Analysis of International Student Mobility by Education Levels

by

Mehmet Zahid Bilgin

A Dissertation Submitted to the
Graduate School of Social Sciences and Humanities
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in

Economics

KOÇ ÜNİVERSİTESİ

September 4, 2025

Analysis of International Student Mobility by Education Levels

Koç University

Graduate School of Social Sciences and Humanities

This is to certify that I have examined this copy of a master's thesis by

Mehmet Zahid Bilgin

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

examining committee have been made.

Committee Members:

Assoc. Prof. Murat Demirci (Advisor)

Prof. Murat Güray Kırdar

Assoc. Prof. Ayşegül Kayaoğlu Yılmaz

Date:

ABSTRACT

Analysis of International Student Mobility by Education Levels
Mehmet Zahid Bilgin
Master of Science in Economics
September 4, 2025

This thesis examines the factors influencing international student mobility and how these effects vary by education level. Drawing on the literature on gravity models of migration, the study analyzes OECD data on international student flows from 2013 to 2020, covering 24 destination countries and 174 origins. Consistent with prior research, the Poisson Pseudo-Maximum Likelihood (PPML) estimator is employed. In addition, the analysis addresses the relatively underexplored issue of multilateral resistance to migration. The key finding is that master's students exhibit a distinct approach to studying abroad: whereas bachelor's and doctoral students tend to follow a migration-based trajectory, master's students adopt a more balanced combination of the migration model and the human capital model. Beyond this overarching result, the study also identifies several education-level-specific variations in the influence of different factors.

ÖZETÇE

Eğitim Düzeylerine Göre Uluslararası Öğrenci Hareketliliğinin Analizi

Mehmet Zahid Bilgin
Ekonomi, Yüksek Lisans
4 Eylül 2025

Bu tez, uluslararası öğrenci hareketliliğini etkileyen faktörleri ve bu etkilerin eğitim düzeyine göre nasıl değiştiğini incelemektedir. Göçün yerçekimi modelleri ile ilgili literatüre dayanarak, çalışma 2013'ten 2020'ye kadar 24 hedef ülke ve 174 menşe ülkeyi kapsayan OECD'nin uluslararası öğrenci akışlarına ilişkin verilerini analiz etmektedir. Önceki araştırmalarla tutarlı olarak, Poisson Pseudo-Maximum Likelihood (PPML) tahmincisi kullanılmaktadır. Ayrıca, analiz, göç konusunda nispeten az araştırılmış olan çok taraflı direnç konusunu ele almaktadır. Temel bulgu, yüksek lisans öğrencilerinin yurtdışında eğitim konusunda farklı bir yaklaşım sergilediğidir: lisans ve doktora öğrencileri göç temelli bir yol izleme eğilimindeyken, yüksek lisans öğrencileri göç modeli ile insan sermayesi modelini daha dengeli bir şekilde birleştirir. Bu genel sonucun ötesinde, çalışma farklı faktörlerin etkisinde eğitim düzeyine özgü çeşitli farklılıklar da tespit etmektedir.

ACKNOWLEDGMENTS

I express my sincere gratitude to my advisor, Assoc. Prof. Murat Demirci, whose feedback and insightful comments gave me confidence to progress during my research. I also thank TÜBİTAK for their financial support through both the 2210-A scholarship program and Project No 121K349. My deepest appreciation goes to my family, who I feel very fortunate to have beside me in times of need. Lastly, the most heartfelt thanks go to my wife. Her endless support helped me overcome the challenges I faced along the way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

CCE	Common Correlated Effects
CEPII	Centre for Prospective Studies and International Information
DIOC	Database on Immigrants in OECD and non-OECD Countries
GDP	Gross Domestic Product
IHS	Inverse Hyperbolic Sine
ILO	International Labour Organization
MRM	Multilateral Resistance to Migration
OECD	Organisation for Economic Co-operation and Development
PPML	Poisson Pseudo-Maximum Likelihood
PPP	Purchasing Power Parity
RUM	Random Utility Model
THE	Times Higher Education

Chapter 1

INTRODUCTION

The number of students pursuing education abroad has increased steadily over the past few decades. With the rise of globalization, students have gained greater access to international educational opportunities. Recognizing the value of internationally mobile students for human capital development, many host countries have actively sought to attract them through targeted policies. While the primary motivation for these policies is often to retain skilled graduates in the labor force, international students also contribute economically during their temporary stay through tuition fees and living expenses. For origin countries, however, the implications are more complex. If students remain abroad permanently, it may result in a loss of human capital—commonly referred to as 'brain drain'. Conversely, if they return with enhanced skills and education, they can significantly contribute to economic growth and human capital development in their home countries.

As the significance of international student mobility has grown, the field of economics has increasingly turned its attention to this topic, leading to a rapidly expanding body of literature. Drawing on insights from the broader migration literature, recent studies have sought to identify the key trends and determinants underlying student mobility across countries. This study contributes to that growing literature by examining the determinants of international student mobility disaggregated by education level. Utilizing detailed OECD data disaggregated by destination, origin, and education level, the analysis aims to uncover how these determinants differ across bachelor's, master's, and doctoral students. Applying a gravity model framework, the results suggest that bachelor's and doctoral students tend to follow similar mobility patterns, while master's students exhibit distinct trends. Interpreted through the theoretical lens developed by Rosenzweig (2006), the mobility of bachelor's and doctoral students aligns more closely with the migration model, which posits that students choose to study abroad primarily with the intent of remaining in the host country for work after graduation. In contrast, the mobility patterns of master's students are more consistent with the human capital model, in which students pursue international education to enhance their skills and subsequently return to their home country.

Chapter 2 of this study reviews the relevant literature. Building on these insights, Chapter 3 details the study's methodology, outlining the underlying theoretical framework and introducing the models employed. Chapter 4 describes the data sources and provides a comprehensive overview of the dataset. Chapter 5 presents the estimation results across different models and compares them in light of the theoretical considerations discussed in Chapter 3. Finally, Chapter 6 offers concluding remarks.

Chapter 2

LITERATURE

The literature on international student mobility has expanded significantly in recent years, driven by improved data availability and growing academic interest. It is important to acknowledge foundational works that laid the groundwork for these advancements. One such seminal contribution is Rosenzweig (2006), which serves as a theoretical cornerstone for the present study.

Rosenzweig proposed two theoretical models to explain the behavior of international students: the constrained domestic schooling model and the migration model. In the constrained domestic schooling model, students are assumed to return to their home countries after graduation, regardless of where they study. Their decision to study abroad depends solely on the cost of acquiring an additional unit of skill. Therefore, if the cost is lower abroad, students choose to study overseas. In this framework, expanding domestic educational capacity would reduce the cost of accessing high-quality education in origin country, which would lead to a decrease in international student mobility.

By contrast, the migration model assumes students are free to remain and work in the destination country after graduation. In this case, labor market conditions become a central factor influencing the decision to study abroad. Students are motivated not only by educational opportunities but also by the prospect of future employment. In this framework, education in the home country complements rather than substitutes for foreign education. As a result, expanding domestic education would increase the likelihood of studying abroad by enabling students to access more favorable labor markets.

Rosenzweig empirically tested these models using three types of U.S.-based data: F-1 visa issuances, the stock of foreign students, and the number of newly naturalized permanent residents who were formerly international students. Each dataset was disaggregated by country of origin and covered a roughly one-year period. His findings provide stronger support for the migration model, indicating that international student mobility to the United States is more closely tied to migration model rather than to constrained domestic schooling model.

Subsequent research has extended the foundational focus on economic and educational factors by adopting multi-origin and multi-destination perspectives—an advancement facilitated by the growing availability of cross-country datasets. For example, Caruso and De Wit (2015) examined the determinants of international student inflows to 33 European countries between 1998 and 2009, concentrating on destination-year variations. Their analysis shows that higher levels of education expenditure, GDP per capita, and trade openness are associated with significantly greater student inflows, whereas inflation has no discernible effect. The relationship between tuition fees and student inflows is nonlinear: inflows initially rise with tuition fees up to a certain threshold, after which they begin to decline. In addition, the study finds that higher crime rates significantly deter international students.

In a related study, Wei (2013) examined international student flows into 48 destination countries from 1999 to 2008. The analysis confirmed the positive influence of trade openness, while the cost of living was found to have an insignificant effect. Similarly, GDP per capita and education expenditure were not statistically significant in explaining student inflows in the overall model, in contrast to the findings of Caruso and De Wit (2015). Wei extended the analysis by disaggregating flows according to combinations of developed and developing countries. A key insight from this approach is that GDP per capita differences have a positive and significant effect on student mobility—except in flows between developed countries. Moreover, educational variables such as the student-teacher ratio and higher education enrollment rate in destination countries were found to be significant determinants, particularly when the development levels of origin and destination countries are similar.

Building on advancements in international trade literature, migration studies have increasingly adopted gravity models to analyze cross-border movements, including international student mobility. These models typically incorporate geographic distance between countries and the population sizes of both origin and destination countries as core explanatory variables. Consistent across all studies in this literature, the coefficient on distance is invariably negative, reflecting the deterrent effect of greater physical separation on student flows. Conversely, the coefficients on population are uniformly positive, indicating that larger populations—whether at the origin or destination—are associated with higher volumes of international student mobility.

Using this framework, Beine et al. (2014) examined student flows between 13 destination countries and 180 origin countries for the period 2004–2007. Their findings underscore the importance of both economic and educational factors, highlighting significant effects of wages, housing prices, and university quality in destination countries. Notably, their inclusion of housing prices as a proxy for living costs revealed a significant negative effect on student inflows—challenging previous studies that found cost of living to be insignificant. In contrast, tuition fees did not appear to deter international students. The study also incorporated cultural variables, revealing that shared official language and larger migrant networks significantly boost student inflows. Interestingly, the inclusion of migrant networks rendered the effect of colonial ties statistically insignificant.

Abbott and Silles (2016) also employed a gravity model framework to examine international student flows between 18 destination countries and 38 origin countries during the period 2005–2011. Consistent with previous literature, their findings indicate that a shared official language and the overall number of outgoing students from an origin country in the current period positively influence bilateral student flows. In contrast, variables such as GDP per capita differences, university rankings, shared borders, real exchange rates, and colonial ties did not exhibit statistically significant effects in the full sample. However, when the analysis was disaggregated by income level of origin countries, GDP per capita differences became significant for non-high-income origins—aligning with the findings of Wei (2013). A notable contribution of this study is the inclusion of time zone differences as a variable. The results suggest that beyond geographic distance, greater time differences between countries also serve as a deterrent to international student mobility.

While some studies mentioned above included distance and population related variables to resemble their empirical specification into gravity equation, a theoretically driven equation includes the multilateral resistance term (Beine et al., 2016). An important yet often underestimated element in the gravity model framework is therefore the multilateral resistance term. Originally developed in international trade literature, this concept has gained increasing relevance in migration studies, including analyses of international student mobility.

Various methodological strategies exist to account for multilateral resistance. The most common approach in this literature is the inclusion of fixed effects. A prominent example is Didisse et al. (2019), who employed origin and destination fixed effects to study average student flows from 89 developing countries to 30 European destinations between 2004 and 2013. Their findings align with the theoretical framework in Rosenzweig (2006)—specifically, they observed that higher origin GDP per capita and origin youth unemployment are associated with lower student outflows, while origin tertiary enrollment has a positive effect. Going beyond standard gravity variables, the authors introduced alternative measures of migrant networks, such as informal networks and bilateral agreements, in place of traditional migrant stock indicators. The informal network variable—proxied by the average flow of tertiary students from a given origin to a destination over the preceding four-year period—proved to be the most significant determinant. In addition, the study disaggregated language-related cultural proximity into shared official language, common spoken language, and common native language. Of these, only common spoken language retained statistical significance when all were included simultaneously. Lastly, the authors incorporated novel monadic variables capturing youth perceptions (as push factors) and public service quality in destination countries (as pull factors), both of which had significant explanatory power for student mobility patterns.

Kaushal and Lanati (2019) also used destination-origin, destination-year and origin-year fixed effects when their specification allowed. They analyzed the student mobility between 141 destination countries and 206 origin countries between 1999 and 2015 through an extended version of framework in Rosenzweig (2006) which also includes interconnectedness. Interconnectedness was conceptualized through bilateral trade and migrant networks, each aligned with distinct theoretical models:

bilateral trade was posited to be more relevant for the human capital model, while migrant networks were expected to play a more prominent role in the migration model. Empirical results confirmed that both factors exert a positive and significant influence on student flows in the full sample. However, the relative strength of these effects varied by context: migrant networks had a stronger effect in developed destination countries, whereas bilateral trade exerted greater influence in developing destination contexts. Similar heterogeneity was observed for other variables, such as origin GDP per capita and educational infrastructure. Notably, student mobility patterns aligned more closely with the migration model when the destination was developed and non-English-speaking, whereas the human capital model was more applicable to developing or English-speaking destinations. Lastly, the quality of institutions—proxied through various governance indicators—was found to significantly affect student mobility, though only within specific subsamples.

In addition to studies treating international students as a single, homogenous group, only a few studies have disaggregated the analysis by education level, offering comparative insights across student cohorts. A notable example is Bird and Turner (2014), who examined undergraduate and graduate international student flows to the United States from 165 origins between 1998 and 2010. Their findings reveal that home-country economic conditions exert a stronger influence on undergraduate student mobility compared to that of graduate students. Specifically, while higher GDP per capita in the origin country is associated with increased outbound student mobility, a weaker home-country currency (i.e., an unfavorable exchange rate) reduces the likelihood of studying abroad. Additionally, limited higher education capacity in the origin country was found to drive student mobility at both educational levels, though the effect was more pronounced for undergraduates.

Focusing specifically on doctoral-level student mobility within Europe, Mathies and Cantwell (2022) analyzed the determinants of cross-border flows among 27 countries using a gravity model framework. Their study aimed to disentangle the effects of academic, economic, and social variables on international mobility at the doctoral level. The findings show that academic factors, particularly R&D expenditures and the presence of top-ranked universities in the destination country, significantly increase the inflow of doctoral students. On the origin side, higher tertiary attainment is associated with greater student outflows. In terms of economic factors, the results support a dual role of GDP per capita: higher income levels in destination countries act as a pull factor, while lower income levels in origin countries serve as a push factor. By contrast, social indicators such as the gender pay gap and life expectancy were found to have no statistically significant effect on doctoral student mobility.

Finally, Van Bouwel and Veugelers (2013) compared the mobility patterns of doctoral students with those of all tertiary-level students in Europe for the year 2007, using a dataset comprising 29 destination and 31 origin countries. The key variable of interest—the quality of education in the destination—was found to have a positive and significant effect on student inflows for both levels. However, contrary to the authors' expectations, this effect was weaker for doctoral students compared to tertiary students overall. Differences were also observed with respect to economic variables: higher public expenditure on higher education and lower tuition fees were associated with increased inflows of tertiary students, but these factors had no statistically significant impact on doctoral student mobility. Lastly, educational opportunities, proxied by the ratio of enrollment at a given education level to that at the immediately lower level, had a significant negative effect on student outflows. This finding suggests that limited educational capacity at lower levels drives doctoral students to seek study opportunities abroad.

Following the literature above, the primary contribution of this study lies in its comprehensive analysis of international student mobility within a multi-education-level framework with data from multiple destinations and origins. To the author's knowledge, this is the first study to systematically examine mobility patterns across different education levels—bachelor's, master's, and doctoral—within a single empirical setup. Additionally, the study introduces a novel approach to account for multilateral resistance to migration without sacrificing the origin-time dimension. Drawing on methodologies from the international trade literature, a new proxy for the multilateral resistance term is incorporated into the model. This innovation enables more precise estimation of key coefficients and offers a refined understanding of the structural factors shaping student mobility.

Chapter 3

METHOD

3.1 Theory

3.1.1 Gravity Model

Newton's theory of gravitation marked a major breakthrough in physics. According to Newton's law, any two objects attract each other with a gravitational force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them. This force is expressed by the following formula:

$$F = G\frac{m_1 m_2}{r^2} \tag{3.1}$$

where m denotes masses, r denotes distance, and G denotes gravitational constant.

After exerting a profound and lasting influence on physics, the simplicity and broad applicability of the gravity model began to attract attention in other disciplines. Viewing international trade as a form of gravitational interaction between countries, scholars in economics adopted the model to study trade flows. Tinbergen (1962) was the first to apply the gravity model in this context, laying the groundwork for a vast empirical literature. In the years that followed, simplified versions of the gravity model were widely employed—even in the absence of formal theoretical underpinnings. Extending this framework to the context of international education, student mobility from country j to country k at time t—m_{jkt}, the dependent variable in this study—can be modeled as follows:

$$m_{jkt} = \alpha_0 \left(Pop_{kt}\right)^{\alpha_1} \left(Pop_{jt}\right)^{\alpha_2} \left(Dist_{jkt}\right)^{\alpha_3} \tag{3.2}$$

In this specification, Pop denotes the population of the respective country, indexed accordingly, while $Distance$ represents the geographic distance between the destination and origin countries. The α terms correspond to the elasticities of the associated variables.

3.1.2 RUM Model

Due to the initial lack of a solid theoretical foundation in applying the gravity model to migration, researchers began turning to more rigorously defined frameworks. One such framework is the Random Utility Model (RUM), which offers a structured approach to understanding individual decision-making regarding migration. The RUM conceptualizes the migration decision as a choice between remaining in the home country or relocating, based on utility maximization. Given that international student mobility represents a specific type of migration, the RUM is well-suited to analyzing student decisions. Following the approach of Beine et al. (2016), the utility function for a student i from origin country j at time $t-1$, choosing at time t whether to study in destination k from a set of possible options D, is specified as follows[1]:

$$U_{ijkt} = w_{jkt} - c_{jkt} + \epsilon_{ijkt} \tag{3.3}$$

in this formulation, w_{jkt} and c_{jkt} represent the deterministic components of utility, where the former captures the benefits associated with migration and the latter reflects the cost of migrating. The stochastic component of utility, ϵ_{ijkt}, is assumed to follow an independently and identically distributed (i.i.d.) extreme value type I distribution, as specified by McFadden (1974).

Based on individuals' utility-maximizing decisions, the probability that a student migrates from origin j to destination k can be expressed as:

[1] While the theoretical framework is constructed using migration flow data, empirical estimation will employ yearly stock data due to limitations in data availability.

$$E\left(p_{jkt}\right)=\frac{e^{w_{jkt}-c_{jkt}}}{\sum_{d\in D}e^{w_{jdt}-c_{jdt}}} \tag{3.4}$$

When the population of the origin j, denoted by s_{jt}, is included in the model, the total number of migrants from origin j to destination k can be estimated using the following equation:

$$E\left(m_{jkt}\right)=\frac{e^{w_{jkt}-c_{jkt}}}{\sum_{d\in D}e^{w_{jdt}-c_{jdt}}}s_{jt} \tag{3.5}$$

If it is assumed that the migration benefit w_{jkt} does not vary across origin countries, the equation can be reformulated in a manner that resembles a gravity model:

$$E\left(m_{jkt}\right)=\frac{l_{jkt}y_{kt}}{\Omega_{jt}}s_{jt} \tag{3.6}$$

where $y_{kt}=e^{w_{kt}}$, $l_{jkt}=e^{-c_{jkt}}$ and $\Omega_{jt}=\sum_{d\in D}l_{jdt}y_{dt}$. In the migration context, the size parameters in the standard gravity equation are represented by y_{kt} for the destination and s_{jt} for the origin. The trade cost parameter is replaced by the migration cost parameter, denoted l_{jkt}.

Unlike the gravity model of trade, which includes two multilateral resistance terms—one for exporters and one for importers—the migration model includes only a single multilateral resistance term, which is Ω_{jt} and varies by origin. This difference stems from the number of decisions involved in each model. In trade, both the exporter and importer make decisions (to sell and to buy), whereas in migration, the decision lies solely with the individual at the origin; the destination country does not actively choose migrants in this framework.

The model relies on several key assumptions regarding individual decision-making. First, it assumes that individuals do not make decisions dynamically; in other words, when deciding whether to migrate at time t, they rule out any other decision after time t. While this assumption may be problematic in the broader migration literature, it is more justifiable in the context of international student mobility, where individuals typically make one-time decisions about where to pursue a post-secondary degree. Second, the model assumes that the deterministic component of utility does not vary across origin countries or individuals, implying a uniform structure of perceived benefits and costs among all potential migrants[2].

3.2 Empirical Specification

Building on the discussion above, two primary model specifications are developed: the traditional gravity model and the theoretically consistent gravity model. The former is introduced first, based on Equation 3.2. In this specification, α_0, α_1, α_2, and α_3 represent coefficients to be estimated. The destination size variable, Pop_{kt}, denotes the population aged 15 to 29 in the destination country (in hundreds of thousands), while Pop_{jt}, represents the corresponding age group in the country of origin (in hundreds of thousands). The dyadic trade cost variable, $Dist_{jkt}$, is measured as the population-weighted harmonic mean[3] distance between the largest cities of the origin and destination countries. The term ϵ_{jkt} captures the stochastic error component. To facilitate estimation, the equation is transformed into a log-

[2] Ortega and Peri (2013) allow the deterministic component of utility to vary across individuals, based on the assumption that stayers and migrants differ in unobserved ways not captured by the model. While this relaxes a key assumption, it does not fundamentally alter the model's estimation. Multilateral resistance to migration can still be accounted for using origin-time fixed effects, although the interpretation of the coefficients changes.

In contrast, Bertoli and Fernández-Huertas Moraga (2013) introduce variation in destination attractiveness across origin countries by grouping destinations into distinct nests (e.g., English-speaking vs. non-English-speaking, developed vs. developing countries). This approach causes multilateral resistance to migration to vary by destination as well, leading to a more substantial shift in the model's structure and estimation. To address this, the authors employ the Common Correlated Effects (CCE) estimator developed by Pesaran (2006), which accounts for unobserved common factors. However, a key limitation of this method is its reliance on high-frequency data.

[3] Harmonic mean is used in the literature because empirically, distance elasticity is found as -1. While this value is added to the population-weighted distance equation, it transforms into harmonic mean (Conte et al., 2023).

linear form:

$$ln\,(m_{jkt}) = ln\beta_0 + \beta_1 ln\,(Pop_{kt}) + \beta_2 ln\,(Pop_{jt}) + \beta_3 ln\,(Dist_{jkt}) + ln\,(u_{jkt}) \quad (3.7)$$

To address the issue of zero values in certain variables, a value of one is added prior to applying the logarithmic transformation. The same approach is applied to variables introduced later in the analysis. However, the dependent variables are not log-transformed, in accordance with the estimation strategy described below.

While this simple model has yielded valuable insights in both trade and migration literature, enhancing the analysis by incorporating additional variables is essential. The basic equation overlooks several important factors influencing individuals' decisions to pursue education abroad. To account for these factors, new variables are introduced into the model. The extended equation can be specified as follows:

$$\begin{aligned} ln(m_{jkt}) = {} & ln\delta_0 + \delta_1 ln(Pop_{kt}) + \delta_2 ln(Pop_{jt}) + \delta_3 ln(Dist_{jkt}) + \delta_4 ln(GDP_{jkt}) \\ & + \delta_5(Uni_{kt}) + \delta_6(Uni_{jt}) + \delta_7(YouthUnemp_{kt}) + \delta_8(YouthUnemp_{jt}) \\ & + \delta_9 ln(Migrants_{jk}) + \delta_{10}(Contig_{jk}) + \delta_{11}(Col_{jk}) + \delta_{12}(Lang_{jk}) \\ & + \delta_{13}(Eng_k) + \delta_{14}(Eng_j) + ln(v_{jkt}) \end{aligned} \quad (3.8)$$

The variables Pop_{kt}, Pop_{jt}, $Dist_{jkt}$, and ϵ_{jkt} are retained from the traditional specification as introduced above. A key addition in this model is the variable GDP_{jkt}, which captures the difference in GDP per capita (in constant 2015 US dollars) between the destination and origin countries. To address the issue of taking logarithms of non-positive values in the difference in GDP per capita, the inverse hyperbolic sine (IHS) transformation is applied. The variable Uni represents the number of top-ranked universities per per 100,000 people aged 15–29, indexed by destination or origin as appropriate. Similarly, $YouthUnemp$ denotes the youth unemployment rate in tens of percent for the destination and origin countries, again depending on the index. $Migrants_{jk}$ measures the stock of migrants from origin j residing in destination k, expressed in 1,000 people, based on data from 2010 and 2011. $Contig_{jk}$ is a binary variable indicating whether countries j and k share a common border. $Lang_{jk}$ equals 1 if the two countries share an official language, and

Col_{jk} indicates whether a colonial relationship existed between them after 1945. Additionally, two dummy variables, indexed by Eng, are included to control for whether the destination or origin country is English-speaking.

Although the traditional gravity model of trade had a significant influence on the literature, it omitted a crucial element: the multilateral resistance term. This term intuitively captures the overall trade environment encountered by both exporters and importers. Trade between any two countries can be influenced by this broader context. For instance, if either the exporter or importer has numerous alternative trade partners, the volume of trade between them may decrease—and vice versa. Anderson and van Wincoop (2003) formally introduced this concept and demonstrated its analytical importance within the model. Drawing from the trade literature, migration studies have also incorporated the multilateral resistance term into their frameworks. Building on Equation 3.6, a new theoretically grounded migration model can be developed. Once log-linearized, the equation can be estimated as follows:

$$\begin{aligned} ln(m_{jkt}) = ln\theta_0 + \theta_1 ln(GDP_{jkt}) + \theta_2 ln(Pop_{jt}) + \theta_3 ln(Dist_{jkt}) + \theta_4 ln(\Omega_{jt}) \\ + ln(\xi_{jkt}) \end{aligned} \tag{3.9}$$

Estimating the effect of multilateral resistance to migration is a central issue in the migration literature. The most common and direct approach to addressing this problem is the use of origin-time fixed effects. By incorporating the origin-time dummy variable, γ_{jt}, into the equation, the estimation takes the following form:

$$\begin{aligned} ln(m_{jkt}) = ln\omega_0 + \omega_1 ln(GDP_{jkt}) + \omega_2 ln(Pop_{jt}) + \omega_3 ln(Dist_{jkt}) + \gamma_{jt} \\ + ln(\zeta_{jkt}) \end{aligned} \tag{3.10}$$

However, the inclusion of origin-time fixed effects presents a significant limitation: it prevents the identification of the effects of origin-specific variables. To overcome this issue, alternative methodologies have been proposed in both the trade and migration literatures that allow for the control of multilateral resistance while retaining the origin-time dimension. Following the approach developed by Baier and Bergstrand (2009) and later implemented in the migration context by Czaika and Parsons (2017), this study constructs a proxy for the multilateral resistance to migration term, which is calculated as follows:

$$\Omega_{jt} = \sum_{k=1}^{N} \frac{share_{kt}}{ln\,(Dist_{jkt})} \tag{3.11}$$

In this equation, $share_{kt}$ denotes the ratio of destination k's GDP per capita to the world GDP per capita (both measured in constant 2015 US dollars) at time t. By multiplying this term by the inverse of the logarithm of the distance between origin and destination, and summing across all destinations, an origin-time specific proxy variable for multilateral resistance is obtained. This method allows for the estimation of Equation 3.9 while preserving the origin-time dimension of the dataset.

This basic theoretical model can be extended following the same approach as the traditional specification. By introducing additional explanatory variables to Equations 3.9 and 3.10, we obtain the following modified expressions:

$$\begin{aligned} ln(m_{jkt}) = {} & ln\eta_0 + \eta_1 ln(Pop_{kt}) + \eta_2 ln(Dist_{jkt}) + \eta_3 ln(GDP_{jkt}) + \eta_4(Uni_{kt}) \\ & + \eta_5(YouthUnemp_{kt}) + \eta_6 ln(Migrants_{jk}) + \eta_7(Contig_{jk}) \\ & + \eta_8(Col_{jk}) + \eta_9(Lang_{jk}) + \eta_{10}(Eng_k) + \gamma_{jt} + ln(\kappa_{jkt}) \end{aligned} \tag{3.12}$$

$$\begin{aligned} ln(m_{jkt}) = {} & ln\lambda_0 + \lambda_1 ln(Pop_{kt}) + \lambda_2 ln(Pop_{jt}) + \lambda_3 ln(Dist_{jkt}) + \lambda_4 ln(GDP_{jkt}) \\ & + \lambda_5(Uni_{kt}) + \lambda_6(Uni_{jt}) + \lambda_7(YouthUnemp_{kt}) + \lambda_8(YouthUnemp_{jt}) \\ & + \lambda_9 ln(Migrants_{jk}) + \lambda_{10}(Contig_{jk}) + \lambda_{11}(Col_{jk}) + \lambda_{12}(Lang_{jk}) \\ & + \lambda_{13}(Eng_k) + \lambda_{14}(Eng_j) + \lambda_{15} ln(\Omega_{jt}) + ln(\psi_{jkt}) \end{aligned} \tag{3.13}$$

3.3 Estimation Strategy

The estimation strategy for international student mobility data requires special attention due to the structure of the dataset. The first challenge is the non-normal distribution of the dependent variable. By definition, the number of students migrating from one country to another cannot be negative. In addition, given the substantial number of small countries included in the dataset, it is expected that a considerable portion of the dependent variable will be zero. As a result, the distribution of student flows is confined to non-negative values, with a high concentration at zero that gradually declines as values increase. The second challenge concerns the error term. Since the model employs a log-linearized specification, the log-transformation of the error term requires statistical independence for consistent coefficient estimation. However, this assumption is difficult to satisfy in practice due to the heteroskedastic nature of the data.

Silva and Tenreyro (2006) addressed similar challenges in the context of international trade data, where the dependent variable is non-negative and often characterized by a large number of zeros. They compared various estimation methods and concluded that the Poisson Pseudo-Maximum Likelihood (PPML) estimator is the most suitable in the presence of heteroskedasticity and non-normality. As a Poisson-based approach that does not require the dependent variable to follow a Poisson distribution, PPML operates under weaker distributional assumptions, making it compatible with the characteristics of international student mobility data. Crucially, PPML allows the conditional variance of the dependent variable to vary with its conditional mean. This is particularly important in this context: when the expected number of students approaches zero, the possible range of outcomes becomes narrower due to the non-negativity constraint; conversely, when the expected value increases, the potential range widens. A robust estimator must account for this heteroskedasticity in the data-generating process. Given these advantages and the structural similarity between trade and international student mobility data, PPML has become the most widely used estimation technique in recent studies of student flows (e.g., Beine et al. (2014), Abbott and Silles (2016), Didisse et al. (2019), Kaushal and Lanati (2019)). The same estimation strategy is adopted in this study as well.

Chapter 4

DATA

In this study, we use the data about international students provided by the OECD. The OECD records international student mobility within OECD member countries. In particular, the records show the total number of international students at the origin-destination level for each OECD country as a destination. Starting in 2013, the data encompass detailed records categorizing students by gender (male, female, and a combined total), country of origin (inclusive of global nations), destination (limited to OECD countries), educational level (segmented into seven distinct tiers, including three for short-cycle tertiary education, bachelor's, master's, doctoral levels, and an aggregate tertiary level). We focus on the years 2013 to 2020 in this study. This comprehensive dataset enables a nuanced analysis of trends and patterns in international student flows, offering invaluable insights for educational policy and research.

In the initial stage of data cleansing, entries pertaining to students at short-cycle tertiary education levels were excluded due to a higher incidence of incomplete data. Additionally, data from non-country origins (i.e., broader world regions, such as Europe, Asia and Northern America), with the sole exception of "World", were removed to align with the research focus. The "World" category reports data on international students of a specific gender, attending a certain education level, during a particular year from all origins. Lastly, observations whose origins are the same with the destinations were removed from the dataset.

To rectify issues of unbalanced panel data, missing observations were systematically added. If the origin was "World" (in particular, for six observations of destination Switzerland) for these newly included years, the total count of international students was derived by summing up the data across all origins, while maintaining consistency in destination, education level, year, and gender. Moreover, one negative observation is replaced with zero based on the comparison between observations at the same destination-origin-year-level intersection but in remaining two sex categories.

During the second phase of data cleaning, the focus shifted to evaluating the quality and completeness of data for destination countries. This step resulted in the exclusion of countries that failed to meet predefined criteria related to data availability. The primary exclusion criterion was the number of missing years, defined as the absence of origin-specific data for any level of study. Countries that lacked such data for more than two years were excluded from the analysis. As a result, Colombia, Costa Rica, Greece, Mexico, the Slovak Republic, and Slovenia were removed from the dataset.[1] However, Germany and Spain, despite exhibiting similar gaps, were retained due to their large volumes of incoming international students, which were deemed critical for maintaining representativeness. For Germany, where doctoral level origin data was missing for all years except 2020, the missing data was estimated based on the origin distribution from 2020. For Spain, the data for each origin and the aggregate total were missing for the doctoral level in 2014 and 2015, and they were estimated using linear interpolation between 2013 and 2016. Conversely, the United States was excluded because its reported data aggregates master's and doctoral students under a single graduate-level category. As there was no reliable method to disaggregate these figures, the U.S. data could not be incorporated meaningfully into the level-specific analysis and was therefore omitted.

The second criterion in the second phase of data cleaning focused on the quality of origin data for each destination country. This was assessed by comparing two figures: the total number of international students of a particular gender, education level, and year from all origins to a specific destination, and the corresponding number from the "World" origin. Ideally, these figures should match but there are discrepancies. So, the difference as a percentage of the "World" total was calculated.

[1]Missing years for countries: 2013-2015 at master's and doctoral level for Colombia, 2013-2018 at all levels for Costa Rica, 2013-2014 at master's and doctoral level for Greece, 2013-2016 at bachelor's, master's and doctoral level for Mexico, 2013-2015 at master's level for Slovak Republic, all years at all levels for Slovenia

A threshold of 15% was set, and countries exceeding this difference for any level of interest were removed. Belgium, Finland, Israel, Lithuania, Netherlands, and Sweden were excluded based on this criterion.

In the final phase, specific issues in the dataset were addressed. For the remaining destination OECD countries, a few had missing data about origin distribution in certain levels of study only in a single year. From the available information in the missing year and the adjacent years, we infer these missing values. There are five countries with such inferred data: Brazil, Türkiye, Norway, Ireland, and Luxembourg. In the cases of Brazil and Türkiye, the dataset lacked observations for all origins at the master's and doctoral levels in 2013, but the aggregate total by level and gender in 2013 (i.e., the origin "World") was available. Then, we infer the missing origin-specific data by multiplying the number of international students from the "World" origin for each gender in 2013 by the proportion of each origin's contribution to the 2014 aggregate at each level and gender, separately. A similar method was applied for Norway, where data for all origins except the aggregate total (i.e., the origin "World") were quite inconsistent at the bachelor's and master's levels in 2019 with those in the adjacent years. The data for 2019 in Norway were adjusted. Here, the adjustment involved multiplying the "World" origin numbers for each gender in 2019 by the average proportion of each origin in the adjacent years (i.e., 2018 and 2020).

For Ireland, the data for 2013 lacked sex-specific details despite totals of two sexes were available per origin. We make the inference by using the sex ratio of students from other years. Sex ratios for each origin were calculated using the nearest year with available data (i.e., the nearest year where the total number of students coming from this particular origin is non-zero), then applied to the 2013 total numbers per origin. If the total for all sexes was zero in all years except 2013, a 1:1 sex ratio was assumed for that origin and level. Luxembourg's 2014 data, showing similar issues, was addressed using a comparable approach. The average sex ratio for that origin from neighboring years (i.e., from 2013 and 2015) was used if available; otherwise, the ratio from the closest year with positive number of incoming students was used for estimation. If no such year existed other than 2014, a 1:1 sex ratio assumption

Table 4.1: Discrepancy ratios (in percentage)

Destination Countries	Bachelor's	Master's	Doctoral	Tertiary
Panel A: Countries with a discrepancy ratio below 15%				
Australia	1.36	1.15	1.71	1.29
Austria	4.6	5.75	6.94	5.28
Brazil	0.94	0.99	0.88	0.94
Canada	1.89	2.38	1.19	1.91
Chile	9.83	5.29	3.49	7.06
Czech Republic	0.76	0.57	2.33	0.82
Denmark	0.24	0.35	0.72	0.35
Estonia	0.19	0.76	1.56	0.58
France	5.74	6.33	1.99	5.62
Germany	14.61	4.49	8.04	8.65
Hungary	8.95	4	6.71	6.54
Iceland	1.05	0.84	1.03	0.98
Ireland	2.37	6.7	5.38	4.04
Italy	1.67	1.92	2.07	1.79
Japan	1.77	-1.11	6.64	1.68
Korea	1.77	0.8	0.43	1.36
Latvia	0.14	0.18	0.6	0.17
Luxembourg	0.13	0.61	0.77	0.52
New Zealand	1.92	0.01	0.13	1.46
Norway	5.1	2.61	1.67	3.46
Poland	1.42	1.53	2.62	1.49
Portugal	0.08	0.23	0.39	0.21
Spain	0.11	0.25	0.4	0.24
Switzerland	0.51	13.72	0.87	5.51
Türkiye	3.7	3.45	3.7	3.64

Continued on next page

Table 4.1 – continued from previous page

Destination Countries	Bachelor's	Master's	Doctoral	Tertiary
United Kingdom	1.11	1.03	0.95	1.06
United States	1.83	2.73		2.22
Panel B: Countries with a discrepancy ratio above 15%				
Belgium	23.82	31.61	47.27	29.16
Finland	23.77	14.76	14.35	19.12
Israel	15.48	17.71	32.01	17.21
Lithuania	0.37	28.28	54.38	13.92
Netherlands	23.81	21.58	10.93	22.30
Sweden	30.76	22.97	10.01	21.55

Notes: OECD data. The table shows the discrepancy ratios in percentages. The discrepancy ratios were derived by subtracting the number of international students from all origins from those reported under the "World" origin, and then dividing this difference by the number of students coming from the "World" origin.

was applied for each origin and level.

Besides problems specific to destinations, there were some problems related to specific origins in OECD data. First, Kosovo and Serbia were excluded because these two countries separated within our period of analysis which makes data for both origins unreliable. Another origin-specific problem was that data related to some origins were written in another origin's data for some destinations. These countries have been detected by using flag codes in the dataset. These flag codes were explanatory notes about why these specific observations were missing. Two related flag codes were "data included in another category" and "includes data from another category". If an origin was labeled as "includes data from another category" while another one was labeled as "data included in another category" for the same destination, education level, year and sex, then it was considered that the latter observation was included in the former. Such origins were Macao, Hong Kong and Democratic People's Republic of Korea. For Japan, data for Hong Kong and Macao were included in China's data, and data for the Democratic People's

Republic of Korea were included in the Republic of Korea's data. To rectify this, the distribution of these origins in a country with complete and accurate data for these regions was used as a reference. Australia, having full data and geographical proximity to Japan, served as this reference. The number of students from Hong Kong and Macao to Japan was inferred and the number of students from China to Japan was recalculated by using their proportion in Australia's data relative to China. Similarly, students from the Democratic People's Republic of Korea to Japan were estimated and students from the Republic of Korea were recalculated by using Australia's data as the reference. For France, where the same issue occurred with Hong Kong and Macao origins in 2013 and 2014, Switzerland's comprehensive data was utilized in a similar manner to correct the discrepancies because it was the nearest country with full data about these origins.

Upon completion of these data cleaning and estimation procedures, the OECD dataset was finalized by calculating the tertiary education level for each unique combination of destination, origin, and year. This was accomplished by summing the data across all other educational levels (i.e., bachelor's, master's and doctorates), thereby providing a comprehensive and accurate representation of international student mobility in the context of tertiary education within OECD countries.

In addition to OECD data, other sources were utilized to obtain information on covariates. One key source is the CEPII dataset (Centre d'Études Prospectives et d'Informations Internationales, or the Centre for Prospective Studies and International Information). Widely used in trade literature, this dataset provided relevant variables for our analysis. These include various measures of distance—such as the simple distance between capital cities, the simple distance between the most populated cities, and population-weighted distance (both arithmetic and harmonic mean). Additionally, the dataset captures linguistic ties (common official language), colonial relationships (whether two countries had a colonial relationship after 1945), and shared borders.

The World Bank database is utilized for macroeconomic indicators and population data. The macroeconomic indicators include unemployment (total and youth separately) and GDP (including PPP, per capita, and per capita PPP) for both destination and origin countries. The dataset provides two estimates for total and youth unemployment: national and ILO estimates. Our analysis primarily uses national estimates, but missing values are supplemented with ILO estimates due to their strong correlation. Lastly, population data cover total and gender-specific populations, both in aggregate and across specific age groups: 15-19, 20-24, and 25-29.

Data on top universities were sourced from the Times Higher Education (THE) World University Rankings. To maintain consistency across years, only the top 400 universities were considered, as this represents the maximum number of institutions consistently reported throughout the study period. For each destination and origin country, the number of universities appearing in this top 400 list was calculated on an annual basis.

Lastly, data on migrant stock from origin countries residing in destination countries were sourced from the Database on Immigrants in OECD and non-OECD Countries (DIOC). This dataset provides origin–destination-specific figures for the reference year 2010/11—the only version available that aligns with the temporal scope of this study. Consequently, the migrant stock variable is treated as time-invariant. Missing values, which are likely to reflect negligible or zero migrant flows, were replaced with zeros. Furthermore, two destination countries—Brazil and South Korea—were entirely absent from this dataset. To maintain consistency in the sample across different model specifications, these countries were excluded from the analysis.

In total, 24 OECD countries remained in the dataset as destination countries. These are: Australia, Austria, Canada, Chile, Czech Republic, Denmark, Estonia, France, Germany, Hungary, Iceland, Ireland, Italy, Japan, Latvia, Luxembourg, New Zealand, Norway, Poland, Portugal, Spain, Switzerland, Türkiye, and the United Kingdom. The analysis sample was further refined by excluding origin countries[2]

[2]Excluded origins: Andorra, Anguilla, Antigua and Barbuda, Aruba, Bermuda, British Virgin Islands, Cayman Islands, Cook Islands, Cuba, Curacao, Dominica, Eritrea, Gibraltar, Grenada, Holy See, Kiribati, Liechtenstein, Marshall Islands, Micronesia, Monaco, Montserrat, Nauru, Niue, North Korea, Palau, Saint Kitts and Nevis, San Marino, Seychelles, Sint Marteen, South Sudan,

with insufficient data on key macroeconomic indicators. After these exclusions, the final dataset comprised 174 origin countries for each destination.

4.1 Descriptive Statistics

Summary statistics for the dataset used in the analysis are presented in Table 4.2. The mean number of international students per origin-destination-year observation is approximately 243, while the median is only 5, and the maximum value reaches 138,850. This stark difference highlights the highly skewed distribution of the dependent variable, with a large concentration of small values. A similar pattern is observed in the migrant network size variable, which also exhibits a right-skewed distribution.

In contrast, the population and distance variables display relatively symmetrical distributions, as indicated by the similarity between their mean and median values. The GDP per capita difference (destination minus origin) is highly left-skewed, reflecting the fact that, in most cases, destination countries have significantly higher GDP per capita than origin countries—although there are some exceptions where the origin has a higher GDP. This skewed pattern is also visible in the university quality variables. On average, destination countries have five times more top-ranked universities than origin countries, and the median number of top universities for origin countries is zero. Regarding youth unemployment, the destination and origin variables have similar left-skewed distributions, with comparable means and medians. However, the range is wider for origin countries, suggesting more variation in youth unemployment levels among origins. For the dummy (binary) variables, only the proportions can be interpreted. Approximately 2% of origin-destination pairs share a border, and a similar proportion have a colonial relationship after 1945. Common official language is shared in about 10% of cases. Around 27% of origin countries and 21% of destination countries have English as an official language.

As illustrated in Figure 4.1, the number of international students has steadily increased over the study period (2013–2020). The total number of internationally

Syria, Tokelau, Turks and Caicos Islands, Tuvalu, Venezuela, Yemen

Table 4.2: Summary statistics

Statistics	Mean	Standard Deviation	Minimum	Median	Maximum
Students	243	1977.564	0	5	138850
Year	2016.5	2.291	2013	2016.5	2020
Education level	3.5	1.118	2	3.5	5
(L) Distance	8.601	0.861	4.007	8.828	9.882
(L) Population (d)	7.727	1.555	4.243	7.449	9.916
(L) Population (o)	7.616	1.805	3.166	7.586	12.832
(I) GDP per capita difference	7.845	7.41	-12.177	10.852	12.284
Top universities (d)	0.302	0.379	0	0.203	2.566
Top universities (o)	0.062	0.194	0	0	2.566
Youth unemployment (d)	1.637	0.938	0.367	1.337	5.548
Youth unemployment (o)	1.755	1.3	0.029	1.418	8.186
(L) Migrant network	0.809	1.286	0	0.139	7.177
(L) Resistance term	2.471	0.071	2.312	2.451	2.688
Shared border	0.019	0.135	0	0	1
Colonial tie	0.021	0.142	0	0	1
Common language	0.101	0.302	0	0	1
English language (d)	0.208	0.406	0	0	1
English language (o)	0.276	0.447	0	0	1

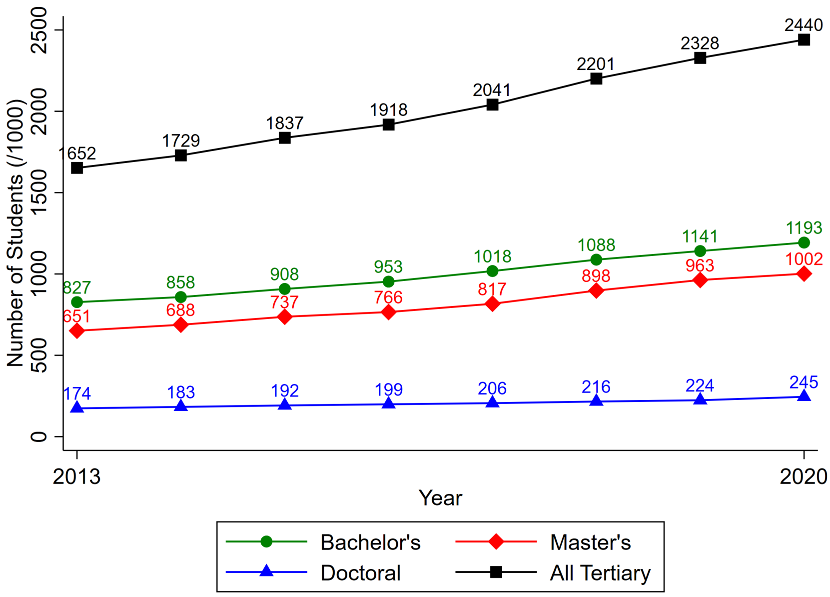
Continued on next page

Table 4.2 – continued from previous page

Notes: OECD data for 2013–2020. This table presents summary statistics for the variables used in the analysis of international student migration, with each row representing a variable and each column a specific statistical measure. All statistics are based on the transformed versions of the variables used in the empirical analysis, with 132,864 observations for each variable. Variables ending in lowercase d and o denote destination- and origin-specific measures, respectively; an uppercase L at the start indicates a log transformation, and an uppercase I indicates an inverse hyperbolic sine transformation. Distance refers to the population-weighted harmonic mean distance between the most populated cities, and population variable measures individuals aged 15–29 in hundreds of thousands. GDP per capita difference between destination and origin is expressed in constant 2015 USD. The "top universities" variable counts institutions ranked in the top 400 of the Times Higher Education World University Rankings per 100,000 people aged 15–29. Youth unemployment is reported in tens of percent, migrant network size in thousands, and the last five variables are dummies. Colonial tie refers only to ties established after 1945, and language denotes the official language.

mobile tertiary students rose from 1.65 million in 2013 to 2.44 million in 2020, marking a 48 percent increase. Throughout this period, bachelor's students constituted the largest group, growing from 827 thousand to 1.193 million—an increase of 44 percent. In contrast, master's students recorded the highest relative growth, rising by approximately 54 percent, from 651,000 to just over 1 million. Even the smallest group—doctoral students—experienced notable growth, increasing by 41 percent from 174,000 to 245,000. As a result of these differing growth rates, the share of master's students among the total tertiary student population increased, while the shares of other levels declined. These variations suggest distinct trends in international student mobility by education level. To fully understand these patterns, disaggregated analyses by degree level are necessary.

To gain a clearer understanding of international student mobility by education level, the top 20 migration corridors for each degree level are presented in Tables 4.3, 4.4, 4.5, and 4.6. A consistent pattern emerges across all levels: China ranks as the leading country of origin, occupying multiple top spots in each table. Specifically, four of the top five migration corridors at every level involve Chinese students. Australia, the United Kingdom, Japan, and Germany are among the most common



(a)

Figure 4.1: Trends of education levels during the period of study

Notes: OECD data. This figure illustrates the global number of internationally mobile students enrolled in tertiary education programs from 2013 to 2020, disaggregated by degree level. The values on the y-axis represent the number of students in thousands.

destinations for Chinese students overall. Notably, France and Germany become more prominent destinations at the graduate level, particularly for master's and doctoral students, but are less represented among the top destinations for Chinese undergraduate students. Conversely, New Zealand appears as a popular destination for Chinese undergraduates but not for graduate-level students. These differences may reflect distinct preferences or decision-making criteria between undergraduate and graduate students when evaluating destination country characteristics.

In contrast to migration corridors that are consistently popular across all education levels, some corridors are specific to particular levels of study. For instance, certain origins appear among the top corridors for undergraduate students but are absent from the graduate-level lists. These origins of popular corridors include Hong Kong, Malaysia, Azerbaijan, Turkmenistan, and South Korea. Conversely, other countries—such as Algeria, the United States, Nigeria, Russia, and Tunisia—emerge

Table 4.3: Bachelor's students' migration corridors

Destination	Origin	Yearly Average
Australia	China	50517
Canada	China	45391
United Kingdom	China	40619
Japan	China	36639
Poland	Ukraine	14540
Austria	Germany	14263
United Kingdom	Hong Kong	12958
Czech Republic	Slovakia	11734
United Kingdom	Malaysia	11599
Australia	India	11042
Turkey	Azerbaijan	10056
New Zealand	China	10029
Australia	Malaysia	9955
Canada	France	8852
Turkey	Turkmenistan	8774
Australia	Nepal	8658
France	Morocco	8449
Canada	India	8342
Japan	South Korea	8082
Germany	China	8027

Notes: OECD data for the 2013-2020 period. This table presents the top international migration corridors for bachelor's-level students, ranked by the average annual number of migrating students over the sample period.

as prominent origins at the graduate level but are not represented among the top undergraduate corridors. Similarly, some destination-origin pairs stand out uniquely, appearing in the top rankings without any comparable corridor involving a different partner country. Examples include Poland–Ukraine, Austria–Germany, Czech Re-

Table 4.4: Master's students' migration corridors

Destination	Origin	Yearly Average
United Kingdom	China	52882
Australia	China	48175
Australia	India	32597
Japan	China	19179
Germany	China	18792
France	Morocco	16839
United Kingdom	India	14709
France	China	12860
Germany	India	12537
Austria	Germany	11692
France	Algeria	11493
Poland	Ukraine	9165
Czech Republic	Slovakia	9110
Australia	Nepal	8322
Canada	China	7793
United Kingdom	United States of America	7257
United Kingdom	Nigeria	6732
Canada	India	5705
Germany	Russia	5666
France	Tunisia	5354

Notes: OECD data for the 2013-2020 period. This table presents the top international migration corridors for master's-level students, ranked by the average annual number of migrating students over the sample period.

public–Slovakia, and Australia–Nepal. These patterns underscore the importance of considering both origin-specific factors and destination-origin dynamics when analyzing international student mobility.

Another notable indicator presented in the migration corridors tables is the

Table 4.5: Doctoral students' migration corridors

Destination	Origin	Yearly Average
Japan	China	6314
United Kingdom	China	6295
Australia	China	3764
Switzerland	Germany	3624
Germany	China	3189
Canada	China	2900
United Kingdom	United States of America	2719
Portugal	Brazil	2632
United Kingdom	Saudi Arabia	2602
United Kingdom	Italy	2548
France	China	2508
Canada	Iran	2410
United Kingdom	Germany	2328
Austria	Germany	2122
United Kingdom	Nigeria	2026
France	Italy	1855
Czech Republic	Slovakia	1788
France	Tunisia	1781
Turkey	Iran	1641
France	Algeria	1631

Notes: OECD data for the 2013-2020 period. This table presents the top international migration corridors for doctoral level students, ranked by the average annual number of migrating students over the sample period.

concentration of student flows. Across all education levels, the top two or three migration corridors—including those originating from China—account for a disproportionately large share of total student mobility. A closer look reveals significant variation in the distribution of flows beyond the top-ranked corridors. For instance,

Table 4.6: All tertiary students' migration corridors

Destination	Origin	Yearly Average
Australia	China	102457
United Kingdom	China	99797
Japan	China	62132
Canada	China	56084
Australia	India	44822
Germany	China	30008
Austria	Germany	28076
France	Morocco	26399
Poland	Ukraine	23950
United Kingdom	India	23611
Czech Republic	Slovakia	22632
France	China	21672
France	Algeria	19158
Australia	Nepal	17167
United Kingdom	United States of America	15822
United Kingdom	Hong Kong	15310
Canada	India	15051
United Kingdom	Malaysia	14987
Germany	India	14477
United Kingdom	Nigeria	14260

Notes: OECD data for the 2013-2020 period. This table presents the top international migration corridors for all tertiary students, ranked by the average annual number of migrating students over the sample period.

the number of master's students in the 20th-ranked corridor is only 10 percent of that in the top corridor. In contrast, the same ratio is 16 percent for bachelor's students and 26 percent for doctoral students. This suggests that international student mobility at the doctoral level is more evenly distributed across corridors than at the master's level, where student flows are more heavily concentrated in a few dominant routes. These differences highlight the need to analyze international student mobility separately by education level, as the underlying patterns and drivers may differ substantially.

Chapter 5

RESULTS

5.1 Model Fit

As discussed in the previous section, this study analyzes international student flows using several different models, with the results presented in Tables 5.1, 5.2, 5.3, and 5.4. When comparing the explanatory power of these models based on their R-squared values, the theoretical model with fixed effects emerges as the best. This model consistently yields higher R-squared scores than the others across both the basic and extended specifications, regardless of the education level. For instance, in the basic and extended models with fixed effects for the bachelor's level, the R-squared values are 0.264 and 0.734, respectively. They are 0.171 and 0.471 in the corresponding models in traditional gravity specification. This outcome is expected, as the inclusion of origin-time fixed effects allows the model to control for many unobserved factors that could otherwise bias the results.

On the other hand, the explanatory power differs between the basic and extended versions of the remaining models. The basic traditional model yields higher R-squared values (e.g., 0.171 at the bachelor's level) than the basic theoretical model that incorporates the multilateral resistance term (e.g., 0.116 at the bachelor's level) for all education levels. However, this pattern reverses in the extended versions: the extended theoretical model with the multilateral resistance term outperforms the extended traditional model in explaining student mobility for two out of the three education levels (0.471 versus 0.486 at the bachelor's level). This difference suggests that the youth population in the country of origin has greater explanatory power over international student mobility than the combined influence of the GDP per capita gap and the multilateral resistance term. However, the multilateral resistance term still makes a meaningful contribution to the model, particularly when other

Table 5.1: Basic models

Panel A: Basic Traditional Model				
	Bachelor's	Master's	Doctorate	All
(L) Distance	-0.827***	-0.882***	-0.805***	-0.846***
	(0.0280)	(0.0261)	(0.0258)	(0.0242)
(L) Population (d)	0.689***	0.704***	0.631***	0.689***
	(0.0198)	(0.0216)	(0.0166)	(0.0184)
(L) Population (o)	0.717***	0.865***	0.755***	0.781***
	(0.0278)	(0.0273)	(0.0151)	(0.0244)
Constant	0.414	-0.917**	-1.202***	0.654*
	(0.372)	(0.414)	(0.306)	(0.360)
Number of observations	33,216	33,216	33,216	33,216
R-squared	0.171	0.163	0.284	0.203
Origin-time fixed effects	NO	NO	NO	NO
Panel B: Basic Theoretical Model with MRM Term				
	Bachelor's	Master's	Doctorate	All
(L) Distance	-1.021***	-0.955***	-0.805***	-0.973***
	(0.0597)	(0.0411)	(0.0502)	(0.0474)
(L) Population (o)	0.708***	0.877***	0.763***	0.781***
	(0.0312)	(0.0310)	(0.0179)	(0.0280)
(I) GDP per capita difference	0.0107***	0.0120***	0.0213***	0.0121***
	(0.00269)	(0.00288)	(0.00353)	(0.00254)
(L) Resistance term	-3.431***	-1.029*	0.355	-2.134***
	(0.776)	(0.540)	(0.667)	(0.612)
Constant	16.30***	7.935***	2.925	12.68***
	(2.464)	(1.817)	(2.027)	(1.970)
Number of observations	33,216	33,216	33,216	33,216
R-squared	0.116	0.117	0.180	0.138

Continued on next page

Table 5.1 – continued from previous page

	Bachelor's	Master's	Doctorate	All
Origin-time fixed effects	NO	NO	NO	NO
Panel C: Basic Theoretical Model with Fixed Effects				
	Bachelor's	Master's	Doctorate	All
(L) Distance	-0.907***	-0.850***	-0.670***	-0.861***
	(0.0425)	(0.0339)	(0.0392)	(0.0339)
(I) GDP per capita difference	0.0442***	0.0347***	0.0585***	0.0421***
	(0.00289)	(0.00279)	(0.00304)	(0.00261)
Constant	13.87***	13.40***	10.12***	14.20***
	(0.341)	(0.267)	(0.327)	(0.272)
Observations	33,216	33,216	33,192	33,216
R-squared	0.264	0.187	0.311	0.255
Origin-Time FE	YES	YES	YES	YES

Notes: Notes: The table reports regression estimates for different education levels using three basic model specifications. Standard errors are shown in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Variables ending in lowercase d and o denote destination- and origin-specific measures, respectively. A leading uppercase L indicates a log transformation, while an uppercase I denotes an inverse hyperbolic sine transformation. Distance is the population-weighted harmonic mean distance between the largest cities. The population variable represents individuals aged 15–29, measured in units of 100,000. The GDP per capita difference between destination and origin is expressed in constant 2015 USD.

dimensions of the phenomenon are captured through additional control variables in the extended specifications. Although the coefficient of the multilateral resistance term is difficult to interpret directly, its sign is theoretically consistent—negative and statistically significant in most cases—except for the basic theoretical model applied to doctorate students.

When evaluating the overall explanatory power of the models across different education levels, doctoral student mobility emerges as the most accurately captured. Regardless of model type—basic or extended—the highest R-squared values are consistently observed for doctorate-level student flows. Bachelor's students follow,

Table 5.2: Extended traditional model

	Bachelor's	Master's	Doctorate	All
(L) Distance	-0.437***	-0.361***	-0.318***	-0.391***
	(0.0307)	(0.0277)	(0.0222)	(0.0245)
(L) Population (d)	0.478***	0.687***	0.601***	0.574***
	(0.0314)	(0.0369)	(0.0262)	(0.0303)
(L) Population (o)	0.470***	0.664***	0.588***	0.559***
	(0.0220)	(0.0264)	(0.0151)	(0.0213)
(I) GDP per capita difference	0.0126***	0.00320	0.0200***	0.00929**
	(0.00484)	(0.00384)	(0.00373)	(0.00385)
Top universities (d)	-0.460**	1.248***	0.865***	0.497***
	(0.228)	(0.145)	(0.110)	(0.167)
Top universities (o)	0.652***	0.471**	1.134***	0.641***
	(0.233)	(0.215)	(0.126)	(0.200)
Youth unemployment (d)	-0.278***	-0.162***	-0.173***	-0.215***
	(0.0247)	(0.0200)	(0.0175)	(0.0181)
Youth unemployment (o)	-0.118***	0.0722***	0.144***	-0.0127
	(0.0190)	(0.0209)	(0.0156)	(0.0172)
(L) Migrant network	0.455***	0.414***	0.391***	0.430***
	(0.0198)	(0.0155)	(0.0129)	(0.0155)
Shared border	0.410***	0.302***	0.322***	0.356***
	(0.101)	(0.0915)	(0.0769)	(0.0851)
Common language	0.137*	0.651***	0.733***	0.392***
	(0.0759)	(0.0870)	(0.0668)	(0.0692)
Colonial tie	0.460***	0.245**	0.0288	0.336***
	(0.111)	(0.120)	(0.0944)	(0.0991)
English language (d)	1.045***	0.0675	0.321***	0.555***
	(0.0616)	(0.0743)	(0.0462)	(0.0520)

Continued on next page

Table 5.2 – continued from previous page

	Bachelor's	Master's	Doctorate	All
English language (o)	-0.791***	-0.894***	-1.103***	-0.855***
	(0.0799)	(0.103)	(0.0601)	(0.0776)
Constant	0.426	-4.486***	-4.880***	-1.158**
	(0.456)	(0.599)	(0.366)	(0.485)
Observations	33,216	33,216	33,216	33,216
R-squared	0.471	0.407	0.561	0.504
Origin-Time FE	NO	NO	NO	NO

Notes: The table reports regression estimates for different education levels using three basic model specifications. Standard errors are shown in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Variables ending in lowercase d and o denote destination- and origin-specific measures, respectively. A leading uppercase L indicates a log transformation, while an uppercase I denotes an inverse hyperbolic sine transformation. Distance is the population-weighted harmonic mean distance between the largest cities. The population variable represents individuals aged 15–29, measured in units of 100,000. The GDP per capita difference between destination and origin is expressed in constant 2015 USD. The "top universities" variable counts institutions ranked in the top 400 of the Times Higher Education World University Rankings per 100,000 people aged 15–29. Youth unemployment is reported in tens of percent, migrant network size in thousands, and the last five variables are dummies. Colonial tie refers only to ties established after 1945, and language denotes the official language.

with their mobility generally better explained than that of master's students. In all models except for the basic theoretical model with the multilateral resistance term, bachelor's student flows exhibit higher R-squared scores than those of master's students. For example, in the extended theoretical model with fixed effects, the R-squared values are 0.734, 0.498, and 0.761 respectively at the bachelor's, master's and doctoral levels. The differences in explanatory power across education levels within the same model are generally notable, and these gaps tend to become even more pronounced in the extended versions.

Table 5.3: Extended theoretical model with MRM term

	Bachelor's	Master's	Doctorate	All
(L) Distance	-0.664***	-0.476***	-0.457***	-0.570***
	(0.0375)	(0.0387)	(0.0257)	(0.0317)
(L) Population (d)	0.453***	0.674***	0.584***	0.554***
	(0.0316)	(0.0369)	(0.0260)	(0.0300)
(L) Population (o)	0.423***	0.641***	0.561***	0.523***
	(0.0203)	(0.0261)	(0.0150)	(0.0206)
(I) GDP per capita difference	0.0116**	0.00463	0.0189***	0.00964**
	(0.00462)	(0.00379)	(0.00355)	(0.00375)
Top universities (d)	-0.347	1.262***	0.883***	0.552***
	(0.222)	(0.141)	(0.107)	(0.161)
Top universities (o)	0.932***	0.799***	1.339***	0.967***
	(0.168)	(0.183)	(0.0974)	(0.153)
Youth unemployment (d)	-0.264***	-0.155***	-0.174***	-0.206***
	(0.0242)	(0.0206)	(0.0172)	(0.0188)
Youth unemployment (o)	-0.0717***	0.0958***	0.170***	0.0233
	(0.0180)	(0.0200)	(0.0162)	(0.0166)
(L) Migrant network	0.485***	0.426***	0.407***	0.451***
	(0.0199)	(0.0161)	(0.0127)	(0.0155)
Shared border	0.257**	0.246**	0.246***	0.251***
	(0.105)	(0.0975)	(0.0804)	(0.0901)
Common language	0.104	0.624***	0.721***	0.362***
	(0.0802)	(0.0893)	(0.0670)	(0.0720)
Colonial tie	0.439***	0.253**	0.0155	0.330***
	(0.102)	(0.117)	(0.0894)	(0.0935)
English language (d)	1.080***	0.1000	0.368***	0.594***
	(0.0625)	(0.0749)	(0.0463)	(0.0529)

Continued on next page

Table 5.3 – continued from previous page

	Bachelor's	Master's	Doctorate	All
English language (o)	-0.995***	-0.965***	-1.218***	-0.992***
	(0.0818)	(0.107)	(0.0623)	(0.0801)
(L) Resistance term	-5.164***	-2.688***	-3.232***	-4.110***
	(0.438)	(0.484)	(0.328)	(0.399)
Constant	15.61***	3.373**	4.624***	10.90***
	(1.342)	(1.575)	(1.028)	(1.256)
Observations	33,216	33,216	33,216	33,216
R-squared	0.487	0.396	0.579	0.505
Origin-Time FE	NO	NO	NO	NO

Notes: The table reports regression estimates for different education levels using three basic model specifications. Standard errors are shown in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Variables ending in lowercase d and o denote destination- and origin-specific measures, respectively. A leading uppercase L indicates a log transformation, while an uppercase I denotes an inverse hyperbolic sine transformation. Distance is the population-weighted harmonic mean distance between the largest cities. The population variable represents individuals aged 15–29, measured in units of 100,000. The GDP per capita difference between destination and origin is expressed in constant 2015 USD. The "top universities" variable counts institutions ranked in the top 400 of the Times Higher Education World University Rankings per 100,000 people aged 15–29. Youth unemployment is reported in tens of percent, migrant network size in thousands, and the last five variables are dummies. Colonial tie refers only to ties established after 1945, and language denotes the official language.

5.2 *Migration vs Human Capital Model*

Following the model comparison, it is useful to examine the distinct patterns associated with each education level. Before delving into the detailed estimation results, a general framework can be outlined as follows: the international mobility of bachelor's and doctoral students tends to align more closely with the migration model, whereas master's students exhibit a more balanced pattern that reflects elements of both the migration and human capital models.

The most critical variable for identifying the theoretical framework underlying student mobility is the quality of education in the country of origin. According to Rosenzweig (2006), the migration model predicts a positive relationship between the quality of education and the number of students studying abroad. The findings of this study provide support for the migration model at all education levels. The number of students going abroad increases with the number of top universities per youth in their home country. But the effect is stronger for the bachelor's and doctoral levels, with being particularly pronounced for doctoral students, likely because admission to doctoral programs requires a higher level of academic preparation. Students from countries with fewer high-quality universities may be at a disadvantage.

In contrast, master's students exhibit a more nuanced pattern. The effect of number of top universities per youth in the home country on their likelihood of studying abroad is the weakest among all education levels. This remains true despite that admission to master's programs generally requires higher qualifications than undergraduate programs. If both bachelor's and master's students followed the migration model to a similar extent, the effect of home country education quality would be expected to be stronger for master's students. Therefore, this finding suggests that master's students are less influenced by the migration model compared to those at other education levels. Instead, they appear to adopt a more balanced approach that reflects elements of both the migration and human capital models. The latter predicts a negative relationship between the quality of domestic education and the number of students pursuing studies abroad, as improved local education opportunities reduce the incentive to leave.

Another critical factor influencing student mobility patterns is the earnings gap between origin and destination countries, which is proxied in this study by differences in GDP per capita. According to the migration model proposed by Rosenzweig (2006), student flows are expected to increase as income disparities widen, since individuals from lower-income countries are more motivated to pursue education in wealthier destinations. The empirical findings of this study support this prediction: in all models, the GDP per capita gap is positively and significantly associated with

Table 5.4: Extended theoretical model with fixed effects

	Bachelor's	Master's	Doctorate	All
(L) Distance	-0.636***	-0.381***	-0.410***	-0.502***
	(0.0306)	(0.0422)	(0.0199)	(0.0308)
(L) Population (d)	0.506***	0.698***	0.596***	0.592***
	(0.0220)	(0.0320)	(0.0234)	(0.0223)
(I) GDP per capita difference	0.0236***	0.0133***	0.0513***	0.0220***
	(0.00369)	(0.00383)	(0.00305)	(0.00307)
Top universities (d)	-0.132	1.331***	0.578***	0.663***
	(0.150)	(0.115)	(0.103)	(0.117)
Youth unemployment (d)	-0.142***	-0.114***	-0.104***	-0.123***
	(0.0211)	(0.0239)	(0.0150)	(0.0193)
(L) Migrant network	0.414***	0.410***	0.372***	0.410***
	(0.0204)	(0.0191)	(0.0135)	(0.0165)
Shared border	0.389***	0.352***	0.00285	0.330***
	(0.0897)	(0.109)	(0.0699)	(0.0877)
Common language	0.761***	0.892***	1.133***	0.822***
	(0.0772)	(0.0930)	(0.0615)	(0.0695)
Colonial tie	0.411***	0.208	0.290***	0.317***
	(0.0892)	(0.151)	(0.0648)	(0.0991)
English language (d)	1.094***	0.0458	0.304***	0.550***
	(0.0416)	(0.0677)	(0.0351)	(0.0403)
Constant	6.025***	2.172***	1.698***	4.862***
	(0.293)	(0.455)	(0.274)	(0.322)
Observations	33,216	33,216	33,192	33,216
R-squared	0.734	0.498	0.761	0.669
Origin-Time FE	YES	YES	YES	YES

Continued on next page

Table 5.4 – continued from previous page

Notes: Notes: The table reports regression estimates for different education levels using three basic model specifications. Standard errors are shown in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Variables ending in lowercase d denote destination-specific measures. A leading uppercase L indicates a log transformation, while an uppercase I denotes an inverse hyperbolic sine transformation. Distance is the population-weighted harmonic mean distance between the largest cities. The population variable represents individuals aged 15–29, measured in units of 100,000. The GDP per capita difference between destination and origin is expressed in constant 2015 USD. The "top universities" variable counts institutions ranked in the top 400 of the Times Higher Education World University Rankings per 100,000 people aged 15–29. Youth unemployment is reported in tens of percent, migrant network size in thousands, and the last four variables are dummies. Colonial tie refers only to ties established after 1945, and language denotes the official language.

outbound student flows for both bachelor's and doctoral students, with the effect being stronger for doctoral students. This stronger relationship may be attributed to the fact that higher-income destination countries often entail greater living and tuition costs—barriers that doctoral students are more likely to overcome through scholarships and funding opportunities. In contrast, bachelor's students typically face more limited financial support, which may reduce the impact of income differentials on their mobility. As a result, the net positive effect of the GDP gap is more pronounced for doctoral students.

In contrast to standard migration models, the human capital framework does not yield clear predictions regarding the impact of cross-country earnings gaps on student mobility (Rosenzweig, 2006). This stems from its foundational assumption that individuals invariably return to their home countries to work after completing their education. Theoretically, this implies that income disparities between origin and destination countries should exert a weaker – if any – influence on mobility decisions, since students' expected returns depend primarily on home-country labor market conditions rather than those abroad. The empirical results for master's students align with this theoretical expectation. Except one model, the GDP per capita gap demonstrates a smaller positive effect on master's-level mobility compared to other educational tiers. Notably, this relationship becomes statistically insignificant in two of these specifications: the extended traditional model and the extended theoretical model with MRM term. This pattern suggests that master's students' mobility decisions reflect a hybrid rationale, incorporating both human capital accumulation and potential migration considerations, rather than being driven predominantly by income differentials.

The differing migration patterns across education levels are further supported by the influence of additional control variables, particularly the number of top universities in the destination country. This variable shows the strongest positive effect on the mobility of master's students across all models, highlighting the greater relevance of the human capital model at this level. For doctoral students, the effect is also positive and statistically significant, suggesting that the presence of high-quality institutions continues to play a meaningful role in shaping their mobility decisions. However, for bachelor's students, the effect is either negative and insignificant or, in one model, negative and significant. This outcome may reflect the higher tuition costs typically associated with top-ranked universities, which can be a substantial barrier for bachelor's students who generally have fewer opportunities to access financial aid or tuition waivers. In contrast, doctoral students are more likely to benefit from scholarships and institutional funding, making them better positioned to pursue education at prestigious institutions abroad.

A similar pattern emerges when examining the role of youth unemployment in the country of origin. Based on the theoretical frameworks discussed earlier, youth unemployment is expected to have a stronger push effect on bachelor's and doctoral students compared to master's students, given their more migration-oriented motivations. This expectation is confirmed for doctoral students; however, the effect for bachelor's students runs counter to the theoretical prediction. In their case, higher youth unemployment in the origin country is associated with a decrease in outbound mobility. This puzzling result can be explained by the same rationale in the last paragraph. Unlike their postgraduate counterparts, bachelor's students have more limited access to scholarships or external funding and often rely on self-financing. As a result, access to local employment becomes crucial for accumulating the necessary resources to study abroad. When the youth unemployment rises, their ability to fund international education declines, leading to lower outbound mobility.

Another dimension of this pattern is reflected in the effect of youth unemployment in the destination country on international student mobility. Although rising youth unemployment in the destination generally leads to a decline in incoming student numbers across all education levels, the effect is strongest for undergraduate students. As previously discussed, undergraduate students are more sensitive to labor market conditions in the host country, given their limited access to external funding. In contrast, the negative effect of youth unemployment is weakest for master's students in most models. This further reinforces the idea that the migration model is less dominant in explaining the mobility behavior of master's students.

5.3 Other Control Variables

In addition to monetary and educational factors, several variables can be categorized under the umbrella of cultural influences. For instance, when two countries have had a recent colonial relationship, it is likely that they share various cultural, linguistic, or institutional traits. Similar cultural proximity may also exist when countries share a common border or when there is a significant migrant population from one country residing in the other. As with other variable groups, the impact of cultural factors varies across education levels. Across all models, cultural variables exhibit the strongest effect on the mobility of bachelor's students. This can be attributed to their relatively younger age and lower level of independence, which may make them more sensitive to cultural familiarity and less equipped to adapt to significant cultural differences. Thus, shared cultural characteristics likely serve as a key facilitator of mobility at the undergraduate level.

In contrast to the broader influence of cultural variables, sharing a common official language appears to play a relatively minor role in shaping bachelor-level student mobility. In fact, the effect of a shared official language is insignificant in one of the models for bachelor's students. On the other hand, the increase in the number of incoming students is highest for bachelor's students when the destination country is English-speaking. These two findings can be interpreted as follows: for bachelor's students, the effect of a shared official language may be largely captured by other cultural factors, such as colonial ties or existing migrant networks. However, the influence of English as the official language of the destination likely reflects broader advantages that go beyond the scope of other cultural indicators. This distinction suggests that the role of English warrants further investigation as a unique factor influencing international student mobility, particularly at the undergraduate level.

Chapter 6

CONCLUSION

International students have attracted increasing attention over the past several decades. This study contributes to the literature on international student mobility by examining the determinants of student flows across different education levels. Using OECD data disaggregated by destination, origin, and education level, and applying a gravity model framework, the analysis reveals that master's students exhibit distinct mobility patterns compared to bachelor's and doctoral students. Based on the theoretical framework introduced by Rosenzweig (2006), the findings suggest that master's students are more likely to pursue international education with the intention of returning to their home countries upon completion. In contrast, bachelor's and doctoral students appear more motivated by opportunities to remain and work in the host country after their studies.

Students' varied motivations offer valuable insights for institutions and governments looking to improve their strategies for international students. Universities, in particular, benefit from a diverse student body, which is believed to enhance both the quality of education and research. This study's findings can help universities better allocate resources to achieve an optimal number and distribution of international students. On a broader scale, governments also have a keen interest in international students. Because these students are both a new source of resources (e.g., human capital, tuition fees) and a user of existing ones, governments face a trade-off. The findings of this study can help them better manage this balance, ensuring they do not lean too heavily in one direction.

BIBLIOGRAPHY

Abbott, A. and Silles, M. (2016). Determinants of international student migration. *The World Economy*, 39(5):621–635.

Anderson, J. E. and van Wincoop, E. (2003). Gravity with gravitas: a solution to the border puzzle. *The American Economic Review*, 93(1):170–192.

Baier, S. L. and Bergstrand, J. H. (2009). Bonus vetus OLS: a simple method for approximating international trade-cost effects using the gravity equation. *Journal of International Economics*, 77(1):77–85.

Beine, M., Bertoli, S., and Fernández-Huertas Moraga, J. (2016). A practitioners' guide to gravity models of international migration. *The World Economy*, 39(4):496–512.

Beine, M., Noël, R., and Ragot, L. (2014). Determinants of the international mobility of students. *Economics of Education Review*, 41:40–54.

Bertoli, S. and Fernández-Huertas Moraga, J. (2013). Multilateral resistance to migration. *Journal of Development Economics*, 102:79–100.

Bird, K. and Turner, S. (2014). College in the states: foreign student demand and higher education supply in the US. Working Paper 23, Curry School of Education, University of Virginia.

Caruso, R. and De Wit, H. (2015). Determinants of mobility of students in Europe: empirical evidence for the period 1998-2009. *Journal of Studies in International Education*, 19(3):265–282.

Conte, M., Cotterlaz, P., and Mayer, T. (2023). The CEPII Gravity Database.

Czaika, M. and Parsons, C. R. (2017). The gravity of high-skilled migration policies. *Demography*, 54(2):603–630.

Didisse, J., Nguyen-Huu, T. T., and Tran, T. A.-D. (2019). The long walk to knowledge: on the determinants of higher education mobility to Europe. *The Journal of Development Studies*, 55(6):1099–1120.

Kaushal, N. and Lanati, M. (2019). International student mobility: growth and dispersion. Technical Report 25921, National Bureau of Economic Research, Cambridge, MA.

Mathies, C. and Cantwell, B. (2022). Intra-regional mobility of PhD students in the European Union: the outcomes of region-making policy? *Policy Reviews in Higher Education*, 6(2):230–251.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, pages 105–142. Academic Press, New York.

Ortega, F. and Peri, G. (2013). The effect of income and immigration policies on international migration. *Migration Studies*, 1(1):47–74.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4):967–1012.

Rosenzweig, M. R. (2006). Global wage differences and international student flows. *Brookings Trade Forum*, 9:57–86.

Silva, J. M. C. S. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658.

Tinbergen, J. (1962). *Shaping the World Economy: Suggestions for an International Economic Policy*. The Twentieth Century Fund, New York.

Van Bouwel, L. and Veugelers, R. (2013). The determinants of student mobility in Europe: the quality dimension. *European Journal of Higher Education*, 3(2):172–190.

Wei, H. (2013). An empirical study on the determinants of international student mobility: a global perspective. *Higher Education*, 66(1):105–122.