Labor market outcomes of natives and immigrants: Evidence from the ECHP

Franco Peracchi and Domenico Depalo
Tor Vergata University

OECD, May 29, 2006

1 Introduction

- This paper analyzes the evidence on **labor market outcomes** of **natives** and **immigrants** provided by the **European Community Household Panel** (**ECHP**), a longitudinal household survey which covers a wide range of topics and gives comparable information across the 15 member states of the European Union before the 2004 enlargment (EU-15).

- Our goal is to provide a detailed **description of the differences between natives and immigrants** in **activity rates**, **employment rates**, **unemployment rates,** and **earnings**, controlling for a variety of personal characteristics, including the **country of origin** and the **length of stay in the host country**.

- In particular, we ask two sets of questions:

  - Are there significant **differences in labor market outcomes** of natives and immigrants? To what extent these differences may be accounted for by **differences in the observed characteristics** of the two groups?
  - Is there any evidence of **integration** of immigrants into the labor markets of Western European countries? How much of the residual differences in labor market outcomes of natives and immigrants **persist** after a sufficiently long residence of immigrants in the host country?

2 The data

2.1 Brief description of the ECHP

The ECHP is an **annual longitudinal survey**, based on a **standardized questionnaire**, whose **target population** consists of all **private households** throughout the national territory of each country. The total length of the ECHP is **8 years** (from 1994 to 2001). In the first wave, a sample of almost 130,000 people aged 16+ was interviewed in the (then) 12 member states of the EU. Austria, Finland and Sweden were **added later** (respectively in 1995, 1996 and 1997).

The following aspects of the ECHP are relevant for this paper:

- Depending on the sampling frame adopted, we may have **noncoverage** of small portions of the target population, such as **households recently arrived in a country** (Ireland, Italy) or **nonresidents unable to speak the national language** (Greece, Netherlands).
- **Very little** is known about **unit nonresponse in the first wave** of the ECHP.
- The ECHP aims at being both **cross-sectionally** and **longitudinally** representative, with changes in the population over time reflected by the continuous evolution of the sample – through births to sample households and the formation of new households from the split off of existing ones.
- In fact, because the ECHP does **not** employ refreshment samples, its **cross-sectional** representativeness tends to **fade away over time** due both to **non random attrition** and to the presence of demographic changes arising from the **arrival of new waves of immigrants**.
- The initial rules for following households and people have been **modified** in some countries, especially between the first and the second wave. In particular, most countries did **not** follow people who moved to another EU country.

2.2 Definition of immigrant status

- **Three variables** in the ECHP questionnaire provide information on immigrant status:
  - "previous foreign country of residence before coming to the present country",
  - "foreign country of birth",
  - "citizenship".
- For the first two variables, **three versions** are available:
  - The **most detailed** version provides a breakdown into **11 geographical areas**. This version is **not** available for Austria, Finland, Germany, Greece, Italy, Luxembourg, Netherlands and Sweden.
  - The **intermediate** version provides a breakdown into **7 geographical areas** (Africa; Asia; America; EU-15; other European countries; Australia, Oceania and other countries; any country not elsewhere classified). This version is **not** available for Finland, Germany, Greece, Luxembourg, Netherlands and Sweden.
  - The **coarsest** version provides a breakdown into **2 geographical areas** (EU-15; other foreign country). This version is **not** available for Germany, Greece, Luxembourg, Netherlands and Sweden.
- We work with the **intermediate** version (breakdown into **7 geographical areas**) because it provides some geographical detail on the country of origin without loosing too many countries.

2.3 Data checks

2.3.1 Comparing the various definitions

- To assess the degree of correspondence between last foreign country of residence and foreign country of birth we proceed as follows:
  - First we look for inconsistencies across waves.
  - After dropping the observations for which there are inconsistencies across waves, we tabulate the joint distribution of the two variables for the people ever in the sample. We find that 97.2% of the observations lie on the main diagonal.
- **Citizenship** may also be used to define immigrant status. Because this variable has a rather **coarse classification**, its use would lead to a considerable loss of information on the country of origin. Its main advantage is that data are **available for all countries and all waves**.
- **Except for the UK**, a definition based on citizenship generally leads to a **much lower** % of immigrants (**Table 1**).
- We decided to **drop the UK** because the data look "strange".
- Relying on **foreign country of birth** , as we do, has the following **advantages**:
  - It **conforms to the international standard definition** of immigrant.
  - It is **not affected by problems of return migration** of those who lived abroad and come back to their home country.
  - It is **not affected by naturalization**, a first step in the process towards citizenship.

Table 1: Percentage of immigrants by country according to different definitions of immigrant (weigthed data, first ECHP wave of each country).

Country	Citizenship	Foreign born	Foreign residence
Denmark	3.00	4.34	6.60
Netherlands	1.37		
Belgium	6.65	8.53	10.07
France	5.34	10.11	11.26
Ireland	1.53	4.69	10.46
Italy	.08	1.90	2.64
Greece	1.01		
Spain	.77	1.92	5.32
Portugal	1.36	3.58	8.85
Germany (SOEP)	5.74		
Luxembourg(PSELL)	32.41	34.00	34.97
Austria	5.85	10.10	10.10
Finland	1.30		
Sweden	4.87		
UK (BHPS)	2.20	.40	.45

2.3.2 Descriptive statistics

- Number of people in the sample by foreign country of birth (**Table 2**).
- Distribution of the sample by immigrant status (**Table 3**).
- Distribution of immigrants by country of birth (**Table 4**).

Table 2: Distribution of the sample by foreign country of birth.

Country	Natives	EU-15	Other	Africa	America	Asia	Austr.	Total
Denmark	7063	89	74	18	20	88	1	7353
Belgium	7459	385	87	137	21	21	0	8110
France	16107	480	114	558	17	80	0	17356
Ireland	11420	463	3	9	35	13	9	11952
Italy	21470	139	120	64	54	7	6	21860
Spain	21910	163	30	46	168	4	1	22322
Portugal	14913	109	4	215	70	5	0	15316
Austria	8539	165	493	14	13	38	0	9262
Total	108881	1993	925	1061	398	256	17	113531

Table 3: Distribution of the sample by immigrant status (percentage relative frequencies).

Country	Natives	Immigrants	Total
Denmark	96.06	3.94	100
Belgium	91.97	8.03	100
France	92.80	7.20	100
Ireland	95.55	4.45	100
Italy	98.22	1.78	100
Spain	98.15	1.85	100
Portugal	97.37	2.63	100
Austria	92.19	7.81	100

Table 4: Distribution of immigrants by country of birth (percentage relative frequencies).

Country	EU-15	Other Euro.	Africa	America	Asia	Austr.	Total
Denmark	30.69	25.52	6.21	6.90	30.34	.34	100
Belgium	59.14	13.36	21.04	3.23	3.23	.00	100
France	38.43	9.13	44.68	1.36	6.41	.00	100
Ireland	87.03	.56	1.69	6.58	2.44	1.69	100
Italy	35.64	30.77	16.41	13.85	1.79	1.54	100
Spain	39.56	7.28	11.17	40.78	.97	.24	100
Portugal	27.05	.99	53.35	17.37	1.24	.00	100
Austria	22.82	68.19	1.94	1.80	5.26	.00	100

2.3.3 Comparison with other data sources

To what extent is the picture from the ECHP **consistent** with the information obtained from other sources?

- The picture from the first ECHP wave (**Table 5**) is roughly consistent with the information provided by the OECD for the year 1994 (**Table 6**).
- The discrepancies between the ECHP and the OECD data **increase over time**. The main reasons are:
  - the **lack of refreshment samples** in the ECHP;
  - the **differential rate of panel attrition** between natives and immigrants (**Table 7).**

Table 5: Fraction of immigrants by country and wave (percentage relative frequencies).

	Year							
Country	1994	1995	1996	1997	1998	1999	2000	2001
Denmark	4.00	3.93	3.68	3.41	3.30	2.79	2.92	2.93
Belgium	9.25	8.82	8.27	8.14	7.83	7.63	7.34	6.72
France	9.82	8.94	8.21	7.65	6.81	6.42	5.79	5.52
Ireland	4.65	4.69	4.38	4.25	3.98	3.72	3.62	3.46
Italy	1.85	1.87	1.83	1.75	1.64	1.55	1.52	1.43
Spain	1.93	1.83	1.71	1.55	1.57	1.55	1.53	1.52
Portugal	2.81	2.73	2.53	2.54	2.33	2.38	2.34	2.11
Austria		8.31	7.65	6.74	6.72	6.34	6.45	6.19

Table 6: Fraction of foreign people on total population by country and year (percentage relative frequencies). Source: OECD.

	Year							
Country	1994	1995	1996	1997	1998	1999	2000	2001
Denmark	3.80	4.20	4.70	4.70	4.80	4.88	4.80	5.00
Belgium	9.10	9.00	8.97	8.86	8.70	8.80	8.40	8.21
France						5.60		
Ireland	2.69	2.69	3.20	3.13	3.00	3.20	3.30	3.90
Italy	1.60	1.70	2.00	2.10	2.10	2.17	2.43	2.36
Spain	1.18	1.27	1.37	1.60	1.83	2.00	2.20	2.74
Portugal	1.58	1.70	1.70	1.76	1.78	1.90	2.08	2.17
Austria	8.90	9.00	9.04	9.08	9.13	9.20	9.34	9.40

Table 7: One-year attrition rates by year and immigrant status.

Country	Immigrant status	Year 1995	1996	1997	1998	1999	2000	2001
Denmark	Natives	.04	.07	.08	.09	.07	.06	.04
	Immigrants	.09	.11	.10	.12	.13	.09	.04
Belgium	Natives	.03	.04	.05	.06	.07	.07	.10
	Immigrants	.02	.04	.04	.08	.07	.10	.16
France	Natives	.03	.03	.05	.06	.06	.07	.06
	Immigrants	.03	.04	.06	.07	.08	.10	.06
Ireland	Natives	.05	.06	.06	.07	.11	.17	.12
	Immigrants	.05	.09	.07	.08	.15	.18	.14
Italy	Natives	.02	.02	.05	.06	.05	.07	.09
	Immigrants	.02	.03	.06	.07	.08	.07	.14
Spain	Natives	.04	.05	.06	.07	.07	.08	.07
	Immigrants	.06	.07	.11	.09	.08	.09	.11
Portugal	Natives	.02	.03	.03	.04	.04	.05	.04
	Immigrants	.03	.06	.04	.10	.05	.05	.09
Austria	Natives		.03	.04	.06	.06	.08	.06
	Immigrants		.06	.09	.08	.09	.06	.08

2.4 Sample selection

- Since we are mainly interested in labor market outcomes, we restrict attention to the **working age population** (people aged 20–64).
- To keep a balance between consistency and detail in the definition of immigrant, we work with the following 8 countries: **Austria, Belgium, Denmark, France, Ireland, Italy, Portugal, Spain**.
- The resulting sample consists of **89,799 individuals** observed from a minimum of one year to a maximum of eight years.

2.5 Construction of labor market outcomes

2.5.1 Current employment status

- To simplify the analysis, we only distinguish between employed and non-employed people. Based on the ILO main activity status at the time of the interview, we classify people **as employed** if they are "currently working" or "normally working", and as **non-employed** otherwise.

- For **employed** people, we further distinguish between **full-time (FT)** and **part-time (PT)** workers, and between **employees** and **self-employed**:

  - The distinction between FT and PT workers is based on **hours worked per week** in the main and additional jobs (our threshold for FT work is **35 hours/week**).
  - The distinction between employees and self-employed is based on the **type of employment**. People are classified as self-employed if they are in self-employment or are unpaid workers in a family enterprise, and as employees otherwise.

2.5.2 Employment status last year

We classify a person as **full-year** (FY) worker, **part-year** (PY) worker or **non-worker** using the information on main activity status in each single month of last calendar year:

- a person is classified as a **FY worker** if her main activity status was employment in all 12 months last year;

- a person is classified as a **PY worker** if she is not a FY worker and her main activity status was employment in at least one month;

- a person who is neither a FY nor a PY worker, is classified as a **non-worker**.

2.5.3 Earnings

- The ECHP contains information on **two earnings concepts**:
  - monthly earnings on the current main job ("**current monthly earnings**" for short),
  - **annual earnings in the last calendar year**.

  All amounts are in **national currencies** and **current prices** and, except for France, are **net** of social security contributions and income taxes.
- All monetary amounts have been **converted to Euros and adjusted using PPPs** for the year considered.
- **Imputation** is an issue, because little is known about ECHP imputation at the personal income level. Nicoletti and Peracchi (2005) show that:
  - **wages and salaries of full item-nonrespondents** appear to be **underestimated** in the ECHP. However, the **number of cases involved is small**, and so statistics computed for the full sample and the subset of respondents do **not** differ much;
  - for self-employment income, instead, **full item-nonresponse is very frequent**, but we find **no evidence of non-response bias**.

2.6 Construction of the covariates

2.6.1 Year of birth

Year of birth is **top-coded** for all countries and all waves with people born earlier than 1909 recorded as born in 1909. As a result, a person's age is top-coded at 85 years in wave 1, 86 years in wave 2, etc.

2.6.2 Marital status

The reference period for marital status is the time of the interview. The ECHP distinguishes between 5 categories but, to simplify the analysis, we only distinguish between people **with a spouse** ("married") and people **without a spouse** (other).

2.6.3 Education

- The ECHP provides **three** alternative measures:
  - the highest level of general or higher education completed,
  - the age when the highest level of general or higher education was completed,
  - the age when full-time education was stopped.
- The first variable is relevant if educational attainments are conventionally defined following the **ISCED**. The other two variables would be relevant if educational attainments were measured in terms of **years of education**.
- Combining this information with that on the "year of arrival in the country of present residence", one could in principle **distinguish between the type of education received in the home and the host country**.
- Unfortunately, the information on the last two variables is **incomplete**. Age when full time education was stopped is not given until 1998. For both variables, the panel documentation also reports problems in wave 5 for Portugal.

2.6.4 Labor market experience

- For those who ever worked, we construct a measure of labor market experience by taking the **difference between the age in current wave and the age when the person started her first job or business**.
- For immigrants, the years of labor market experience **in the host country** cannot be determined precisely. A useful piece of information would be the **year of start of current job**, but:
  - this question is only asked to those who report themselves as "normally working",
  - the codes are rather imprecise. Of the 6 categories available, only one ("1981 to 2002") looks fairly reliable. For these reasons, we do not distinguish between labor market experience abroad and in the country of present residence.

2.6.5 Length of stay

Our measure of the length of stay in the present country is the difference between the **year of the current wave** and the **year of arrival in the country of present residence**.

2.7 Data availability

2.7.1 Labor market outcomes

- **Working status** contains a **small** fraction of missing values (most of them for Belgium and the subsample of natives).
- The indicator for **FY/PY employment** also contains a **small** fraction of missing values (about 2/3 of them are for France, mostly for the sample of natives and concentrated in the 2nd wave).
- The **self-employment indicator** contains a **small** fraction of missing values but a **large** fraction of "not applicable" (because the question is only asked if the main activity status is either "currently working" or "normally working").
- There are no missing values for **current monthly earnings**, but this is because missing data have been **imputed**. Unfortunately, there is **no flag** for imputed values.

2.7.2 Covariates

- **Year of birth** and **gender** contain **no missing values** and appear to be **highly reliable** for all countries. We checked for inconsistencies across waves, but found none.
- **Marital status** has **very few** missing values (somewhat surprisingly, all missing values are for natives).
- **Education level** is available for more than 99% of the observations. Missing values are more frequent for natives than for immigrants.
- **Labor market experience** contains no "missing values" nor "not applicable values", since we constructed the variable only for those who ever worked.
- **Length of stay in the host country** has **very few** missing values (except in France for the sample of natives) or inconsistencies.

3 Descriptive statistics

For simplicity, we only present the results obtained using the **sample weights**.

3.1 Labor market outcomes

All statistics are computed for the subsample of **people aged 20–64**, after dropping cases with missing values for the relevant variables.

3.1.1 Employment status

- **Current employment status**: Employment rates are always higher for men than for women. With some exceptions, they are also higher for natives than for immigrants. Male/female differentials tend to be larger for immigrants than for natives.
- **FT/PT status**: For both natives and immigrants, PT appears to be more important for women than for men. PT tends to be more important for immigrants than for natives, especially for men.
- **Employee/self-employment status**: For both natives and immigrants, self-employment appears to be more important for men than for women. For men, it tends to be more important for immigrants than for natives, whereas for women the opposite tends to be true.
- **Employment status last year**: The fraction of non-workers is always much higher for women than for men, irrespective of immigrant status. Except for Italy, it is higher for immigrants than for natives.
- **FY/PY status**: For both natives and immigrants, the fraction of FY workers is much higher for men than for women. In general, this fraction is greater for natives than for immigrants. Women are more likely to be PY workers than men, no matter the immigrant status. There is some evidence that PY work is more frequent for immigrants than for natives.

3.1.2 Earnings

We consider 2 measures of earnings: **current monthly earnings** and **average monthly earnings in the last calendar year**.

All statistics considered have been computed for the subsample of **workers aged 20–64**, after **symmetrically trimming 1% of the sample on each side** for each country.

We distinguish between various categories of workers. For current monthly earnings we distinguish between **all workers**, **FT and PT workers**, and **employees and self-employed**. For average monthly earnings last year, we only distinguish between **FY and PY workers**.

- Current monthly earnings are always **higher for men than for women**, irrespective of immigrant status. The **gender gap** varies by country, but is around 20% for FT or FY workers, and somewhat larger for PT or self-employed workers. Natives tend to have higher mean earnings than immigrants, but differences are not large.
- **Within-country variability** tends to be higher for men than for women. It also tends to be higher for immigrants than for natives.
- **Cross-country variability** is substantial. In Denmark, Belgium, France and Ireland, the mean of current monthly earnings of a male worker is above 1300 Euros, in Austria it is 50–100 Euros lower, while in Spain and Italy it is about 150–200 Euros lower. Earnings are always lowest in Portugal.
- Confining attention to FT or FY workers, the **distribution** of monthly earnings of immigrants is shifted a little to the right and possibly **more spread** out than for natives, but **differences tend to be small**.

3.1.3 Length of stay

- The length of stay in the country of present residence is a **key** covariate in our analysis.
- Table 8: Sample statistics based on the **last wave available** for each individual (year 2001 for the majority of cases). Each statistic has been computed using the subsample of people aged 20–64, excluding a few cases with inconsistency between the year of arrival in the present country and the year of birth.
- **Main findings**:
  - The immigrants in our sample have spent a **considerable amount of time** in the host country, the **mean** length of stay being **at least 17 years**, with large differences across countries.
  - In Belgium, France and Italy, **men** have a higher length of stay than women. In Denmark and Ireland, **women** have a higher length of stay than men. In Austria, Portugal and Spain there is virtually **no gender difference**.

Table 8: Statistics for length of stay of immigrants by country and gender.

Country	Mean	SD	p25	p50	p75
Male					
Denmark	17.3	11.6	8.0	16.0	25.0
Belgium	27.6	13.6	17.0	29.0	38.0
France	28.0	12.5	20.0	29.0	36.0
Ireland	22.6	11.9	16.0	22.5	30.0
Italy	25.6	13.7	16.0	25.0	35.0
Spain	18.3	11.5	8.0	19.0	25.0
Portugal	20.4	8.1	16.0	21.0	25.0
Austria	18.6	14.4	8.0	14.0	26.0
Female					
Denmark	19.0	13.4	8.0	17.0	27.0
Belgium	25.2	12.9	15.0	26.0	36.0
France	25.9	12.6	16.0	25.0	35.0
Ireland	23.6	12.6	17.0	22.0	31.0
Italy	22.2	12.4	12.0	23.0	30.0
Spain	18.2	11.6	8.0	19.0	27.0
Portugal	20.3	8.4	15.0	21.0	26.0
Austria	18.8	15.8	7.0	12.0	27.0
Total					
Denmark	18.2	12.6	8.0	16.0	26.5
Belgium	26.4	13.3	16.0	28.0	36.0
France	26.9	12.6	18.0	27.5	36.0
Ireland	23.2	12.3	16.0	22.0	31.0
Italy	23.6	13.1	13.0	24.0	32.0
Spain	18.2	11.5	8.0	19.0	26.0
Portugal	20.3	8.3	15.0	21.0	26.0
Austria	18.7	15.2	7.0	13.0	27.0

4 Regression analysis

- This section presents the results of fitting simple regression models to the individual data in order to **summarize** the way in which labor market outcomes vary between natives and immigrants depending on the **country of residence** and **observable personal characteristics**.

- The labor market outcomes that we consider are:

  - **Labor force state probabilities**, with labor force states defined on the basis of "ILO main activity status at the time of interview": unemployed, employed (normally working or currently working), and active (employed or unemployed).

  - **Mean log monthly earnings** (current monthly earnings and average earnings last year).

- For each labor market outcome, we consider both **models for the pooled data** and **separate models for immigrants and natives** (also distinguishing between immigrants from **EU-15** and from **non EU-15** countries). The models for immigrants contain a **richer** set of covariates than the models for natives, as we control for **additional** variables (typically a measure of the **length of stay in the host country** and indicators for the **area of origin**).

- All models are fitted **separately for men and women**.

4.1 Labor force status

- The basic model for **labor force state probabilities** is the linear logit model

$$\eta(X) = \ln \frac{\pi(X)}{1 - \pi(X)} = \alpha + \beta X,$$

  where $\pi(X)$ is the conditional probability of being in a certain labor force state given a vector X of covariates which contains age and its square, schooling dummies, a dummy for not having a spouse, and dummies for the present country and the calendar year.

- The model for the **pooled data** also includes an **immigrant dummy**. The model for the **immigrant sample** includes instead a set of dummies for the **area of origin** and the **length of stay** in the host country.

- The intercept α corresponds to the **log-odds for the reference person** (an individual aged 35, with basic education only, with a spouse, observed in year 2001 and residing in Italy). For the model fitted to the **pooled data**, the reference person is a **native**. For the model fitted to the **immigrant data**, is instead an **immigrant from the EU-15** who has been living in Italy for **less than 5 years**.

- Each model has been estimated by **ML**, after dropping cases with missing covariates. Significance levels are based on **robust standard errors**.

- The **estimation sample** consists of **87,901 individuals aged 20–64 (445,138 observations)**. Of these, 84,081 individuals (427,074 observations) are **natives** and 3,820 individuals (18,064 observations) are **immigrants** – 1,625 (7,951 obs.) from **EU-15** countries and 2,195 (10,113 obs.) from **non EU-15** countries – with a ratio of immigrants to natives of 4.5% in terms of individuals and 4,2% in terms of observations.

4.1.1 Activity rate

- Activity rates are higher for men than for women. They are also higher for natives than for immigrants, but the differences are **much smaller** than those by gender.
- For both men and women, the estimated relationship between activity rate and **age** has an **inverse U-shape**. Because the coefficient on the linear age term is **larger** for immigrant men than for native men, the profile of the age-activity relationship is initially steeper but then falls less rapidly for immigrant men than for native men.
- Activity rates are **highest for people with tertiary education**. The coefficient on the dummy for tertiary education is **larger for women and for immigrant men**.
- **Not having a spouse** is associated with **lower** activity rates for men but **higher** activity rates for women. The size of the marital status effect is larger for native men than for native women, and for immigrant women than for immigrant men.
- The dummies for the **country of residence** are strongly significant for natives but **not** for immigrants, whereas the dummies for the **country of origin** are **never significant**.
- Downward time trend for **native men**, upward time trend for **native women**, but no clear time trend for **immigrants**.
- For immigrants, especially **non EU-15 women**, the **length of stay** in the host country is important. Initially, they tend to have **lower** activity rates than natives. The gap progressively **dimishes** as the length of stay in the country **increases**.

4.1.2 Employment rate

- The results obtained are **very similar** to those for the activity rate. A noticeable difference is the larger size of the coefficient on the linear age trend for women.
- Another difference is the fact that, employment rates **increase** with educational attainemnts.
- The effect of **calendar time** and **length of stay** in the host country are **stronger** than in the case of the activity rate. The dummies for length of stay in the host country are now always significant for non EU-15 men.
- The effect of the other variables in the model (marital status, country of residence, and country of origin) are much the **same** as for the activity rate.

4.1.3 Unemployment rate

- In general, the sign of the coefficients is the **opposite** than for the activity rate and the employment rate.
- Other things being equal, unemployment rates **decline** with educational attainemnts.
- Not having a spouse is associated with **higher** unemployment rates for **men** but **lower** unemployment rates for **women**.
- The dummies for the **country of residence** are strongly significant for natives but **not** for immigrants.
- Downward time trend, somewhat stronger for **natives** tahn for **immigrants**.
- The **longer** an immigrant has been residing in the host country, the **lower** is the probability of unemployment. Length of stay is not significant for non EU-15 women.

4.2 Earnings

- The basic model for the **mean of log monthly earnings** is the linear model

$$\mu(X) = \alpha + \beta X,$$

where $\mu(X)$ is the conditional mean given a vector X of covariates which contains years of labor market experience and its square, and dummies for schooling attainments, for not having a spouse, the host country and the calendar year.

- The model for the **pooled data** also includes an **immigrant dummy**. The model for the **immigrant sample** includes instead a set of dummies for the **area of origin** and the **length of stay** in the host country.

- The intercept α corresponds to the mean of log earnings for the reference person (an individual with 20 years of labor market experience, basic education only, with a spouse, observed in year 2001, residing in Italy). For the model fitted to the **pooled data**, the reference person is a **native**. For the model fitted to the **immigrant data**, is instead an **immigrant from the EU-15** who has been living in Italy for **less than 5 years**.

- We consider **four different definitions of earnings**, namely:
  - **current monthly earnings** of an **employed person** and a **FT employee**,
  - **average monthly earnings last year** of an **employed person** and a **FY employee**.

- Each model is estimated by **OLS**, after dropping observations with missing covariates and with monthly earnings **below 100 Euros**(less than 2% of the sample in Belgium, Denmark and Portugal, less than 1% in the other countries). Significance levels are based on **robust standard errors**.

- The **estimation sample** consists of 214,659 observation for current monthly earnings (79.7% of the sample are FT employees) and 226,749 observations for average monthly earnings last year (79.6% of the sample are FY employees).

4.3 Current monthly earnings

- The gap between male and female earnings is about 35–50%, and is **remarkably similar** for natives and EU-15 immigrants.
- For male workers, the "college premium" is about 50% and the "high-school premium" is about 20%. Educational premia are **larger for female workers**. Although different by gender, educational premia are **remarkably similar** for natives and immigrants.
- Estimated earnings-experience profile are **concave** and **surprisingly similar** both for men and women, and for natives and immigrants.
- The **premium for having a spouse** is positive for men but negative for women, and the positive premium for men is larger in size than the negative premium for women. For non EU-15 immigrant men there is essentially no premium.
- For natives, cross-country differences are **large**, with France and Portugal as the two extremes. For immigrants, cross-country differences are **smaller**.
- Earnings growth is **slower for immigrants** than for natives, mainly because of the slower growth of earnings for immigrant men, but these differential trends do **not** change much the earnings differentials by immigrant status.
- The relative position of non EU-15 male immigrants depends strongly on the **length of stay** in the host country. For men, a longer stay is associated with a narrower earnings gap and, **after 20+ years**, there is essentially **no earnings gap**. The length of stay is not statistically significant for non EU-15 female immigrants and for EU-15 immigrants.
- The area of origin does **not** appear to be statistically significant.

4.4 Average monthly earnings last year

The sign and magnitude of most coefficients are **very similar** to those obtained for current monthly earnings.

5 Summary

The ECHP provides useful information on the differences in labor market outcomes of natives and immigrants. However:

- detailed information is available for **only 8 EU-15 countries** (Austria, Belgium, Denmark, France, Ireland, Italy, Portugal, and Spain).
- Since it has **no refreshment sample**, the ECHP allows us to follow the process of economic integration of the cohorts of immigrants that reached Western Europe before the mid-1990s but can tell us **very little** about later cohorts of immigrants.

Labor market outcomes differ significantly between natives and immigrants but **differences are small compared to those between men and women**. In particular, **other thinsg being equal**:

- Natives tend to have higher activity rates, higher employment rates, lower unemployment rates and higher monthly earnings than **newly arrived** immigrants.
- **Given gender and immigrant status**, important predictors of labor market outcomes are **age** (or **labor market experience**), **educational attainments**, and **marital status**. The effects of these variables are **remarkably similar** between natives and immigrants (maybe not surprising since, with a few exceptions, immigrants from other EU-15 countries represent more than 30% of our sample).
- **Cross-country differences** are sizeable for natives, but much smaller for immigrants. For immigrants, the effects of the **area of origin** tend to be small.
- For immigrants from **non EU-15** countries, another important predictor is the **length of stay in the host country**. A **longer** stay is associated with **smaller** differences relative to natives. For immigrants who have being in the host country for 25+ years, there is essentially **no difference** relative to a native with similar characteristics.

6 Warnings

- Our positive conclusions about integration of immigrants may **not** generalize to the cohorts of immigrants that reached Western Europe **after the mid-1990s**.
- They may be difficult to generalize to the non-negligible fraction of immigrants who **dropped out of the ECHP sample** (to move to another country or for other reasons). Their labor market outcomes may **not** have been as favorable as those of the "survivors" into the ECHP.
- They may also be difficult to generalize to another group of immigrants, about which we know nothing, namely those **not included into the 1st ECHP wave** because of problems with the sampling frame, non contact, language problems, or refusal to participate.