

The intergenerational mobility of Immigrants : How persistent is pre-migration parental background?

Pascal Achard

September 3, 2016

Abstract

This paper studies the heterogeneity in schooling and labour market achievements among second generation immigrants in France. I focus on immigrants who have a low background (as measured by father's occupation) in France but who come from different backgrounds in the origin country (as measured by grandparents' occupation, grandparents' education and parents' occupation before migration). I find that the second generation of previously high and currently low perform substantially better than those originally from a low background. This effect is strong irrespective of the country of origin and stronger for people living in poor neighborhoods. Focusing on children with the same social situation but very different parents also allow me to study within household parental transmission. I test several hypotheses as of what can explain this difference. Results show that it is not due to parental investment (either material or time investment), nor to parental schooling strategy.

PRELIMINARY AND INCOMPLETE. DO NOT QUOTE

1 Introduction

Motivation Around 20% of the French population is first or second generation immigrant. Children of immigrants perform (on average) worse than their native counterparts; they are more likely to drop out of high school, more likely to be unemployed, less likely to have white collar jobs. This has fueled a lot of media and political discussion around the issues of immigration and social integration. However, recent work¹ has highlighted the heterogeneity among this population. One common element in most of the research done on immigrants is to only consider their situation in the destination country and leave aside elements that pre-existed migration. Works that have included pre-migration characteristics only look at the country of origin as if it was a sufficient statistics to anything that happened before migrating. But there are many sources of heterogeneity within countries. Not every immigrant from a country with an average low skilled workforce is necessary low skilled, not every immigrant from a socially conservative country is necessary conservative. **This paper looks at whether the nature of the heterogeneity that existed before migration with the parents (and that went unnoticed when the parents arrived) explains the heterogeneity among the children.**

I look in particular at how much does the social background of the family in the origin country help understand why some second generation immigrants are more successful than others. This question would have a very intuitive answer if immigrants from middle/high strata in their country of origin remain in the same social situation once in France. Differences would not be due to pre-migration household characteristics but to differences in background in France. However, migration corresponds to a downward shock for many immigrants who move from high to low backgrounds when they cross borders. **To my knowledge, this paper is the first to introduce household level pre-migration characteristics in the study of the future of second generation immigrants.**

Once the population of interest is narrowed down to second generation immigrants with a low background in France, it is less clear whether or not the social status of the parents before arrival should matter. In a qualitative study on black West Indian immigrants in the U.S., Mary Waters found that second generation move away from their parents' high aspiration to adopt the lifestyle and behaviors of their low achieving peers:

These different interpretations of the role of race will play in one's life also create a gulf between parents and children. Parents will tell their children to strive for upward social mobility and have high aspirations, but often the peer group and the children's own day-to-day experience tell them the color of their skin might make it difficult or impossible to meet those aspirations. Waters (2009)

Research Question Does everything restart when you migrate? For a similar (and low) social status in France but from a different status in the origin country, how different/similar are the achievements of the children? And if they are different, what are the non monetary elements that parents transmit their children?

¹see Beauchemin et al. (2015) as a general reference

Review of the literature and Contribution This paper is at the junction of different literatures; on the long-term persistence of parental background, on social mobility of immigrants and on parental transmission to second generation immigrants. Because on each of these subjects the literature is very large, I only mention a limited number of papers and I detail what is my contribution to each literature.

On long term persistence of parental background. Black and Devereux (2011) provide a summary of the old debates and recent contributions to the literature on intergenerational mobility. Lindahl et al. (2015) use data compiling educational achievements for 4 generations in Sweden to show that adding generations increase the estimates of persistence of family background.

The focus of this literature has been on persistence through time, in this paper I look at persistence through space in a situation where immigrants have little material assets to transfer to their children.

On immigration and social mobility. Besides seminal work of Borjas (1993, 1992), Aydemir et al. (2009) have more recently looked at the intergenerational mobility among second generation immigrants in Canada and compare it with that of natives.

The focus of this literature has been on relating success of the children to the country of origin of the parents. This approach has limitations if you consider that immigrants from the same country are heterogenous and that the cost of adjusting to the destination country is different for different strata of the origin country. In this paper, I look at the social dimension beyond the country of origin.

On parental transmission to second generation immigrants. Bisin and Verdier (2010) have developed theoretical models explaining when first generation immigrants transmit their cultural heritage to their children. Fogli and Fernandez (2009) showed in an empirical work that female labor supply among immigrants in the US was related to average female participation in the origin country.

This is probably the literature that is closest to the focus of this paper. In this paper, I try to look at specific and direct channels of parental transmission (through schooling strategy and parental investment) to explain the differences in achievements.

Data The dataset used in this paper is “Trajectoires et Origines”, (TeO) collected by INED and INSEE in 2008/2009. This data is a very rich source of information and has several advantages as compared to other commonly used data sources in this literature (labor force survey or census).

1. It has been designed with the specific purpose of studying the population of first and second generation immigrants, so it asks very relevant question for this population.
2. It is the unique dataset (at least in France) to have specifically sampled 2nd generation. People don't appear as immigrants in civil registries if they were born in France. So it required the designers of the survey to sample this population from registries of the parents. This ensures that the sample surveyed is representative of the entire population.
3. Most importantly for this study, it has some information on the social situation prior to migration.

Figure 1: How to generate two samples from one survey

		Part of the Survey		
		Personal History	Current Situation	Children
Sampled as	first generation	grandparents parents before migration	parents in France	population of interest
	second generation	parents in France parents before migration	population of interested	

Note: In each cell is described the population that I am able to target by looking at the sample in the survey (rows) and the questions asked to him/her (columns)

In everything that follows, second generation immigrants refer to people whose both parents are immigrants. The reason why I choose to look at families with both parents immigrants (as opposed to those with one or two parents) is to make sure that I capture the influence of the pre-migration background. By focusing on children of mixed couples (one immigrant and one native), the influence of the pre-migration status is mixed with the influence of the background of the parent who is a native².

Three different populations are being surveyed in TeO: (i) first generation immigrants, (ii) second generation immigrants and (iii) natives. The questions refer to different moments in their lives: the personal history (a long part of the survey), the current situation (the longest part of the survey) and the children (a small part of the survey) and can be used differently depending on the population used. Figure 1 shows how different sections of the survey can be used to generate two samples of the population that I am interested in.

When using the sample that surveys first generation (from now on referred to as sample 1), I have a lot of information on the situation prior to migration, a lot of information on the parents' situation in France but little on the situation of the children and little on transmission mechanisms. When using the sample that surveys second generation (from now on referred to as sample 2), I have a lot of information on the situation of the children and on the potential transmission mechanisms but little on the parents' situation in France and prior to migration.

Being able to use two different samples from the same survey has several advantages: first it allows for a robustness 'out of sample' check and second it allows to use each sample to answer the question for which the information contained in the sample surveyed is the most relevant.

In what follows, I will use sample 1 to establish the difference in success among second generation immigrants and sample 2 to look into the potential explanations for this difference. When possible, I will show that there results are similar in the two samples.

Structure of the paper The paper is structured as follows: in section 2, I will show that migration reduces the heterogeneity of the population creating a difficulty to read correctly the social status of the family when only looking at the situation in France. In section 3, I will show that there is a strong difference in achievements among children of immigrants depending on the social status of their parents in the origin country (for a given low background in France). In section 4, I will dig into the transmission mechanisms that can explain this difference. Section 5 does several robustness checks and section 6 concludes.

²However, considering that it can be a limitation, I detail how the populations of children with one and two parents immigrants differ in section 5. I also address as a robustness check the question of whether results change if I define second generation immigrants as individuals with only one parent immigrant.

2 Migration reduces variance

2.1 Notation and Definition

2.1.1 Notation

S stands for status, subscripts P and C stand for parents and children. For parents $t = 0$ refers to pre-migration, $t = 1$ refers to the time of arrival in France, $t = 2$ refers to the situation at the time of the survey. So parents' status is observed three times $S_{P,0}$, $S_{P,1}$, $S_{P,2}$ and children's status only once S_C . Status can take two values H or L (for high and low).

2.1.2 Definition of the statuses

How to define high and low status in France? It is based on the occupation of the father (or the last occupation for those retired or unemployed). The National Statistical Agency (INSEE) defines six type of socio-professional category: Self Employed Agricultural, Self Employed Non Agricultural, High Managerial, Supervisory Occupations, Lower Services and Lower Technical. Are considered high status, the individuals whose occupation is classified as High Managerial or Supervisory Occupations, the rest is classified as low status. I look at occupation rather than at wages, because I don't observe wages of the parents and the children.

How to define high and low status in the country of origin? It is not an easy task to come up with a definition of what it means to be from a high background in the country of origin. Ideally, this definition should be country and time specific, having a high school diploma means something different in a developed and in a developing country, it means something else for someone who grew up in the 70s than someone who grew up in the 40s. It is even more difficult since there is little information on the distribution of educational achievements and occupations for developing countries in the 30s, 40s and 50s when some of the parents or grandparents in the population of interest were children³.

Instead, I use definitions that are common to all countries and time periods. In the statistical analysis, I use country of origin fixed effects or look at a subset of country of origins to account for the difference in high/low that are due to the general level of development of the country and not the relative social position of the family in the country of origin.

To define high and low I use information on both (when possible) education and occupation of the parents and the grandparents. For occupation, I use the same categories that I use for France (i.e. same classification) and for education, I use the criteria having or not finished primary school. I do this choice for the following reasons: (i) it makes intuitive sense to use the

³Barro and Lee (2013) provide mean educational attainment (in years of education) for every period of five years in most countries in our sample for the population aged 15 to 64 at the time. It is however of little help since it tells, for instance, in 1910 what was the average education not what was the average education for those born in 1960 (the information I would ideally like to have), and this information cannot be reconstructed easily. The other problem with this data source is that there is little variation in the measure of education since it is averaged over people 15 to 64 and thus is not very sensitive to changes for one cohort. The third problem is that values for average education are very low for many countries in our sample (zero for some countries and some years) meaning that very few people would be classified as low. However to make sure that the results don't depend on a trend in education in the countries of origin, I use for every year of birth the corresponding average number of years of schooling in that particular year as a control, see table 27.

end of an education cycle⁴, (ii) it corresponds to roughly the top 20% of my primary sample of analysis (the $S_{P,2} = L$ in sample 1) and so is neither too high nor too low. As I have two samples with difference variables, I detail below the definitions for both. To check that my results are not due to my definition of high and low in the origin country, I show results when I use variations of the definitions.

For sample 1 I develop a so-called large definition that has two subcomponents education and socio-professional status. The education component works as follows: if the grandmother or the grandfather (of the person being interviewed, whether it is the mother or the father) has finished primary school, the pre-migration status is considered high. Otherwise it is considered low. The socio-professional component works as follows: if the person being interviewed (whether it is the mother or the father) has a high occupation (high managerial or supervisory occupation) before migrating or if the grandfather (of either the mother or the father's side) had a last known occupation to be high, the pre-migration status is considered to be high. Otherwise it is considered low. The large definition is high if either the education or the socio-professional definition is high. Otherwise it is considered low.

The benefits of this definition is that it includes several dimensions of the pre-migration status. The criteria are also relatively far in time for the perspective of second generation immigrants.

For sample 2 I develop a so-called narrow and a large definition. The large definition works as follows: if one parent (and not grandparents as in the definition of sample 1) has finished primary school the pre-migration status is high. Otherwise it is low. The narrow definition works as follows: if both parents have finished primary school, the pre-migration status is high. Otherwise it is low.

The drawbacks of this definition is that while it includes pre-migration information only (since our sample of parents have had schooling in their country of origin), it is only based on education and doesn't go back as far in time as the definitions of sample 1.

Is it possible to map the population of both samples into each other? By definition it is not possible to map observations of sample 2 into sample 1 since information on grandparents are missing from sample 2. However it is possible to look in sample 1 at the observations that would be classified as high or low based on the criteria of sample 2. When I map the definitions of sample 2 into sample 1, 70% of the observations remain in the same category⁵. The results using the observations from sample 1 and the definition of sample 2 are extremely close to the results using both sample 1 observations and definitions⁶.

In the rest of section 2, I use observation from sample 1. In section 3, I use observations from sample 1 in my main specifications. In section 4, I use observations from sample 2.

⁴I cannot use years of schooling since TeO only reports discrete categories

⁵For example using the narrow definition on the 3661 observations used in the baseline regression reported in table 11, 2003 observations are characterized as low according to the definitions of the two samples, 474 are characterized as high according to both definitions, 547 are low according to the definition of sample 1 and high according to the definition of sample 2 and 520 the other way around, 117 observations are missing one of the definitions. The results are available in the appendix.

⁶See the first four columns of table 26

2.2 There was heterogeneity before migration

Table 1 describes the percentages of immigrants coming from a high or low background by country (or region) of origin⁷. There is clearly a difference between immigrants coming from developed versus developing countries, reflecting the difference in levels of education.

However not all immigrants from developed countries are characterized as high and not all immigrants from developing countries are characterized as low. There is always a substantial fraction of immigrants coming from a different background than the one commonly associated with their country of origin. For example, at least 25% of the immigrants from Maghreb are from a high background and even for European immigration Spain and Italy provide a large share (more than 65%) of low background immigrants. The situation for African countries is very heterogenous, with countries such as Mali and Senegal providing mostly low background immigrants and countries such as Cameroon or Congo Brazzaville providing mostly high background immigrants. Many countries provide more or less equal share of high and low background people, such as Cambodia, Ivory Cost, Laos and Central American countries.

Table 1 both shows, (i) that there is some ground in digging beyond the information of the country of origin and (ii) that the definition of high and low applies very differently to different countries. I restrict later some analysis to certain country of origins or include country fixed effects in linear regressions to make sure that I estimate a difference on comparable populations.

2.3 This heterogeneity was reduced at the time of migration

Table 2 shows the transition matrix both in absolute numbers and in percentages between the pre migration status and the status upon arrival⁸ for men⁹. It also reports marginal distributions. When most people from a low background in the country of origin remain in a low background, a very important share (almost 70%) suffer a downward shock from high to low¹⁰.

To check that these numbers are not due to European immigrants being characterized as high because my threshold (especially for education) is relatively low, I look at two subsamples; immigrants from North Africa and Maghreb and immigrants living in poor neighborhoods¹¹. The numbers in table 3 are very close to the general population. For those living in Urban

⁷For clarity, I report countries (or regions) for which there are more than 20 observations. I don't have disaggregated data by countries for some regions but those are not the main countries contributing to immigration in France.

⁸Results are very similar when $S_{P,2}$ is used instead of $S_{P,1}$ but my argument is on the process of migration and is thus better captured with the information closest to the time of arrival.

⁹To not capture issues related to non-participation, I define high based on the occupation of the father. When the person being interviewed is a woman, I use information on the occupation of the partner provided they report the partnership having begun before the birth of the child I am looking at. TeO contains current occupation, but not first occupation in France, of the partner so I have to limit the sample to men for transition matrices 2, 3, 4.

¹⁰There is a fraction of people moving from low to high when migrating. This situation doesn't fit the hypothesis that I am testing in this paper. This fraction is anyway very small (from 4.5 to 8% depending on the subsample I am looking at).

¹¹I define poor neighborhoods as being labelled a Urban Sensitive Area

Table 1: Descriptive statistics

Country of origin	$S_{P,0=L}$	$S_{P,0=H}$	Nb of Observations
Algeria	73.6	26.4	318
Germany	12.8	87.2	39
Central America	50.0	50.0	32
South America	17.9	82.1	28
Africa (Other)	33.7	66.3	83
Europe (Other)	31.0	69.0	71
Belgium	20.4	79.6	49
Cambodia	55.1	44.9	107
Cameroun	21.7	78.3	23
Congo B	20.0	80.0	35
Ivory Cost	43.8	56.2	32
Spain	67.7	32.3	158
Italy	68.8	31.2	157
Laos	54.1	45.9	98
Mali	70.8	29.2	48
Marocco	75.8	24.2	326
Middle East	12.1	87.9	58
Poland	32.0	68.0	25
Portugal	76.0	24.0	549
RDC	24.5	75.5	49
Senegal	63.7	36.2	80
Tunisia	65.9	34.1	132
Turkey	67.3	32.7	278
UK	9.4	90.6	32
Vietnam	34.1	65.9	123
Asia (Other)	21.3	78.7	61

Note : The Observations are first generation immigrants (sample 1) who have at least one child who is at least 18 and was born in France or arrived before age 10. High and Low refer to the general definition used for sample 1

Sensitive Areas (in table 4), the numbers culminate to 90% having faced a downward shock. This is not a surprise since the probability of $S_{P,2} = H$ and living in a poor neighborhood is very low (only 6.51%). However it shows that among this subpopulation almost all the $S_{P,0} = H$ have suffered from a downward shock.

Table 2: Transition matrix men - Before migration to first job

	Absolute Numbers			Percentages		
	$S_{P,1}=L$	$S_{P,1}=H$	Total	$S_{P,1}=L$	$S_{P,1}=H$	Total
$S_{P,0}=L$	833	69	902	92.35	7.65	56.66
$S_{P,0}=H$	480	210	690	69.57	30.43	43.34
Total	1 313	279	1 592	82.47	17.53	

Note : The Observations are first generation male immigrants (sample 1) who have at least one child who is at least 18 and was born in France or arrived before age 10. High and Low refer to the general definition used for sample 1

Table 3: Transition matrix - Before migration to first job - Maghreb and African origin

	Absolute Numbers			Percentages		
	$S_{P,1}=L$	$S_{P,1}=H$	Total	$S_{P,1}=L$	$S_{P,1}=H$	Total
$S_{P,0}=L$	329	29	358	91.90	8.10	60.07
$S_{P,0}=H$	175	63	238	73.53	26.47	39.93
Total	504	92	596	84.56	15.44	

Note : The Observations are first generation male immigrants (sample 1) from Maghreb or Sub Saharan Africa who have at least one child who is at least 18 and was born in France or arrived before age 10. High and Low refer to the general definition used for sample 1

Table 4: Transition matrix - Before migration to first job - Urban Sensitive Area

	Absolute Numbers			Percentages		
	$S_{P,1}=L$	$S_{P,1}=H$	Total	$S_{P,1}=L$	$S_{P,1}=H$	Total
$S_{P,0}=L$	189	9	198	95.45	4.55	67.81
$S_{P,0}=H$	84	10	94	89.36	10.64	32.19
Total	273	19	292	93.49	6.51	

Note : The Observations are first generation male immigrants (sample 1) living in Urban Sensitive Area who have at least one child who is at least 18 and was born in France or arrived before age 10. High and Low refer to the general definition used for sample 1

2.4 Once in France, the population of first generation remained very homogeneous

Table 1 established that our population of interest was heterogenous and tables 2, 3, 4 that this population was homogenized when it migrated to France. Tables 5, 6, 7 show that this population for its most part remained homogeneous in France.

These tables show (for the same subsamples) that between 87% to 92% of male immigrants remained in a low status. So the population of $S_{P,0} = H$, $S_{P,1} = L$ has difficulties in France being distinguished from the population of $S_{P,0} = L$, $S_{P,1} = L$. Somehow what makes the population of previously high different from the previous low (level of education, abilities, networks in the origin country...) failed to have been identified in France or to have operated once in France. For comparison with natives, the last row of table 5 reports distribution among natives.

A question that arises at this stage is whether I should use $S_{P,1}$ instead of $S_{P,2}$ in latter analysis. As I am interested in looking at the influence of pre-migration social status for a given (and low) background in France, I don't want my estimates to be contaminated by the improvement in status from the population of $S_{P,1} = L$, $S_{P,2} = H$. For this reason, I stick to the criteria of $S_{P,2}$. As explained in footnote 9, using $S_{P,1}$ would also reduce my sample size and thus my statistical power. I would only be able to look at general specification (baseline model and with country fixed effects) and not at more specific subsamples (people living in urban sensitive areas, from certain origins, from certain origins in urban sensitive area).

3 The heterogeneity reappeared with the second generation

In this section, I look at how much the achievements of the children of immigrants differ depending on the pre-migration status of the parents. I have a series of achievements at adulthood. I can present my results with both transition matrices or linear regressions. Both have their strengths and weaknesses

1. Transition matrices are very easy to read and give a clear picture
2. Linear models allow to test the null hypothesis that conditioning on potential co-founders, the difference is significantly different from zero.

In this section, I will use both ways of presenting. The results are qualitatively similar.

3.1 Why focus on the previously high, currently low and the previously low, currently low?

The objective is to fix the situation in the destination country and vary the situation in the origin country. That left only two possibilities: comparing $S_{P,0} = L$, $S_{P,2} = L$ with $S_{P,0} = H$, $S_{P,2} = L$ or $S_{P,0} = L$, $S_{P,2} = H$ with $S_{P,0} = H$, $S_{P,2} = H$. Since the population of $S_{P,1} = L$, $S_{P,2} = H$ is a very small fraction of our sample, I find this comparison less relevant and focus on the first one instead. Moreover, the political debate is centered around immigrants from a poor background and the hypothesis to be tested is how to link the success of some with long term family achievements. So, I also find this comparison more interesting from a public policy perspective.

Table 5: Transition matrix - from arrival to current job

	Absolute Numbers			Percentages		
	$S_{P,2=L}$	$S_{P,2=H}$	Total	$S_{P,2=L}$	$S_{P,2=H}$	Total
$S_{P,1=L}$	1 130	183	1 313	86.06	13.94	82.47
$S_{P,1=H}$	34	245	279	12.19	87.81	17.53
Total	1 164	428	1 592	73.12	26.88	
Among Natives	584	356	940	62.13	37.87	

Note : The Observations in the first three rows are first generation male immigrants (sample 1) who have at least one child who is at least 18 and was born in France or arrived before age 10. High and Low refer to the general definition used for sample 1. The observations in the last row are natives who have at least one child who is at least 18

Table 6: Transition matrix - from arrival to current job - Maghreb and African origin

	Absolute Numbers			Percentages		
	$S_{P,2=L}$	$S_{P,2=H}$	Total	$S_{P,2=L}$	$S_{P,2=H}$	Total
$S_{P,1=L}$	436	68	504	86.51	13.49	84.56
$S_{P,1=H}$	12	80	92	13.04	86.96	15.44
Total	448	148	596	75.17	24.83	

Note : The Observations are first generation male immigrants (sample 1) from Maghreb or Sub Saharan Africa who have at least one child who is at least 18 and was born in France or arrived before age 10. High and Low refer to the general definition used for sample 1

Table 7: Transition matrix - from arrival to current job - Urban Sensitive Area

	Absolute Numbers			Percentages		
	$S_{P,2=L}$	$S_{P,2=H}$	Total	$S_{P,2=L}$	$S_{P,2=H}$	Total
$S_{P,1=L}$	252	21	273	92.31	7.69	93.49
$S_{P,1=H}$	7	12	19	36.84	63.16	6.51
Total	259	33	292	88.70	11.30	

Note : The Observations are first generation male immigrants (sample 1) living in Urban Sensitive Area who have at least one child who is at least 18 and was born in France or arrived before age 10. High and Low refer to the general definition used for sample 1

As one of the contribution of this paper to break down the immigrant population according to criteria at the junction of their status in the origin and the destination country, I think it is interesting to show some descriptive statistics in order to have a better understanding of the populations we are talking about.

Table 8 gives more information at a geographical disaggregated level, for immigrants from Europe, Asia, Maghreb and Africa and for the three following combinations : (1) $S_{P,0} = H$, $S_{P,2} = H$, (2) $S_{P,0} = L$, $S_{P,2} = L$, (3) $S_{P,0} = H$, $S_{P,2} = L$. It is also a way to check that within geographical areas: subgroups (1), (2), (3) differ/resemble along the expected lines (have similar/different characteristics in the country of origin and/or in France) for characteristics that are not used in the definitions of high and low. I report the mean year of arrival, the proportion of male, the proportion of people who reported good or very good level of (spoken) French when they arrived, the proportion of immigrants who completed secondary and higher education. I also report the mode of occupation of the grandfather, the mode of occupations¹² for fathers and mothers.

¹²I use here a two digit description of the occupation

Table 8: Descriptive statistics

	Mean			Number of Obs		
	$S_{P,0=H}$ (1)	$S_{P,0=H}$ (2)	$S_{P,2=L}$ (3)	(1)	(2)	(3)
Year of arrival	1 976	1 974	1 971	193	254	561
Male	0.36	0.41	0.44	193	254	561
Fluency in French	0.35	0.12	0.05	151	218	471
Secondary Education	0.68	0.27	0.06	193	254	553
Higher Education	0.50	0.10	0.01	193	254	553
Occupation Grandfather	Skilled Craft Workers	Skilled Craft Workers	Skilled Craft Workers	193	254	561
Occupation Father	Technical Managers	Skilled Industrial Workers	Skilled Industrial Workers	69	104	249
Occupation Mother	Managerial Occupations (gov)	Workers - direct services	Workers - direct services	124	150	312
Europe						
Year of arrival	1 979	1 981	1 981	92	211	295
Male	0.51	0.46	0.47	92	211	295
Fluency in French	0.35	0.15	0.06	85	202	286
Secondary Education	0.74	0.36	0.06	92	208	269
Higher Education	0.59	0.11	0.01	92	208	269
Occupation Grandfather	Retailers and Related	Retailers and Related	Farmers	92	211	295
Occupation Father	Technical Managers	Skilled Industrial Workers	Skilled Industrial Workers	47	98	140
Occupation Mother	No activity	No activity	No activity	45	113	155
Asia						
Year of arrival	1 977	1 979	1 977	78	124	486
Male	0.56	0.43	0.37	78	124	488
Fluency in French	0.73	0.75	0.38	60	112	428
Secondary Education	0.61	0.24	0.07	77	121	397
Higher Education	0.57	0.12	0.03	77	121	397
Occupation Grandfather	Skilled Craft Workers	Retailers and Related	Unskilled Industrial Workers	78	124	488
Occupation Father	Technical Managers	Skilled Industrial Workers	Skilled Industrial Workers	44	53	180
Occupation Mother	No activity	Workers - direct services	No activity	34	71	308
Maghreb						
Year of arrival	1 982	1 984	1 983	61	141	175
Male	0.54	0.35	0.35	61	141	175
Fluency in French	0.94	0.79	0.38	50	120	158
Secondary Education	0.80	0.53	0.20	60	136	107
Higher Education	0.62	0.32	0.08	60	136	107
Occupation Grandfather	Managerial Occupations (gov)	Managerial Occupations (gov)	Farmers	61	141	175
Occupation Father	Workers - direct services	Skilled Industrial Workers	Skilled Industrial Workers	0	50	62
Occupation Mother	Workers - direct services	Civilian Employees	Workers - direct services	28	91	113
Africa						

Note : The Observations are parents of second generation immigrants (sample 1) who are at least 18, were born in France or arrived before age 10. High and Low refer to the general definition used for sample1

Once broken by region of origin and pre-migration status (based on a definition that is not country specific and which show a great deal of variation between developed and developing countries), the picture of immigrants look strikingly similar in terms of education of the parents (recall that the definition of high and low looked at grandparents education and not at parents' education). The only exception being African immigrants being more educated. It shows that the definition captures well parental background. The European immigration is older and more masculine.

The populations of $S_{P,0} = H, S_{P,2} = L$ and $S_{P,0} = L, S_{P,2} = L$ have features that fit the proposed story: they have different characteristics in France, the mode of occupations is similar for men and women when broken down by region of origins (for men the mode is skilled industrial workers in all cases but for women the mode of occupation is the same for Europe and Asia), however they have different characteristics for elements that are related to pre-migration status, namely their education, their fluency in French upon arrival and their fathers' occupation.

The population of $S_{P,0} = H, S_{P,2} = L$ is however different from that of $S_{P,0} = H, S_{P,2} = H$ in most dimensions in France and in the country of origin (education and fluency in French which is likely to be correlated to social status in former French colonies). It shows that the definition of high and low in France and in the country of origin is able to identify a population in between the very high and the low strata of the population in the origin country.

3.2 Results from transition matrices

For convenience, I only report transition matrices for sample 1. For this sample, I only look at one outcome variable which at age 18, has the child successfully completed high school (obtained the baccalauréat) or obtained a higher education degree¹³. If yes, then $S_C = H$, if no $S_C = L$.

The differences in achievements are very big in both tables 9 and 10. For comparison with natives, the last row of table 9 reports distribution among natives. There is a 14 percentage points difference when I look at the entire population (meaning that $S_{P,0} = H, S_{P,2} = L$ are almost 30 percent more likely to successfully finish high school than are $S_{P,0} = L, S_{P,2} = L$). The difference reaches 20 percentage points when one focuses on people living in urban sensitive area, meaning that children from $S_{P,0} = H, S_{P,2} = L$ are fifty percent more likely to complete high school or get a higher education degree than children from $S_{P,0} = L, S_{P,2} = L$.

To put these numbers into perspective, I will detail in subsection 3.4 how much of a difference in monetary resources available to the household has to exist in the population of native French to observe similar differences in probabilities of educational success.

¹³Another potential outcome variable available in TeO is being employed at the time of the interview. For this outcome, all immigrants from a low background in France do equally poorly. I choose not to include this outcome for the following reason, by definition, being employed is influenced by business cycle or by discrimination and is not a life cycle type variable as are education achievements or the type of occupation when employed. It is a very interesting question to understand why the difference is so sharp for outcomes other than unemployment, however it is a different question than the one this paper aims at answering to.

Table 9: Transition matrix - Children and Before Migration status

	Absolute Numbers			Percentages		
	$S_C=L$	$S_C=H$	Total	$S_C=L$	$S_C=H$	Total
$S_{P,0}=L$ $S_{P,2}=L$	1 814	1 705	3 519	51.55	48.45	71.99
$S_{P,0}=H$ $S_{P,2}=L$	515	854	1 369	37.62	62.38	28.01
Total	2 329	2 559	4 888	47.65	52.35	
Among Natives	740	1 189	1 929	38.36	61.64	

Note : The Observations in the first three rows are second generation immigrants (from sample 1) who are at least 18 were born in France or arrived before age 10. High and Low refer to the general definition used for sample 1. the observations in the last row are children of native French

Table 10: Transition matrix - Children and Before Migration status - Urban Sensitive Area

	Absolute Numbers			Percentages		
	$S_C=L$	$S_C=H$	Total	$S_C=L$	$S_C=H$	Total
$S_{P,0}=L$ $S_{P,2}=L$	586	402	988	59.31	40.69	79.10
$S_{P,0}=H$ $S_{P,2}=L$	104	157	261	39.85	60.15	20.90
Total	690	559	1 249	55.24	44.76	

Note : The Observations are second generation immigrants (from sample 1) who are at least 18 were born in France or arrived before age 10. Parents live in Urban Sensitive Areas High and Low refer to the general definition used for sample 1

3.3 Results from linear regressions

3.3.1 From the first subsample

I estimate a linear probability model on a subpopulation of $S_{P,2} = L$ second generation immigrants :

$$y_i = \alpha + \beta 1\{S_{P,0} = H\} + \gamma X + \varepsilon_i \quad (1)$$

where $1\{\}$ is an indicator function taking value 1 when $S_{P,0} = H$ and 0 otherwise. The coefficient of interest β expresses the difference $E[Y|S_{P,0}=H, S_{P,2}=L] - E[Y|S_{P,0}=L, S_{P,2}=L]$, X is a vector of controls that include age, gender, a dummy for each potential number of siblings and resources available to the household. Controlling for household resources enables me to

look at parental effect beyond general material well-being of the household¹⁴. It can be argued that number of siblings is more an outcome than a control and thus should not be included. I wanted to have per capita available resources which is done by including number of siblings together with available resources. As can be guessed from the transition matrices 9 and 10,

the results without siblings and without household resources tend to be higher. What is most important is that including these very relevant variables do not make the coefficient of interest disappear. Since I observe multiple children from the same household, I cluster the standard errors at the family level.

Table 11: Digging into pre migration status

	BAC - Baseline	BAC - Country FE	BAC - Maghreb	BAC - African
β	0.10*** [0.05,0.16]	0.10*** [0.04,0.16]	0.14*** [0.04,0.24]	0.11 [-0.03,0.25]
Control Age	Yes	Yes	Yes	Yes
Control Gender	Yes	Yes	Yes	Yes
Country FE	No	Yes	No	No
Nb Sibling	Yes	Yes	Yes	Yes
Ressources Household	Yes	Yes	Yes	Yes
R-Squarred	0.05	0.08	0.07	0.14
Nb of Observations	3082	3082	1237	356
Nb of High	710	710	189	139
Nb of Low	2372	2372	1048	217
Nb of clusters	1260	1260	418	138
Unconditional Mean	0.50	0.50	0.49	0.45

The observations are children of first generation immigrants (sample 1) that are at least 18, who grew up in France or arrived before age 10. High and Low refer to the general definition used for sample 1. Standard Errors are clustered at the household level

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In table 11, besides the baseline specification, I look at the inclusion of country of origin fixed effects and at subsamples where I limit observations to immigrants from Maghreb and Sub-Saharan Africa. The results don't change for specific regions of origin and are somewhat lower

¹⁴In a sense, this conditioning is redundant with estimating the difference on a sample of $S_{P,2} = L$ however I want to account for the potential differences that can exist within this category to make sure that elements related to the material situation in France is accounted for.

but still quantitatively very big when fixed effects are included. I also report the number of observations with $S_{P,0} = H$ and $S_{P,0} = L$ to give an idea of how much the population I am focusing on represent in the population of low background immigrants. Overall they represent 28% of this population, up to 40% for immigrants with African origin and 15% for immigrants with North African origins. I report the number of clusters and the unconditional mean for all $S_{P,2} = L$ to given an idea of the magnitude of the difference.

In table 12, I perform the same regressions but on a subsample of people living in Urban Sensitive Area. In accordance with the results from the transition matrices, the effect is larger for this population. As the sample size shrinks, the standard errors increase but they remain significant in three out of four specifications. For North Africa the estimate is not significantly different from zero anymore but the point estimate hasn't changed. There are few observations with $S_{P,0} = H$ in this subsample which explain why the estimates are not very precise.

Table 12: Digging into pre migration status - Urban Sensitive Area

	BAC - Baseline	BAC - Country FE	BAC - Maghreb	BAC - African
β	0.17*** [0.04,0.30]	0.22*** [0.08,0.35]	0.15 [-0.05,0.36]	0.13 [-0.13,0.38]
Control Age	Yes	Yes	Yes	Yes
Control Gender	Yes	Yes	Yes	Yes
Country FE	No	Yes	No	No
Nb Sibling	Yes	Yes	Yes	Yes
Ressources Household	Yes	Yes	Yes	Yes
R-Squarred	0.06	0.10	0.09	0.15
Nb of Observations	877	877	484	132
Nb of High	143	143	54	45
Nb of Low	734	734	430	87
Nb of clusters	295	295	148	43
Unconditional Mean	0.42	0.42	0.46	0.30

The observations are children of first generation immigrants (sample 1) that are at least 18, who grew up in France or arrived before age 10. Parents live in Urban Sensitive Areas High and Low refer to the general definition used for sample 1.

Standard Errors are clustered at the household level

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.3.2 From the second subsample

Sample 2 has more outcome variables of interest, (i) on education such as dropping out, successfully finishing high school, having a higher education degree and (ii) on achievements in the labour market such as occupation and wage. It also allows to include societal outcomes such as having a partner who is a native French.

In table 13, I estimate equation1 with sample 2. The vector of controls include age, gender, the number of siblings and a dummy for each occupation (at a two digit level of disaggregation)¹⁵ of the father and the mother. The difference between the two populations is of a high order of magnitude for all the educational outcomes, being a white collar and the probability of having a partner that is a native French. This last result means that not only do $S_{P,0} = H$, $S_{P,2} = L$

¹⁵To account for the heterogeneity within this category (as was done in the first sample with household resources)

perform better than $S_{P,0} = L$, $S_{P,2} = L$ but also that they have a better social integration. The effect is small and not statistically significant for log of wage.

Table 13: Digging into pre migration status - Narrow Definition - Baseline

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.06*** [0.03,0.08]	0.10*** [0.05,0.15]	0.16*** [0.09,0.22]	0.11** [0.02,0.19]	-0.00 [-0.08,0.08]	0.11*** [0.06,0.16]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squared	0.06	0.09	0.12	0.10	0.33	0.01
Nb of Observations	1979	1979	1191	708	568	2274
Nb of High	754	754	374	232	182	570
Nb of Low	1225	1225	817	476	386	1704
Unconditional Mean	0.89	0.56	0.37	0.37	7.31	0.46

Education outcomes: Both parents immigrants, not studying, currently low

Labour market outcomes : Males in the labour force, both parents immigrants, currently low

Social outcomes : Both parents immigrants, currently low

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Having two samples within the same survey allows to do an out of sample check. The only comparable outcome is BAC and the estimates in table 13 are in line with those of table 11, suggesting the difference observed between the two groups is not due to the specific sample used.

3.4 Is it possible to quantify the money equivalent of parents from different social background?

Since for sample 1, I have data on the resources available to the household, I am able to calculate for natives, what income groups have an average high school graduation rate of 48% (as the $S_{P,0} = H$, $S_{P,2} = L$) and which group has an average graduation rate of 62% (as the $S_{P,0} = L$, $S_{P,2} = L$).

To observe such a difference among natives, one has to compare households that have an average available resources of below 2600€ (success rate of 48%) with those that have resources from 2600€ to 3300€. The first group has median resources of 1800€¹⁶ and the second of 3000€^{17, 18}. So having parents from a higher social stratum in the country of origin for children of immigrants has an effect equivalent to being 2/3 richer for low background natives.

4 Can we identify the channels of parental transmission?

4.1 Some hypotheses to test

If children of $S_{P,0} = H$, $S_{P,2} = L$ do so much better than children of $S_{P,0} = L$, $S_{P,2} = L$, what are the non-monetary elements that their parents transmit them? Is it that these children

¹⁶which corresponds to the 23rd percentile of the distribution among natives

¹⁷which corresponds to the 55th percentile of the distribution among natives

¹⁸When the median for $S_{P,0} = H$, $S_{P,2} = L$ is 2000€ and for $S_{P,0} = L$, $S_{P,2} = L$ 1860€

are brought up in better environments and are exposed to higher-achieving peers? Is it that parents invest more time or resources in their children’s schooling? Is it that parents transmit different behaviors such as work ethic or patience?

Sample 2 of the data allows me to isolate potential channels and test several hypotheses. One that has already indirectly been tested in the specifications presented above is the quality of the neighborhood. Since the difference between $S_{P,0} = H$, $S_{P,2} = L$ and $S_{P,0} = L$, $S_{P,2} = L$ is highest in Urban Sensitive Areas, it is not just by moving to a ‘better’ neighborhood that parents improve the prospects of their children.

Another series of hypotheses that the data allows me to test (mostly about differences in educational outcomes) are: (i) is it parents’ schooling strategy?, (ii) is it parents’ time investment in their children’s schooling? , (iii) is it the home’s study conditions that parents provide? For each of these three potential channels, there is a question in the survey that can be used, i.e. do parents send their children to a school outside their school district? did parents spend time helping their children doing their homework? did the children have access to a room where they would be alone to study? A set of hypotheses I cannot test are parental transmission of aspiration or of behaviors (such as patience, attitude towards risk)¹⁹.

The way I test these hypotheses is simply to add an extra variable in a linear regression like equation 1 and look at how the coefficient of the difference (β) changes. One question is whether I should include these variables as extra controls or as outcome variables. Below I report both the results when adding a variable in a regression and when looking at how much parents differ in their answers to the questions above.

4.2 The effect of parental background is unchanged by parental investment

There is evidence that $S_{P,0} = H$, $S_{P,2} = L$ are more likely to send their children to a school outside their district than $S_{P,0} = L$, $S_{P,2} = L$ (20% instead of 14%), are more likely to help their children (20% instead of 5%) and to have a place to study alone (77% instead of 60%), which suggest that there is a difference in behavior from the parents.

However, when these variables are included separately in linear regressions, as in tables 14, 15 and 16 the coefficients giving the difference in conditional probabilities don’t change at all. The extra coefficients are statistically significant and have the expected sign for the expected outcomes, however they leave almost unchanged the difference that I am trying to explain. This means that among those that are being sent to a school outside their school district because the parents wanted their children to study in a better environment, there is still a 13 percentage points difference in the probability of successfully finishing high school between $S_{P,0} = H$, $S_{P,2} = L$ and $S_{P,0} = L$, $S_{P,2} = L$. So parents strategy helps but there is something else on top that is not explained by parents schooling strategy, nor by parental investment

¹⁹In addition to the data sources that I detailed, TeO has an extra survey that was filled by children of immigrants (sample 1), where some of the questions are related to educational aspirations. However this survey was only targeted at children aged 15 to 24 living at their parents and was voluntary (it was left after the interview to be filled and sent back). As a result, there is a huge attrition that is not likely to be random. For instance, out the 1738 observations eligible (from the baseline specification of table 11), only 511, so less than 30%, filled this survey. Because of this strong attrition, I don’t use this extension of the survey.

in helping their children with their homework, nor by their study condition at home that determine their children's achievements.

Table 14: Potential explanations - Schooling Strategy

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.05***	0.10***	0.15***	0.10**	-0.00	0.10***
	[0.03,0.08]	[0.05,0.14]	[0.09,0.22]	[0.02,0.19]	[-0.08,0.07]	[0.06,0.15]
Strategy	0.01	0.09***	0.09***	0.03	0.05	0.00
	[-0.02,0.05]	[0.03,0.14]	[0.02,0.17]	[-0.06,0.13]	[-0.03,0.13]	[-0.05,0.06]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Job Mother	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squared	0.06	0.09	0.13	0.10	0.34	0.02
Nb of Observations	1979	1979	1191	708	568	2274
Nb of High	754	754	374	232	182	570
Nb of Low	1225	1225	817	476	386	1704
Unconditional Mean	0.89	0.56	0.37	0.37	7.31	0.46

The population is second generation immigrants (sample 2) that are at least 18 (for drop out and bac) and 25 for higher education.

For the partner outcome there is no restriction on age. High and Low refer to the large definition of sample 2.

All the immigrants are low in France. Standard errors are robust to heteroskedasticity

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Potential explanations - Help for howework

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.05***	0.09***	0.15***	0.11**	-0.00	0.10***
	[0.02,0.08]	[0.05,0.14]	[0.08,0.21]	[0.02,0.19]	[-0.08,0.08]	[0.05,0.15]
Help	0.02	0.05	0.11**	-0.01	0.03	0.02
	[-0.02,0.05]	[-0.01,0.11]	[0.01,0.21]	[-0.12,0.11]	[-0.07,0.12]	[-0.06,0.09]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Job Mother	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squared	0.06	0.09	0.13	0.10	0.33	0.02
Nb of Observations	1965	1965	1184	704	565	2226
Nb of High	753	753	374	232	182	570
Nb of Low	1212	1212	810	472	383	1656
Unconditional Mean	0.89	0.56	0.37	0.38	7.31	0.46

The population is second generation immigrants (sample 2) that are at least 18 (for drop out and bac) and 25 for higher education.

For the partner outcome there is no restriction on age. High and Low refer to the large definition of sample 2.

All the immigrants are low in France. Standard errors are robust to heteroskedasticity

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.3 How to interpret these results?

I am only able to list a series of hypothesis, test those for which I have enough data (i.e. the relevant questions and a non selected sample). If the hypotheses that are tested are rejected, this points in favor of those that cannot be tested but are relevant. Unfortunately I can't test these. Parents can help their children in different ways, provide them with a good environment,

Table 16: Potential explanations - Room to study

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.05***	0.10***	0.15***	0.11**	0.00	0.10***
	[0.02,0.08]	[0.05,0.14]	[0.09,0.22]	[0.02,0.20]	[-0.08,0.08]	[0.06,0.15]
Room alone to study	0.04**	0.06**	0.02	-0.07	-0.05	-0.00
	[0.01,0.08]	[0.01,0.11]	[-0.03,0.08]	[-0.15,0.01]	[-0.11,0.02]	[-0.05,0.04]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Job Mother	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squared	0.06	0.09	0.12	0.11	0.34	0.02
Nb of Observations	1979	1979	1191	707	567	2270
Nb of High	754	754	374	232	182	570
Nb of Low	1225	1225	817	475	385	1700
Unconditional Mean	0.89	0.56	0.37	0.37	7.31	0.46

The population is second generation immigrants (sample 2) that are at least 18 (for drop out and bac) and 25 for higher education.

For the partner outcome there is no restriction on age. High and Low refer to the large definition of sample 2.

All the immigrants are low in France. Standard errors are robust to heteroskedasticity

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

help them directly or influence their behavior by acting as raw models or transmitting positive values and habits. The former is captured in the data and can be tested, not the latter.

It has been shown theoretically in Zilibotti and Doepke (2012) and empirically in Dohmen et al. (2012) that behavioral transmission from parents to children is crucial for their achievements. It can be that these are the relevant channels for immigrant parents. These are hypotheses that could be tested in the future on the population of second generation immigrants.

5 Robustness checks

5.1 Different definitions of high and low and different subsamples

For sample 1, I look at four combinations: the two components of the definition (education and socio-professional) and two subsamples, the entire one and the people living in Urban Sensitive Area. The results remain statistically significant and of important magnitude in all settings but for subpopulation living in Urban Sensitive Area that have parents that are high according to the socio-professional definition. The strong effect observed on people living in Urban Sensitive Area is driven by those having high parents based on educational achievements of their family. These correspond to tables 22, 23, 24, 25

For sample 2, I look at eight combinations : four with the narrow definition and four with the large definition. For the narrow definitions, I look at a specification with country fixed effects, on immigrants from North Africa or Sub-Saharan Africa and on immigrants who went to a junior high school where they were more than half of the students children of immigrants (I use this specification to focus on environments with high concentration of migrants). These correspond to tables 29, 30, 31, 33. The effects are somewhat smaller for Maghreb and less precisely estimated for Maghreb and Sub-Saharan Africa but doesn't change the picture of the baseline model. For the large definition, I estimate the baseline model, one with country

of origin fixed effects, with immigrants from North Africa or Sub-Saharan Africa. In all cases, the difference is strong and significant. These correspond to the tables 34, 35, 36, ???. In all these cases, the effects remain of an important order of magnitude and statistically significant.

Results are robust to the inclusion of frequency weights. These correspond to the last two columns of table 26 and table 32 sample 2, the narrow definition and table 38 for sample 2, large definition.

In table 28, I estimate

$$y_i = \alpha + \beta 1\{S_{P,0} = H\} + \gamma 1\{S_{P,0} = H, Educ\} + \theta X + \varepsilon_i \quad (2)$$

to check whether the effect observed among those that have a high status according to the socio-professional definition are driven by those that are at the same time high according to the educational definition.

5.2 Different definitions of immigrants

5.2.1 How different are children of one vs two parents immigrants

For sample 1, table 18 and sample 2, ??, I look into the differences between immigrants that married natives and immigrants that came with someone from the same country. I compute the probability (for both parents) to have a secondary education, a higher education and the mode of occupations for each. It appears that families with two parents immigrants have lower education level.

Table 17: Descriptive statistics : One vs Both parents are foreigners

		Both parents Percentages	Nb of Obs	Only One parent Percentages	Nb of Obs
Men	Secondary Education	0.19	960	0.10	10
	Higher Education	0.08	960	0.00	10
	Occupation	Skilled Industrial Workers	1,002	Skilled Industrial Workers	10
	Secondary Education - Partner	0.17	901	0.22	9
	Higher Education - Partner	0.06	901	0.11	9
	Occupation - Partner	Workers - Direct Services	1,002	Workers - Direct Services	10
Women	Secondary Education	0.17	1,246	0.47	19
	Higher Education	0.07	1,246	0.32	19
	Occupation	Inactiv	1,433	Civilian Employee	19
	Secondary Education - Partner	0.16	1,050	0.25	8
	Higher Education - Partner	0.08	1,050	0.25	8
	Occupation - Partner	Skilled Craft Workers	1,433		0

Note : The Observations are first generation immigrants (sample 1) who have at least one child who is at least 18 and was born in France or arrived before age 10.

Table 18: Descriptive statistics : One vs Both parents are foreigners

	Both parents %	Nb of Obs	Only Mother %	Nb of Obs	Only Father %	Nb of Obs
Secondary E	0.07	3,787	0.15	770	0.13	1,355
Higher E	0.02	3,787	0.07	770	0.07	1,355
Occupation	Unskilled Industrial Workers	4,162	Skilled Craft Workers	917	Skilled Craft Workers	1,579
Secondary E - Partner	0.06	3,939	0.18	837	0.18	1,452
Higher E - Partner	0.02	3,939	0.08	837	0.09	1,452
Occupation - Partner	Workers - Direct Services	4,162	Workers - Direct Services	917	Workers - Direct Services	1,579

Note : The Observations are second generation immigrants (sample 2) who are at least 18

5.2.2 Is the effect different when immigrants are defined as having one parent non native?

As can be seen from 19, they are surprisingly similar. This is confirmed for sample 1 in table ?? but it doesn't say much since very observations don't have both parents immigrants in that sample.

Table 19: Digging into pre migration status - Narrow Definition - One parent immigrant

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.08*** [0.04,0.11]	0.12*** [0.07,0.18]	0.10*** [0.04,0.15]	0.11** [0.02,0.19]	-0.00 [-0.08,0.08]	0.04* [-0.00,0.08]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squared	0.12	0.19	0.15	0.10	0.33	0.00
Nb of Observations	1962	1962	1687	708	568	1659
Nb of High	1359	1359	1135	232	182	1054
Nb of Low	603	603	552	476	386	605
Unconditional Mean	0.90	0.60	0.37	0.37	7.31	0.77

Education outcomes: One parent immigrant, not studying, currently low

Labour market outcomes : Males in the labour force, one parent immigrant, currently low

Social outcomes : One parent immigrant, currently low.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

There is no systematic differences in behaviors.

5.2.3 Is the effect of parental background different between having one or two parents immigrants?

In table 20, I estimate

$$y_i = \varsigma + \beta 1\{S_{P,0} = H\} + \beta_2 1\{Mother\} + \beta_3 1\{Father\} + \alpha 1\{S_{P,0} = H\} 1\{Mother\} + \gamma 1\{S_{P,0} = H\} 1\{Father\} + \theta X + \varepsilon_i \quad (3)$$

Where high and low is now defined for second generation immigrants having one or both parents immigrants, the terms α and β are differences-in-differences between having your mother or your mother immigrant vs having both your parents immigrants.

Table 20: Digging into pre migration status - Narrow Definition - Differences between one and two parent immigrants

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.04*** [0.02,0.07]	0.09*** [0.05,0.14]	0.15*** [0.09,0.21]	0.09** [0.01,0.17]	0.00 [-0.07,0.08]	0.11*** [0.06,0.16]
α	0.04 [-0.03,0.11]	0.02 [-0.08,0.12]	-0.06 [-0.18,0.05]	0.08 [-0.07,0.24]	0.02 [-0.11,0.15]	-0.16*** [-0.24,-0.07]
γ	0.06** [0.01,0.11]	0.05 [-0.03,0.12]	-0.07 [-0.16,0.02]	-0.06 [-0.18,0.06]	0.05 [-0.06,0.17]	-0.04 [-0.11,0.04]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squared	0.07	0.11	0.13	0.09	0.31	0.10
Nb of Observations	3590	3590	2344	1386	1122	3659
Nb of High	1791	1791	1033	655	534	1377
Nb of Low	1799	1799	1311	731	588	2282
Unconditional Mean	0.89	0.56	0.36	0.38	7.34	0.58

Education outcomes: One or two parent immigrants, not studying, currently low

Labour market outcomes : Males in the labour force, one or two parent immigrants, currently low

Social outcomes : One or two parent immigrants, currently low.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5.3 Do we observe a similar pattern among natives?

5.3.1 On long term persistence

As can be seen in Chan and Bolivier (2013); Mare (2011), grandparents have a direct effect on the achievements of their grandchildren so it makes sense that having high performing grandparents, even when the parents are not, generate a difference among natives. I observe also an effect in the data²⁰. However it is likely to go through different channels than what is observed among immigrants. The family of low background immigrants in France is not likely to be able to transfer assets (when native grandparents can do it more easily) and in addition most of the grandparents don't even live in France so they can hardly have a direct effect. For instance 31% of the grandfathers still alive in our sample 1 live in France, when for natives 96% do. So if there is a grandparents effect, it is qualitatively very different.

5.3.2 On inter-generational mobility

Table 21 presents the transition matrices (using observations from sample 1) of three different populations: natives, all immigrants and immigrants with a low background. When comparing natives with immigrants, one can see that although the marginal distributions are different, the transition dynamics are very similar. However when comparing the two bottom part of the table, one can see that this similarity in dynamics hides a lot of heterogeneity as was discussed in the paper.

²⁰The definition of high and low had to be changed from finishing primary school to finishing junior high school as there would be too few observations otherwise.

Table 21: Transition matrix - Comparing the dynamics between natives and immigrants

	Absolute Numbers			Percentages		
	$S_C=L$	$S_C=H$	Total	$S_C=L$	$S_C=H$	Total
$S_{P,2}=L$	579	643	1 222	47.38	52.62	63.35
$S_{P,2}=H$	161	546	707	22.77	77.23	36.65
Total	740	1 189	1 929	38.36	61.64	
$S_{P,2}=L$	2 329	2 559	4 888	47.65	52.35	80.85
$S_{P,2}=H$	286	872	1 158	24.70	75.30	19.15
Total	2 615	3 431	6 046	43.25	56.75	
$S_{P,0}=L$ $S_{P,2}=L$	1 814	1 705	3 519	51.55	48.45	71.99
$S_{P,0}=H$ $S_{P,2}=L$	515	854	1 369	37.62	62.38	28.01
Total	2 329	2 559	4 888	47.65	52.35	

Note : The Observations in the first 3 rows are children of French natives that are at least 18. The observations in the 3 rows in the middle are second generationn immigrants (from sample 1) who are at least 18 were born in France or arrived before age 10. The observations in the last 3 rows are a subset of those in the middle 3 rows namely with $S_{P,2}=L$. High and Low refer to the general definition used for sample 1

6 Conclusion

In this paper, I look at whether pre migration status of the family is an important element to understand the education and labor market achievements of second generation immigrants. I find that this dimension that is not well captured in most data and not usually considered in the literature is of great importance. The life prospects of the two groups of previously high and currently low and previously and currently low are as different as low class versus middle natives. When I dig into the parental investment, I am unable to find an explanation for this difference. I point towards some more hypotheses to be explored in further research.

References

- Aydemir, A., Chen, W.-H., and Corak, M. (2009). Intergenerational Earnings Mobility among the Children of Canadian Immigrants. *The Review of Economics and Statistics*, 91(2):377–397.
- Barro, R. J. and Lee, J. W. (2013). A new data set of educational attainment in the world, 1950-2010. *Journal of Development Economics*, 104(C):184–198.
- Beauchemin, C., Hamel, C., and Héran, F. (2015). Trajectoires et Origines : enquête sur la diversité des populations en France. Technical report.
- Bisin, A. and Verdier, T. (2010). The Economics of Cultural Transmission and Socialization. In Jess Ben-habib, A. B. and Jackson, M., editors, *Handbook of Social Economics*, volume 1A of *Handbook of Social Economics*, chapter 9, pages 339–416. Elsevier.
- Black, S. E. and Devereux, P. J. (2011). *Recent Developments in Intergenerational Mobility*, volume 4 of *Handbook of Labor Economics*, chapter 16, pages 1487–1541. Elsevier.
- Borjas, G. J. (1992). Ethnic Capital and Intergenerational Mobility. *The Quarterly Journal of Economics*, 107(1):123–50.
- Borjas, G. J. (1993). The Intergenerational Mobility of Immigrants. *Journal of Labor Economics*, 11(1):113–35.
- Chan, T. W. and Bolivier, V. (2013). The Grandparents Effect in Social Mobility: Evidence from British Cohort Studies. *American Sociological Review*, 92(1):662–679.
- Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2012). The Intergenerational Transmission of Risk and Trust Attitudes. *Review of Economic Studies*, 79(2):645–677.
- Fogli, A. and Fernandez, R. (2009). Culture: An Empirical Investigation of Beliefs, Work, and Fertility. *American Economic Journal: Macroeconomics*, 1(1):146–77.
- Lindahl, M., Palme, M., Massih, S. S., and Sjagren, A. (2015). Long Term Intergenerational Persistence of Human Capital : An Empirical Analysis of Four Generations. *Journal of Human Resources*, 50(1):1–33.
- Mare, R. (2011). A Multigenerational View of Inequality. *Demography*, 48(1):1–23.
- Waters, M. C. (2009). *Black identities*. Harvard University Press.
- Zilibotti, F. and Doepke, M. (2012). Intergenerational Transmission of Risk Preferences, Entrepreneurship, and Growth. Technical report.

7 Appendix

7.1 For sample 1

7.1.1 Variation in the definition of immigrants

Table 22: Digging into pre migration status

	BAC - Baseline	BAC - Country FE	BAC - Maghreb	BAC - African
β	0.08**	0.05	0.05	0.14**
	[0.01,0.15]	[-0.02,0.12]	[-0.07,0.18]	[0.00,0.29]
Control Age	Yes	Yes	Yes	Yes
Control Gender	Yes	Yes	Yes	Yes
Country FE	No	Yes	No	No
Nb Sibling	Yes	Yes	Yes	Yes
Ressources Household	Yes	Yes	Yes	Yes
R-Squared	0.04	0.07	0.06	0.15
Nb of Observations	3082	3082	1237	356
Nb of High	303	303	99	83
Nb of Low	2779	2779	1138	273
Nb of clusters	1260	1260	418	138
Unconditional Mean	0.50	0.50	0.49	0.45

The observations are children of first generation immigrants (sample 1) that are at least 18, who grew up in France or arrived before age 10. High and Low refer to the socio-professional definition used for sample 1. Standard Errors are clustered at the household level

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 23: Digging into pre migration status

	BAC - Baseline	BAC - Country FE	BAC - Maghreb	BAC - African
β	0.09***	0.09***	0.15**	0.09
	[0.03,0.15]	[0.03,0.15]	[0.02,0.27]	[-0.05,0.23]
Control Age	Yes	Yes	Yes	Yes
Control Gender	Yes	Yes	Yes	Yes
Country FE	No	Yes	No	No
Nb Sibling	Yes	Yes	Yes	Yes
Ressources Household	Yes	Yes	Yes	Yes
R-Squared	0.05	0.07	0.07	0.14
Nb of Observations	3082	3082	1237	356
Nb of High	567	567	131	119
Nb of Low	2515	2515	1106	237
Nb of clusters	1260	1260	418	138
Unconditional Mean	0.50	0.50	0.49	0.45

The observations are children of first generation immigrants (sample 1) that are at least 18, who grew up in France or arrived before age 10. High and Low refer to the education definition used for sample 1. Standard Errors are clustered at the household level

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 24: Digging into pre migration status - Urban Sensitive Area

	BAC - Baseline	BAC - Country FE	BAC - Maghreb	BAC - African
β	-0.01	-0.00	-0.02	0.16
	[-0.18,0.16]	[-0.18,0.17]	[-0.30,0.25]	[-0.12,0.43]
Control Age	Yes	Yes	Yes	Yes
Control Gender	Yes	Yes	Yes	Yes
Country FE	No	Yes	No	No
Nb Sibling	Yes	Yes	Yes	Yes
Ressources Household	Yes	Yes	Yes	Yes
R-Squarred	0.05	0.07	0.08	0.15
Nb of Observations	877	877	484	132
Nb of High	64	64	29	22
Nb of Low	813	813	455	110
Nb of clusters	295	295	148	43
Unconditional Mean	0.42	0.42	0.46	0.30

The observations are children of first generation immigrants (sample 1) that are at least 18, who grew up in France or arrived before age 10. Parents live in Urban Sensitive Areas High and Low refer to the socio-rprofessional definition used for sample 1.

Standard Errors are clustered at the household level

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 25: Digging into pre migration status - Urban Sensitive Area

	BAC - Baseline	BAC - Country FE	BAC - Maghreb	BAC - African
β	0.21***	0.26***	0.15	0.15
	[0.08,0.35]	[0.12,0.41]	[-0.07,0.38]	[-0.11,0.42]
Control Age	Yes	Yes	Yes	Yes
Control Gender	Yes	Yes	Yes	Yes
Country FE	No	Yes	No	No
Nb Sibling	Yes	Yes	Yes	Yes
Ressources Household	Yes	Yes	Yes	Yes
R-Squarred	0.07	0.10	0.09	0.15
Nb of Observations	877	877	484	132
Nb of High	122	122	46	42
Nb of Low	755	755	438	90
Nb of clusters	295	295	148	43
Unconditional Mean	0.42	0.42	0.46	0.30

The observations are children of first generation immigrants (sample 1) that are at least 18, who grew up in France or arrived before age 10. Parents live in Urban Sensitive Areas High and Low refer to the education definition used for sample 1.

Standard Errors are clustered at the household level

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 26: Robustness Checks

	Narrow Definition Baseline	Large Definition Baseline	Narrow Definition Fixed Effects	Large Definition Fixed Effects	Weights Baseline	Weights ZUS
$\beta_{_1}$	0.11*** [0.06,0.16]		0.11*** [0.06,0.16]			
$\beta_{_2}$		0.10*** [0.07,0.14]		0.10*** [0.05,0.15]		
β					0.11*** [0.11,0.11]	0.17*** [0.17,0.18]
Control Age	Yes	Yes	Yes	Yes	Yes	Yes
Control Gender	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	Yes	Yes	No	No
Nb Sibling	Yes	Yes	Yes	Yes	Yes	Yes
Ressources Household	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.05	0.05	0.09	0.09	0.06	0.08
Nb of Observations	3544	3544	3544	3544	709946	196141
Nb of High	994	994	994	994	1010	203
Nb of Low	2550	2550	2550	2550	2651	830
Nb of clusters	1490	1490	1490	1490	1522	363
Unconditional Mean	0.53	0.53	0.53	0.53	0.52	0.45

The observations are children of first generation immigrants (sample 1) that are at least 18, who grew up in France or arrived before age 10.

High and Low refer to the definition used for sample 2. Fixed effects refer to country of origins

ZUS refers to Urban Sensitive Areas. Standard Errors are clustered at the household level

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 27: When including average level of schooling as a control?

	Baseline
β	0.17*** [0.04,0.30]
Barro and Lee	-0.01 [-0.05,0.03]
Control Age	Yes
Control Gender	Yes
Nb Sibling	Yes
Ressources Household	Yes
R-Squared	0.06
Nb of Observations	877
Nb of High	143
Nb of Low	734
Nb of clusters	295
Unconditional Mean	1.00

The observations are children of first generation immigrants (sample 1 with one or two parents) that are at least 18 who grew up in France or arrived before age 10. High and Low refer to the definition used for sample 1.

ZUS refers to Urban Sensitive Areas. Standard Errors are clustered at the household level

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 28: Is the effect entirely through education?

	Baseline	ZUS
β	0.13*** [0.05,0.21]	-0.01 [-0.19,0.16]
γ	-0.01 [-0.09,0.07]	0.23** [0.03,0.42]
Control Age	Yes	Yes
Control Gender	Yes	Yes
Nb Sibling	Yes	Yes
Ressources Household	Yes	Yes
R-Squarred	0.05	0.07
Nb of Observations	3661	1033
Nb of High	1010	203
Nb of Low	2651	830
Nb of clusters	1522	363
Unconditional Mean	0.52	0.45

The observations are children of first generation immigrants (sample 1) that are at least 18 who grew up in France or arrived before age 10. High and Low refer to the definition used for sample 1. ZUS refers to Urban Sensitive Areas. Standard Errors are clustered at the household level

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

7.1.2 With frequency weights

7.1.3 Including information on mean levels of schooling in the country of origin

7.1.4 Only due to one component of the definition

7.2 For sample 2

7.2.1 Variation in the definition of immigrants

Narrow Definition

Large definition

Table 29: Digging into pre migration status - Narrow Definition - Country Fixed Effects

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.06*** [0.03,0.09]	0.11*** [0.06,0.16]	0.14*** [0.07,0.21]	0.10** [0.01,0.19]	-0.00 [-0.09,0.09]	0.09*** [0.04,0.14]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squarred	0.09	0.13	0.17	0.15	0.38	0.16
Nb of Observations	1979	1979	1191	708	568	2274
Nb of High	754	754	374	232	182	570
Nb of Low	1225	1225	817	476	386	1704
Unconditional Mean	0.89	0.56	0.37	0.37	7.31	0.46

Education outcomes: Both parents immigrants, not studying, currently low

Labour market outcomes : Males in the labour force, both parents immigrants, currently low

Social outcomes : Both parents immigrants, currently low

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 30: Digging into pre migration status - Narrow Definition - Maghreb origin

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.04 [-0.02,0.09]	0.06 [-0.04,0.16]	0.10 [-0.07,0.27]	0.01 [-0.24,0.27]	0.01 [-0.30,0.32]	0.20*** [0.11,0.29]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squarred	0.15	0.16	0.23	0.31	0.54	0.02
Nb of Observations	456	456	261	129	91	829
Nb of High	170	170	81	35	21	144
Nb of Low	286	286	180	94	70	685
Unconditional Mean	0.89	0.62	0.41	0.38	7.27	0.34

Education outcomes: Both parents immigrants, not studying, currently low

Labour market outcomes : Males in the labour force, both parents immigrants, currently low

Social outcomes : Both parents immigrants, currently low

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 31: Digging into pre migration status - Narrow Definition - African origin

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.06 [-0.01,0.13]	0.12 [-0.03,0.27]	0.24* [-0.01,0.49]	0.46*** [0.13,0.78]	0.01 [-0.51,0.54]	0.18** [0.03,0.34]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squarred	0.14	0.26	0.51	0.57	0.75	0.03
Nb of Observations	298	298	112	78	54	193
Nb of High	111	111	34	28	24	52
Nb of Low	187	187	78	50	30	141
Unconditional Mean	0.92	0.57	0.39	0.33	7.07	0.35

Education outcomes: Both parents immigrants, not studying, currently low

Labour market outcomes : Males in the labour force, both parents immigrants, currently low

Social outcomes : Both parents immigrants, currently low

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 32: Digging into pre migration status - Narrow Definition - Frequency Weights

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.07***	0.12***	0.18***	0.11***	0.04***	0.09***
	[0.07,0.07]	[0.12,0.12]	[0.17,0.18]	[0.11,0.12]	[0.04,0.05]	[0.09,0.09]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squared	0.14	0.16	0.18	0.12	0.31	0.01
Nb of Observations	729505	729505	509430	300153	249572	909224
Nb of High	754	754	374	232	182	570
Nb of Low	1225	1225	817	476	386	1704
Unconditional Mean	0.89	0.56	0.37	0.37	7.31	0.46

Education outcomes: Both parents immigrants, not studying, currently low

Labour market outcomes : Males in the labour force, both parents immigrants, currently low

Social outcomes : Both parents immigrants, currently low. I use frequency weights

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 33: Potential explanations - High concentration of immigrants

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.10***	0.16***	0.22***	0.24***	0.07	0.11**
	[0.04,0.16]	[0.06,0.25]	[0.06,0.38]	[0.07,0.42]	[-0.14,0.29]	[0.01,0.21]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Job Mother	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squared	0.14	0.15	0.27	0.29	0.44	0.01
Nb of Observations	545	545	263	168	133	552
Nb of High	199	199	75	65	54	129
Nb of Low	346	346	188	103	79	423
Unconditional Mean	0.86	0.54	0.34	0.35	7.25	0.39

The population is second generation immigrants (sample 2) that are at least 18 (for drop out and bac) and 25 for higher education. For the partner outcome there is no restriction on age.

High and Low refer to the narrow definition of sample 2. All the immigrants are low in France

Standard errors are robust to heteroskedasticity. Immigrants went to a junior high school with predominantly children of immigrants

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 34: Digging into pre migration status - Large Definition - Baseline

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.07***	0.12***	0.13***	0.12***	0.03	0.04**
	[0.04,0.11]	[0.08,0.17]	[0.08,0.19]	[0.04,0.20]	[-0.04,0.10]	[0.00,0.08]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squared	0.07	0.09	0.12	0.11	0.34	0.00
Nb of Observations	1979	1979	1191	708	568	2274
Nb of High	1221	1221	651	400	306	1061
Nb of Low	758	758	540	308	262	1213
Unconditional Mean	0.89	0.56	0.37	0.37	7.31	0.46

Education outcomes: Both parents immigrants, not studying, currently low

Labour market outcomes : Males in the labour force, both parents immigrants, currently low

Social outcomes : Both parents immigrants, currently low

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 35: Digging into pre migration status - Large Definition - Country Fixed Effects

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.08*** [0.04,0.11]	0.12*** [0.07,0.17]	0.12*** [0.06,0.18]	0.12*** [0.04,0.20]	0.03 [-0.04,0.10]	0.04* [-0.00,0.08]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squared	0.09	0.14	0.17	0.15	0.38	0.16
Nb of Observations	1979	1979	1191	708	568	2274
Nb of High	1221	1221	651	400	306	1061
Nb of Low	758	758	540	308	262	1213
Unconditional Mean	0.89	0.56	0.37	0.37	7.31	0.46

Education outcomes: Both parents immigrants, not studying, currently low

Labour market outcomes : Males in the labour force, both parents immigrants, currently low

Social outcomes : Both parents immigrants, currently low

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 36: Digging into pre migration status - Large Definition - Maghreb origin

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.09** [0.01,0.16]	0.12** [0.00,0.23]	0.15* [-0.00,0.29]	0.13 [-0.14,0.39]	0.02 [-0.19,0.24]	0.07** [0.01,0.14]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squared	0.16	0.16	0.24	0.32	0.54	0.01
Nb of Observations	456	456	261	129	91	829
Nb of High	313	313	166	72	46	341
Nb of Low	143	143	95	57	45	488
Unconditional Mean	0.89	0.62	0.41	0.38	7.27	0.34

Education outcomes: Both parents immigrants, not studying, currently low

Labour market outcomes : Males in the labour force, both parents immigrants, currently low

Social outcomes : Both parents immigrants, currently low

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 37: Digging into pre migration status - Large Definition - African origin

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.04 [-0.04,0.13]	0.14** [0.00,0.28]	0.21* [-0.02,0.44]	0.23 [-0.11,0.57]	0.08 [-0.40,0.56]	0.12* [-0.02,0.25]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squared	0.13	0.26	0.50	0.49	0.76	0.02
Nb of Observations	298	298	112	78	54	193
Nb of High	175	175	58	42	30	93
Nb of Low	123	123	54	36	24	100
Unconditional Mean	0.92	0.57	0.39	0.33	7.07	0.35

Education outcomes: Both parents immigrants, not studying, currently low

Labour market outcomes : Males in the labour force, both parents immigrants, currently low

Social outcomes : Both parents immigrants, currently low

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 38: Digging into pre migration status - Large Definition - Frequency Weights

	Not Drop Out	BAC	Higher Education	White Collar	Log Wage	Partner - Native
β	0.07*** [0.07,0.07]	0.14*** [0.14,0.14]	0.14*** [0.14,0.14]	0.10*** [0.10,0.11]	0.03*** [0.03,0.03]	-0.01*** [-0.01,-0.00]
Control Age	Yes	Yes	Yes	Yes	Yes	No
Control Gender	Yes	Yes	Yes	No	No	No
Job Father	Yes	Yes	Yes	Yes	Yes	No
Sibling	Yes	Yes	Yes	Yes	Yes	No
R-Squared	0.14	0.16	0.17	0.12	0.31	0.00
Nb of Observations	729505	729505	509430	300153	249572	909224
Nb of High	1221	1221	651	400	306	1061
Nb of Low	758	758	540	308	262	1213
Unconditional Mean	0.89	0.56	0.37	0.37	7.31	0.46

Education outcomes: Both parents immigrants, not studying, currently low

Labour market outcomes : Males in the labour force, both parents immigrants, currently low

Social outcomes : Both parents immigrants, currently low. I use frequency weights

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$