Data Appendix

Age at Arrival, Parents and Neighborhoods: Understanding the Educational Attainment of Immigrants' Children

Laura Ansala Ulla Hämäläinen Matti Sarvimäki

May 16, 2019

Outcome variables

Education The Finnish education system consists of basic education in comprehensive schools for the whole age group (nine years); upper secondary education, comprising general and vocational education and training (usually three years); and higher education at a university or a polytechnic (also known as universities of applied sciences).

Our first measure of educational attainment is an indicator for having graduated from a general or vocational upper secondary school by the age of 23. A person following the standard curriculum would graduate at age 19. However, graduation is often postponed by switching between different tracks of studies, attending tenth grade after comprehensive school, completing the matriculation examination over multiple semesters, gap years and exchange studies. In our data, almost all individuals who ever complete upper secondary education had graduated by age 23.

Our second measure for educational attainment is an indicator for having been enrolled in college between ages 20 and 23. We use enrolment rather than graduation, because most of those in higher education have not yet completed their studies by the age of 23.

Criminal sentences Our data on criminal activity comes from the Recidivism Register, which contains judgements delivered by district courts and courts of appeal since 1977. Our first measure is an indicator for the person having received a fine from these courts. This does not include on-the-spot fines issued by the police or summary penal judgements (usually fines) and thus the pettiest forms of crime such as littering and minor misdemeanour are excluded from the analysis. According to Marttunen (2006), roughly three quarters of the offences committed by youth are not taken to court but handled through summary penal proceedings. The most common crimes in this category are relatively minor traffic offences.

Our second measure for criminal activity is an indicator for having been sentenced to unconditional or conditional imprisonment or community service by the age of 23. Sentences of imprisonment not exceeding two years may be imposed conditionally, meaning that enforcement is suspended for a probation period of at least one year and at most three years. Community service can be imposed instead of unconditional imprisonment if the sentence does not exceed eight months. Among those sentenced to one of these sanctions in our data (4 % of the sample), three quarters of instances are sentences to either conditional imprisonment (62.6 %) or community service (12 %), and a quarter

to unconditional imprisonment. The most common types of crime leading to community service or conditional imprisonment are traffic offences and offences against property, including aggravated theft, robbery, damage to property, and fraud.

Our measure for the most serious crimes is an indicator for having been senteced to unconditional imprisonment by the age of 23. Almost half of the sentences leading to unconditional imprisonment were due to offences against property.

Medical services Finland provides highly subsidized public health care for all residents. The national health insurance scheme covers everyone living in Finland on a permanent basis and provides reimbursement for e.g. the cost of necessary medicines prescribed by a doctor for the treatment of an illness. The basic rates of reimbursement vary between 65—100 % of the cost of the medicines.

Our first measure of the use of medical services is an indicator for the Social Insurance Institution of Finland having granted the person full reimbursement for her medical costs for treating a severe and long-term illness by age 23. This indicates that the person has a severe illness, as defined in the Health Insurance Act (1224/2004), and that she filed the appropriate application. The most typical diseases covered are diabetes, epilepsy, severe psychosis or other severe mental disorders, different cancers and malign tumours, and behavioural disorders related to mental/intellectual disabilities. However, being entitled to reimbursement does not necessarily mean that the individual has actually used the prescribed medication. Moreover, the interpretation of the outcome may be especially ambiguous among immigrants, because diseases for which the reimbursement is available may be very atypical among some ethnic groups.

Our second health measure is an indicator for having been reimbursed for purchasing psychotropic medication by age 23. Psychotropic medication include antipsychotics, neurosis medication, sleep medication, antidepressants, and central nervous system stimulants. In contrast to medication for a severe illness, it is highly likely that the individual actually has used the reimbursed psychotropic medication, since this reimbursement is received only after purchasing the medication.

Control variables

Parents' months of unemployment We control for parents' months of unemployment using 14 indicator variables. The data on unemployment is constructed differently in different years. Before year 2005, each calendar month is considered to be a full month of unemployment, if it had at least 16 unemployment days. Since 2005, the measurement is based on the number of unemployment days, with 1-14 days registered as zero months of unemployment, 15-44 days as one month of unemployment, and so on. The first category of our control variable consists of parents with missing months of unemployment, i.e. parents who were both employed for the whole year. The next category contains parents with zero months of unemployment, and the remaining 12 categories consist of the averages of their combined months of employment (0-1 months, 1-2 months and so on).

Income decile of the sum of parents' labor income We control for parents' income deciles using ten indicator variables for their taxable income and an additional category for the parents having missing taxable income. Parents with zero taxable income belong

to the lowest decile.[1] The income distributions are year-specific, and we have formed the deciles based on observations in the sample, not in the whole population. We have not applied any equivalence scale to the taxable income.

Age of mother and father The ages of the mother and the father are controlled for using seven indicator variables. One category consists of parents with missing ages (meaning the person's parents were not observed during the year she was 15 years old). The other categories are: younger than 35 years, 35 to 39 years, 40 to 44 years and so on, the last category being older than 55 years.

Number of under 18-year-old children living in the same household We control for the number of under 18-year-old children using six indicator variables. One category consists of missing number of children (the number is not known), and another of households with one child under age 18. Households with 2 to 4 children form three separate categories, and households with 5 or more children under age 18 make up the last category.

Place of residence In the population-wide data, we control for the zip-code (postal code) of individual's place of residence at age 15. In the crime and health data, we do not observe the place of residence on this level of detail. However, the data contain information on the individual's province of residence and the degree of urbanization of the residence municipality. We combine this information into 124 indicator variables capturing the type of location the person was living in at age 15.

[1] Household income in terms of taxable income may be zero when the income consists of non-taxable social security benefits such as social assistance, child benefit and various housing benefits.

Table A1: Observations by age at arrival

	Native		Immigrant-native		Immigrant	
	N	%	N	%	N	%
0	1,026,552	98.77	9,838	75.42	997	7.43
1	2,335	0.22	313	2.40	270	2.01
2	1,909	0.18	240	1.84	431	3.21
3	1,678	0.16	233	1.79	578	4.31
4	1,456	0.14	210	1.61	621	4.63
5	1,263	0.12	236	1.81	768	5.72
6	1,093	0.11	230	1.76	785	5.85
7	967	0.09	274	2.10	847	6.31
8	581	0.06	241	1.85	868	6.47
9	477	0.05	288	2.21	900	6.70
10	339	0.03	319	2.45	1,006	7.49
11	240	0.02	300	2.30	999	7.44
12	180	0.02	327	2.51	1,027	7.65
13	160	0.02	313	2.40	1,140	8.49
14	89	0.01	308	2.36	1,040	7.75
15	22	0.00	275	2.11	1,147	8.54

Table A2: Educational Attainment in Comparison to the Children of Natives, 20 % population sample

	(1)	(2)	(3)	(4)
A: Holds a secondary degree at age 20				
Immigrant-native	-0.090	-0.039	-0.038	-0.013
	(0.011)	(0.011)	(0.011)	(0.011)
Immigrant	-0.313	-0.110	-0.068	-0.025
	(0.012)	(0.018)	(0.017)	(0.017)
Age at migration	.	-0.019	-0.012	-0.012
		(0.001)	(0.001)	(0.001)
B: Holds a secondary degree at age 23				
Immigrant-native	-0.069	-0.032	-0.030	-0.007
	(0.010)	(0.010)	(0.010)	(0.010)
Immigrant	-0.210	-0.064	-0.027	0.015
	(0.011)	(0.016)	(0.016)	(0.016)
Age at migration	.	-0.014	-0.007	-0.007
		(0.001)	(0.001)	(0.001)
C: Enrolled in higher education by age 23				
Immigrant-native	-0.018	0.028	0.017	0.040
	(0.014)	(0.014)	(0.014)	(0.014)
Immigrant	-0.178	0.001	0.056	0.088
	(0.015)	(0.022)	(0.021)	(0.021)
Age at migration	.	-0.017	-0.005	-0.005
		(0.002)	(0.001)	(0.001)
Controlling for:				
Parental characteristics	no	no	yes	yes
Residence location	no	no	no	yes

Note: This table reports estimates for equation (**??**), estimated from 20 % population sample. The outcomes are an indicator for holding a secondary degree at age 20 (panel A) or at age 23 (panel B) and having been enrolled in a university or a polytechnic by age 23 (panel C). Background characteristics are parent's average months of unemployment (14 categories), parents' combined income decile (ten categories), mother's and father's age (seven categories) and number of siblings (six categories). Residential location is measured as the interaction of province of residence and its level of urbanization. See Data Appendix for variable definitions. All background characteristics are measured at age 15.