreign Direct Investment. *Review of International Political Economy* 20(3), 486–514.

Singer, Otto (2003): Auswärtige Kulturpolitik in der Bundesrepublik Detuschland - Konzeptionelle Grundlagen und institutionelle Entwicklung seit 1945. Wissenschaftliche Dienste des deutschen Bundestages.

Sjaastad, Larry A. (1962): The Costs and Returns of Human Migration. *Journal of Political Economy* 70(5, Part 2), 80–93.

Staatssekretariat für Migration (2017): Personenfreizügigkeit Schweiz – EU/EFTA. URL: https://www.sem.admin.ch/sem/de/home/themen/fza_schweiz-eu-efta.html (visited on 11/30/2021).

Stark, Oded and David E. Bloom (1985): The New Economics of Labor Migration. *The American Economic Review* 75(2), 173–178.

Stöhr, Tobias (2015): The returns to occupational foreign language use: Evidence from Germany. *Labour Economics* 32, 86–98.

Ständige Arbeitsgruppe Deutsch als Fremdsprache (2003): Deutsch als Fremdsprache weltweit. Datenerhebung 2000. Berlinet al.

– (2006): Deutsch als Fremdsprache weltweit. Datenerhebung 2005. Berlin et al.

Tinbergen, Jan (1962): Shaping the World Economy: Suggestions for an International Economic Policy. New York: The Twentieth Century Fund.

Tjaden, Jasper, Daniel Auer, and Frank Laczko (2019): Linking Migration Intentions with Flows: Evidence and Potential Use. *International Migration* 57(1), 36–57.

Tobin, James (1959): On the Predictive Value of Consumer Intentions and Attitudes. *The Review of Economics and Statistics* 41(1), 1–11.

Toomet, Ott (2011): Learn English, Not the Local Language! Ethnic Russians in the Baltic States. *American Economic Review* 101(3), 526–31.

Uebelmesser, Silke (2006): To Go or Not to Go: Emigration from Germany. *German Economic Review* 7(2), 211–231.

Uebelmesser, Silke, Matthias Huber, and Severin Weingarten (2018a): The German Language Worldwide: A New Data Set on Language Learning. *CESifo Economic Studies* 64(1), 103–121.

– (2018b): The German Language Worldwide: Data – Course Participation Abroad. URL: `https://doi.org/10.7910/DVN/XVNUY8` (visited on 11/30/2021).

– (2018c): The German Language Worldwide: Data – Course Participation in Germany. URL: `https://doi.org/10.7910/DVN/V5BWMT` (visited on 11/30/2021).

– (2018d): The German Language Worldwide: Data – Language Learning Opportunities Abroad. URL: `https://doi.org/10.7910/DVN/VSQBM7` (visited on 11/30/2021).

Uebelmesser, Silke, Ann-Marie Sommerfeld, and Severin Weingarten (2021): A macro-level analysis of language learning and migration. *German Economic Review*.

United Nations (2017): World Population Prospects 2017. New York.

van Dalen, Hendrik P. and Kène Henkens (2012): Explaining Low International Labour Mobility: The Role of Networks, Personality, and Perceived Labour Market Opportunities. *Population, Space and Place* 18(1), 31–44.

Van der Slik, Frans W. P. (2010): Acquisition of Dutch as a Second Language: The Explanative Power of Cognate and Genetic Linguistic Distance Measures for 11 West European First Languages. *Studies in Second Language Acquisition* 32(3), 401–432.

Wen, Chieh-Hua and Frank S. Koppelman (2001): The generalized nested logit model. *Transportation Research Part B: Methodological* 35(7), 627–641.

246

Williams, Allan M, Calvin Jephcote, Hania Janta, and Gang Li (2018): The Migration Intentions of Young Adults in Europe: A Comparative, Multilevel Analysis. *Population, Space and Place* 24(1), 1–16.

Williamson, Oliver E. (1979): Transaction-Cost Economics: The Governance of Contractual Relations. *The Journal of Law and Economics* 22(2), 233–261.

– (1981): The Economics of Organization: The Transaction Cost Approach. *American Journal of Sociology* 87(3), 548–577.

– (1985): The Economic Intstitutions of Capitalism. Free Press.

– (1989): "Transaction Cost Economics". In: *Handbook of Industrial Organization*. Vol. 1. Elsevier, 135–182.

Wittek, Bernhard (2016): "Gütesiegel ‚Goethe' Zur Geschichte eines Instituts zwischen Außen- und Kulturpolitik". In: *Kulturarbeit in Transformationsprozessen*. Ed. by Wolfgang Schneider and Anna Kaitinnis. Wiesbaden: Springer Fachmedien Wiesbaden, 95–99.

World Bank (2021): World Development Indicator.

Statutory Declaration

Erklärung nach §4 PromO: Hiermit erkläre ich,

1. dass mir die geltende Promotionsordnung bekannt ist;

2. dass ich die Dissertation selbst angefertigt, keine Textabschnitte eines Dritten oder eigener Prüfungsarbeiten ohne Kennzeichnung übernommen und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben habe;

3. dass ich bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskriptes keine unzulässige Hilfe in Anspruch genommen habe;

4. dass ich nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen habe und dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen;

5. dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe;

6. dass ich nicht die gleiche, eine in wesentlichen Teilen ähnliche oder eine andere Abhandlung bei einer anderen Hochschule bzw. anderen Fakultät als Dissertation eingereicht habe.

Unterschrift: