Final Assignment DCM - Young Migrants and Degree Choice

Marc Sparhuber

Table of contents

Introduction

With migration expected to rise drastically in the coming decades, issues of integration continuously rise in relevance [SOURCE]. One aspect of migration is educational attainment and more specifically choice of degree field in tertiary education, as getting a degree serves as a vehicle for upward social mobility [SOURCE]. Especially for individuals who hae migrated at a young age, who have yet to make their decisions regarding their education, different reasons for degree choice may be paramount compared to non-migrant children of their age cohort. Thus, the central research question of this research paper is whether individuals who have migrated at a young age differ in their choice of tertiary education to non-migrants. Connected to this are the questions of whether their access to education is the same as that of their non-migrant peers and what hopes and attributions are connected to certain degree fields.

To answer these questions, binary, multinomial and conditional logistic regressions are employed. The analysis is based on data from the American Community Survey 2015, which is a demographic survey conducted yearly by the United States Census Bureau. It contains individual and household-level information about demographic, educational, and financial information, among other fields. Based on some key variables … [TEASER OF RESULTS HERE].

Background

There is a body of work which has already answered the questions above. Many studies have highlighted the lower the age of migration, the better the educational attainment overall, some even found a causal effect (e.g., Ansala et al., 2019, Glick & Yabucki, 2016; Söhn, 2011, Lemmermann & Riphahn, 2018). Other variables typically associated with educational attainment also play a role here: parental education and income, school quality, social capital, neighbourhood quality, among others (Baum & Flores, 2011, Portes & MacLeod, 1999, Söhn, 2011). Another important variable is being able to speak the target country's language (Baum & Flores, 2011, Lemmermann & Riphahn, 2018). A theory suggesting an explanation for this is child development theory, which posits that the younger an individual is, the easier it is for it to change and adapt to a new environment and the "cost" of adopting a new identity is lower than for older individuals (Beck et al., 2012).

Interestingly, the first two generations of migrants tend to achieve higher levels of tertiary education than natives, which was termed the "immigration advantage" and is hypothesized to be due to either positive selection (immigrants show higher levels of motivation and human capital than the average person from their country)

or immigrant optimism, which ascribes immigrants higher expectations and motivations in the form of psychological ressources to overcome socioeconomic hardships in their target country (Portes & Rivas, 2011). While information on degree choice among migrants seems sparse, ethnicity seems to predict it somewhat, with first- and second generation Asians being significantly more likely to attend university and college than natives (Hagy & Staniec, 2002). This is in line with findings that Asian Americans are tend to purse education leading to higher earnings and requiring higher education and enter into degrees that tend to be more financially rewarding (Xie & Goyette, 2003). This seems to be in part due to the fact that individuals in degree related occupations tend to encourage youths from their own demographic groups to pursue degrees related to their occupations (Ma, 2011).

Something about the specific case: in the US - costs associated with tertiary education high -> might be more utility to get an apprenticeship but there may be greater expectations attached, with people believing in the myth of the american dream? -> I.E., college as the main road to success## > But this must be seen in the context of the educational system (griga, hadjar, 2014)

It is expected that though there may be fewer young migrants choosing to pursue tertiary education due to the associated costs compared to their non-migrant counterparts (binary), those who do are driven by strong sense of utility in their choice of studies, therefore choosing fields commonly associated with higher prestige and better financial stability (multi). Additionally, the attributed outcomes associated with degree choice are expected to primarily be related to financial stability and prestige (clog).

This analysis has a breadth of potential confounders and mediators to take into account, many of which are expected to interact with one another and not all of which are included in the data. Sex is expected to serve as a strong confounder, as several fields of study are strongly gender segregated [IS THAT HOW YOU SAY IT?]. This effect is expected to differ, however, based on the socialization at home, where variables such as religion, country of birth as well as parent related variables may show interactions with sex and have their own confounding effects. Citizenship status may influence whether college is attended at all, as undocumented students may choose to not attend college due to either thinking that they are not allowed to do so or because their funding options by the state are reduced [https://counselors.collegeboard.org/financial-aid/undocumented-students]. Additionally, degree choices may be greatly influenced by the age at which an individual has migrated to the US country. Due to this, some of the following models take this into account.

Table 1: Educational and Economical Indicators: Descriptives Divided by Migration Status

Characteristic	**Migrated as Child (younger than 10 years old)**, N = 28,197	**Migrated as Adolescent (between 10 and 19 years old)**, N = 54,017	**Migrated as Adult (20 and older)**, N = 130,549	**Did not Migrate**, N = 995,505
Age	37 (30, 45)	38 (31, 46)	43 (36, 49)	41 (32, 49)
Sex				
Male	13,433 (48%)	27,842 (52%)	60,324 (46%)	481,931 (48%)
Female	14,764 (52%)	26,175 (48%)	70,225 (54%)	513,574 (52%)
Citizenship Status				
Citizen	1,864 (6.6%)	1,514 (2.8%)	3,132 (2.4%)	981,304 (99%)
Born abroad of American parents	0 (0%)	0 (0%)	0 (0%)	13,278 (1.3%)
Naturalized citizen	19,439 (69%)	28,191 (52%)	47,504 (36%)	668 (<0.1%)
Not a citizen	6,894 (24%)	24,312 (45%)	79,913 (61%)	255 (<0.1%)
Race (simplified)				
White	5,289 (19%)	6,301 (12%)	25,067 (19%)	774,210 (78%)
Black/African American	1,619 (5.7%)	3,801 (7.0%)	10,006 (7.7%)	106,710 (11%)
American Indian/Alaska Native	15 (<0.1%)	21 (<0.1%)	87 (<0.1%)	11,011 (1.1%)
Asian/Pacific Islander	7,363 (26%)	11,997 (22%)	42,126 (32%)	17,276 (1.7%)
Hispanic/Latino	13,911 (49%)	31,897 (59%)	53,263 (41%)	86,298 (8.7%)
Educational Attainment				
Did not Finish High School	2,779 (9.9%)	15,242 (28%)	27,215 (21%)	53,565 (5.4%)
Finished High School	8,465 (30%)	16,897 (31%)	33,688 (26%)	334,764 (34%)
Some College	6,769 (24%)	9,368 (17%)	18,773 (14%)	255,397 (26%)
Bachelor's Degree	6,459 (23%)	8,146 (15%)	26,671 (20%)	231,297 (23%)
Master's Degree or Doctorate	3,725 (13%)	4,364 (8.1%)	24,202 (19%)	120,482 (12%)
Employment Status				
Employed	22,309 (79%)	42,194 (78%)	96,891 (74%)	786,813 (79%)
Unemployed	1,286 (4.6%)	2,194 (4.1%)	5,097 (3.9%)	41,896 (4.2%)
Not in labor force	4,602 (16%)	9,629 (18%)	28,561 (22%)	166,796 (17%)
Usual hours worked per week	40 (28, 40)	40 (25, 40)	40 (20, 40)	40 (28, 43)
Total personal income	31,868 (12,416, 62,081)	25,867 (10,347, 49,044)	24,108 (7,243, 51,734)	34,972 (13,451, 62,081)
Total personal earned income	31,040 (10,347, 62,081)	25,867 (9,312, 46,561)	22,763 (4,656, 51,734)	33,110 (10,347, 61,046)
Poverty status	354 (195, 501)	261 (146, 462)	290 (152, 501)	390 (219, 501)

[1] Median (IQR); n (%)

Note: Shows only those respondents who were born after 1960 and at least 24 years old in 2015. Poverty Status is calculated taking into account household size, household member's age, age of the householder and total family income: Values over 100 indicate being above the poverty threshold. Data from the American Community Survey 2015 (collected 2015).

Data

> The data for the 2015 American Community Survey was gathered over the course of the year 2015, containing more than 3 million individuals and is representative for the inhabitants of the Unites States of America.

Describe how you generated your analytic sample from the data > Of these data, only those individuals are taken into account that were born 1960 or later. This is to ensure that all individuals had the chance to do all types of degrees with especially degrees such as computer science only being widely represented at colleges around the late 70s [SOURCE]. Further, depending on the analysis, only employed individuals are included.

Explain how you treated missing values. > Individuals containing missing values in any of the variables of interest are dropped from the analysis.

Clearly identify the dependent variable(s). > The main dependent variable is the choice of degree. For the multinomial regression, the degrees are grouped into broader fields, whereas the conditional logit model takes into account all degrees as alternatives in which at least 1000 individuals received a degree. [Kind of switched around for clogit].

> (Portes, MacLeod, 1999) > Studies who compare ethnicities have a big problem of lumping together and snapshots -> unavoidable here.

Identify and explain the operationalization of key independent variables. > The key independent variable is the "Migration status". It is a categorical variable with the 4 levels "Migrated as Child", "Migrated as Adolescent", "Migrated as Adult" and, "Did not Migrate". The split into these 4 categories is motivated theoretically, as there is expected to be a greater assimilation of migrants, the younger they enter the country. Individuals who are born abroad to American parents and then moved to the U.S are not considered migrants in this analysis,

due to the assumption that they may adapt to their new environment much easier than other migrants due to the knowledge and [WORD] of their parents. Other independent variables are collapsed versions of the birthplace and race variables. [MAYBE MORE HERE].

Produce and discuss descriptive statistics.

> Table 1 presents the descriptives divided by migration status. For categorical variables percentages are displayed and for continuous variables, the median and interquartile range are supplied. Across the entire table it becomes apparent that inidividuals who migrate as children quite closely resemble the descriptive statistics of those who never migrated at all, compared to those who migrated as adults. There seems to be a trend to be worse off when having migrated as an adult compared to all other groups, with a lower total personal and total earned income than the other groups. However, the educational attainment variable results for this group show that this might not be due to low education, as this group contains the highest percentage of degree holders, hinting toward distinct differences between highly educated individuals migrating to the US and less qualified migrants. Comparing the child and adolescent migrants, it becomes apparent that the latter performs worse than the prior in all educational and financial variables, being the most at risk group of all four compared. [COULD WRITE MORE HERE].

Table 2: Young Migrant Status with associated variables: Binary Logistic Regression Models

	Model 1			**Model 2**			**Model 3**		
Covariate	**Odds**	**95% CI**	**p-value**	**Odds**	**95% CI**	**p-value**	**Odds**	**95% CI**	**p-value**
Field of Degree									
Business	—	—		—	—		—	—	
Sciences	1.20	1.10, 1.30	<0.001	1.17	1.08, 1.27	<0.001	1.16	1.07, 1.26	<0.001
Engineering	1.21	1.10, 1.33	<0.001	1.23	1.12, 1.35	<0.001	1.22	1.11, 1.35	<0.001
Arts	0.64	0.58, 0.70	<0.001	0.61	0.55, 0.67	<0.001	0.61	0.55, 0.67	<0.001
Social Sciences	0.80	0.74, 0.87	<0.001	0.77	0.71, 0.84	<0.001	0.77	0.70, 0.83	<0.001
Education	0.37	0.32, 0.43	<0.001	0.36	0.31, 0.42	<0.001	0.37	0.32, 0.42	<0.001
Medicine	0.98	0.88, 1.10	0.8	0.94	0.84, 1.05	0.3	0.94	0.84, 1.05	0.3
Other	0.73	0.62, 0.86	<0.001	0.71	0.60, 0.83	<0.001	0.71	0.60, 0.84	<0.001
Age				0.98	0.98, 0.98	<0.001	0.98	0.98, 0.98	<0.001
Sex									
Male				—	—		—	—	
Female				1.09	1.03, 1.16	0.003	1.10	1.03, 1.17	0.002
Usual hours worked per week							1.00	0.99, 1.00	<0.001
Total personal income							1.00	1.00, 1.00	0.034
Total personal earned income							1.00	1.00, 1.00	0.001
Poverty status							1.00	1.00, 1.00	<0.001

[1] CI = Confidence Interval

Note: Data from the American Community Survey 2015 (collected 2015). The units for the age term is years.

Table 3: Young Migrant Status and [SOMETHING]: Model Fit Statistics

	Model 1	Model 2	Model 3
Log Likelihood	-21697	-21608.84	-21580.43
AIC	43410	43237.67	43188.86
BIC	43486.53	43333.33	43322.78
Likelihood Ratio		176.33	56.82
Likelihood Ratio (Df)		2	4
Likelihood Ratio (p-value)		5.14e-39	1.35e-11

Note: The Likelihood ratio is always calculated with the nested model to the left. Data from the American Community Survey 2015 (collected 2015)

Table 4: Degree Choice: Multinomial logistic regression

	(1)						
	Sciences	Engineering	Arts	Social Sciences	Education	Medicine	Other
Female (ref. Male)	0.66***	0.26***	1.32**	1.48***	2.99***	4.33***	1.13
	[0.57, 0.77]	[0.22, 0.31]	[1.11, 1.57]	[1.27, 1.73]	[2.35, 3.82]	[3.44, 5.45]	[0.84, 1.51]
Migrated as Child (younger than 10 years old - ref. Did not Migrate)	1.25	1.16	0.98	1.12	0.51*	0.89	1.20
	[0.85, 1.84]	[0.76, 1.78]	[0.65, 1.47]	[0.77, 1.62]	[0.29, 0.88]	[0.54, 1.47]	[0.61, 2.37]
Migrated as Adolescent (between 10 and 19 years old)	1.26	1.42	0.63*	0.73	0.34***	0.90	0.62
	[0.84, 1.88]	[0.92, 2.20]	[0.41, 0.98]	[0.49, 1.08]	[0.19, 0.62]	[0.53, 1.51]	[0.30, 1.32]
Migrated as Adult (20 and older)	1.82**	2.99***	0.77	0.68+	0.89	1.33	0.55
	[1.23, 2.70]	[1.96, 4.58]	[0.50, 1.18]	[0.46, 1.00]	[0.51, 1.53]	[0.80, 2.21]	[0.26, 1.18]
Birthplace: LatAm (ref. United States)	0.57**	0.72	0.63*	0.82	0.97	0.68	1.03
	[0.38, 0.84]	[0.48, 1.09]	[0.42, 0.96]	[0.56, 1.19]	[0.56, 1.68]	[0.41, 1.13]	[0.52, 2.04]
Birthplace: Western Europe & Scandinavia	1.14	1.06	1.68*	1.12	1.02	0.70	0.91
	[0.72, 1.80]	[0.65, 1.72]	[1.05, 2.69]	[0.71, 1.76]	[0.53, 1.97]	[0.37, 1.35]	[0.38, 2.17]
Birthplace: Eastern Europe	1.17	0.89	1.18	0.97	0.90	0.74	0.90
	[0.74, 1.85]	[0.54, 1.49]	[0.71, 1.94]	[0.61, 1.54]	[0.45, 1.78]	[0.39, 1.39]	[0.37, 2.20]
Birthplace: East Asia	1.38	1.20	0.84	0.73	0.29***	0.52*	0.60
	[0.93, 2.05]	[0.78, 1.84]	[0.54, 1.30]	[0.48, 1.09]	[0.14, 0.58]	[0.30, 0.91]	[0.27, 1.33]
Birthplace: South & South-East Asia	1.50*	1.50+	0.47**	0.88	0.55+	1.74*	0.54
	[1.02, 2.19]	[0.99, 2.25]	[0.30, 0.74]	[0.59, 1.29]	[0.30, 1.00]	[1.06, 2.85]	[0.25, 1.17]
Birthplace: Middle East	1.53	1.63+	0.74	0.95	0.69	0.82	0.63
	[0.90, 2.59]	[0.93, 2.84]	[0.38, 1.43]	[0.54, 1.66]	[0.28, 1.70]	[0.39, 1.76]	[0.19, 2.11]
Birthplace: Africa	0.66	0.61+	0.59+	0.69	0.58	1.44	0.11*
	[0.40, 1.10]	[0.36, 1.05]	[0.33, 1.08]	[0.41, 1.17]	[0.26, 1.30]	[0.78, 2.65]	[0.01, 0.85]
Birthplace: Oceania	0.80	0.23+	0.75	0.97	1.01	1.29	1.69
	[0.28, 2.28]	[0.05, 1.15]	[0.22, 2.59]	[0.34, 2.79]	[0.24, 4.18]	[0.39, 4.32]	[0.32, 8.90]
Num.Obs.	7262						
AIC	26 724.3						
BIC	27 351.3						

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Source: Data from the American Community Survey 2015.

Covariate	**Odds**	**95% CI**	**p-value**
Mean hours worked (week) * Migration Status (dichotomized)			
Mean hours worked (week) * Not YM	0.92	0.90, 0.94	<0.001
Mean hours worked (week) * YM	0.82	0.80, 0.84	<0.001
Migration Status (dichotomized) * Mean Poverty			
Not YM * Mean Poverty	1.04	1.04, 1.05	<0.001
YM * Mean Poverty	1.04	1.03, 1.04	<0.001
Migration Status (dichotomized) * Mean earned income (year)			
Not YM * Mean earned income (year)	1.00	1.00, 1.00	<0.001
YM * Mean earned income (year)	1.00	1.00, 1.00	<0.001

[1] CI = Confidence Interval

Note: Data from the American Community Survey 2015 (collected 2015). The units for the age term is years.