Regional Labour Markets in Belgium: The Role of Education*

David Sonnewald†

August 23, 2025

Abstract

This paper delves into labour market differences across the three regions in Belgium. I use the information available in the Labour Force Survey to examine how personal and family characteristics affect labour force participation and unemployment risk, as well as how these effects differ across the three regions, by means of a Heckman selection model. I find a paramount role of education in increasing and decreasing labour force participation and unemployment risk, respectively, although with diminishing marginal returns. Cross-regional comparisons show that these effects are even more pronounced in Wallonia, whereas Flanders does not show significant differences from Brussels.

Keywords: Heckman Selection, Determinants of Unemployment, Regional Unemployment.

JEL Classification: J21, J64, J82, C52.

*I am very thankful to my supervisors Koen Declercq (UCLouvain) and Tom Truyts (UCLouvain) for their constant guidance and support. I also acknowledge the precious comments and suggestions received from Elisa Navarra (IMF), my PhD Supervisory Committee and Jury, participants at the Spring Doctoral Workshop in Economics 2022 (UCLouvain), and the 17th edition of the Belgian Day for Labour Economists (Federal Planning Bureau, Brussels). This work was made possible through the generous support of the Innoviris Prospective Research for Brussels programme (Project 2019-PRB-108), which financed my research.

†Center for Applied Public Economics (CAPE) and Center for Research in Economics (CEREC), UCLouvain. Email: david.sonnewald@uclouvain.be. Personal website: https://davidsonnewald.com.

1 Introduction

There is increasing interest in the significant role that regions play in shaping local labour markets (OECD (2022)). Regional characteristics, including economic structure, educational attainment, industrial composition, and local policies have a deep impact on employment patterns, wage levels, and job opportunities.

Belgium presents a valuable case study for examining cross-regional labour market differences due to its complex federal structure and important socio-economic diversity.[1] Each of its regions differs profoundly from the others, and disparities are observable both in terms of standard labour market indicators and characteristics of the working age population.

The objective of this paper is to drill down into these acknowledged existing interregional differences in the labour market in Belgium with a focus on the role of individual characteristics, rather than macro variables,[2] and their effects on labour force participation and unemployment risk. In particular, I centre around the heterogeneous impact of returns to education.

I exploit the information available in the Labour Force Survey for Belgium for the years 2004-2019 to identify the main personal and household characteristics correlated with labour force participation and unemployment. That is, I examine how they affect the probability for an individual to be active and to be unemployed and how these probabilities differ in the three regions. To do so, I employ a Heckman selection model (Heckman (1976)). The probability of being unemployed is estimated jointly with the probability of being active, which determines the main drivers of labour force participation and identifies the sample of the active population, using the maximum likelihood method. Hence, I identify the major determinants of unemployment, minimising sample selection bias through the Heckman correction. To better highlight and study cross-regional differences, I perform such exercise both for Belgium and for its three regions.

At country level, I find that the great majority of the estimated coefficients of the individual characteristics (gender, age, marital status, country of birth, place of residence, household composition, and education) are significantly correlated with both labour force participation and unemployment risk. I measure regional effects as deviations of the coefficients for Wallonia and Flanders from the baseline region, Brussels. The value and the statistical significance

[1]Belgium is a federal state made up of three regions and three linguistic communities. The three regions are the Flemish Region (or Flanders), the Brussels-Capital Region (hereafter, simply Brussels) and the Walloon Region (or Wallonia), which cover the north, the centre and the south of the country, respectively. The linguistic communities are the Dutch speaking, the French speaking and the German speaking Community. Contrary to the regions, the communities, being based on the concept of language, are not bounded to geographic borders. In fact, the official language changes across regions being Dutch in Flanders, French (and minimally German) in Wallonia and both French and Dutch in Brussels.

[2]This involves mainly variables related to cyclical conditions (*e.g.*, GDP) and institutional framework.

of the coefficients varies across the regions. For Belgium and Brussels (reference region), I find a paramount role of education in increasing and decreasing labour force participation and unemployment risk, respectively, although with diminishing marginal returns[3]. Cross-regional comparisons show that these effects are even more pronounced in Wallonia, whereas Flanders does not show significant differences from Brussels.

This article builds on three main lines of research. First, it adds up to the literature where micro data are used to discuss factors associated with unemployment. In fact, compared to those using macro data, there are still relatively few econometric investigations employing this kind of data to study unemployment. In their estimation of the probability of unemployment and labour force participation in Italy, Kostoris and Lupi (2002) find that youth unemployment strongly depends on family income and wealth. Brooks (2005) shows that, in Canada, chronic unemployment is associated with low levels of education, residence in certain regions and membership to visible minority groups. Schmillen and Möller (2010) find that lifetime unemployment, in Germany, is highly unevenly distributed according to individual characteristics. By using the Labour Force Survey data for Belgium, the present study provides additional evidence that personal characteristics (such as age, country of birth, place of residence, household composition, and education) matter in explaining unemployment differentials.

Secondly, this paper speaks to the literature that explores differences in the unemployment rates and in the composition of the unemployed population at regional (sub-national) level. Marelli et al. (2012) show that regional unemployment differentials are large across regions within European countries. They also find that low unemployment regions tend to cluster close to each other. The OECD (2020) report also proves that these gaps are persistent, as only one-third of the OECD countries have experienced an increase in labour productivity in all their regions since 2008. An interesting case study within this strand of literature, similar to Belgium in many aspects (*e.g.*, federal state structure and multilingualism), is Switzerland. Feld and Savioz (2000) find that the main factors explaining the cantonal dispersion of unemployment in Switzerland are fiscal policies, the number of foreigners with different work permits, and human capital investments. More precisely, they find that high tax burdens and the number of foreign workers are positively correlated with cantonal unemployment, while a negative correlation is found with human capital investments. Within the same context, Brügger et al. (2009) study differences in unemployment duration between the Swiss cantons using language as a proxy for culture, as the main linguistic groups (German, French, and Italian speakers) are historically and socially connected to the respective cultural traditions of neighbouring countries. They find

[3]Diminishing marginal returns on education mean that once some level of education is achieved, additional education reduces the unemployment risk less than proportionally.

that differences in culture (proxied by language) between the Swiss cantons explain differences in unemployment duration on the order of 20%.

Thirdly, this study contributes to the human capital literature pioneered by Mincer (1974). While the existing research predominantly focuses on returns on education in terms of wages and incomes, the present analysis shifts the emphasis towards examining education as a mechanism to enhance labour force participation and reduce unemployment. This paper broadens the scope of human capital theory, highlighting the multifaceted role of education in improving overall labour market outcomes beyond monetary ones.

As for Belgium, several studies are devoted to the differences in terms of demography and main economic figures (*e.g.*, GDP per capita) across the three regions.[4] However, the literature on labour market gaps is not as extensive, although pioneer studies date back to the early 2000s. The IMF (2002) identifies wage compression, low mobility and job matching as the main factors behind persistent labour market regional differences in Belgium. Along the same lines, Estevào (2003) finds that the worsening of the matching of people to jobs in the high-unemployment areas of the country is the main cause of the widening regional labour market discrepancies. Meunier and Mignolet (2005) examine regional employment disparities in Belgium and their evolution by developing a shift-share analysis, which breaks down changes in employment in a region over time into components that identify whether growth is due to national trends, industry-specific trends, or the region's own competitive advantages. They find that differences in employment dynamics across the three Belgian regions are mainly driven by the concentration of certain industries, natural endowments, and non-traded infrastructure within each regional economy. While neither Flanders nor Wallonia have inherited a particularly favourable industrial base (Wallonia especially struggles with the decline of its traditional industries, *e.g.*, steel), the Flemish Region benefits from access to the sea, well-equipped harbours, and a dense highway network. Brussels stands apart with its heavy reliance on public administration and services. By creating a mismatch index, Zimmer (2012) evaluates the size of the qualification mismatch in Belgium and its regions, detecting the highest mismatch in the Brussels-Capital Region. In fact, in the capital, employment is concentrated in the services sector and requires primarily workers with a high educational level or specific skills due to the presence of public administrations and international institutions. However, for the most part, the job-seekers residing in the Capital Region are low-skilled. In a more recent study, Buysse and Saks (2020) examine the drivers of activity (as labour status) and quantify the labour reserve both at federal and regional level.

[4] *E.g.*, Van Oyen et al. (1996) on differences in health expectancy; Plasman et al. (2008) on differences in wages; Vandenberghe (2010) on differences in education; Hoorelbeke (2013) on differences in GDP and disposable income.

They show that, while there is scope for raising the activity rate in Belgium and in each of the three regions, only in the Flemish Region the activity rate bears comparison with the three neighbouring countries (Germany, Luxembourg and the Netherlands). With this article, I aim to shed new light by expanding the analysis to include the drivers of unemployment.

The remainder of the paper is organised as follows. Section 2 explores the employed data and presents a number of summary statistics about the labour market in Belgium and its three regions. In Section 3, I present the identification strategy of my empirical analysis. The main results provided by the Heckman selection model are summarised in Section 4. Section 5 concludes.

2 Data

2.1 The Labour Force Survey

The European Union Labour Force Survey (EU LFS) is a large household sample survey on labour participation of people aged 15 and over and on people outside the labour force.[5] It focuses on individual socio-demographic characteristics and conditions of employment, unemployment, and inactivity.

In this paper, I use the Eurostat version of the LFS data for Belgium for the years 2004-2019. The reason for using such time frame is twofold. On the one hand, in terms of institutional context, Belgium nearly completed its federalisation process with the Fourth and Fifth State Reform,[6] in 1993 and 2001, respectively. Therefore, earlier to this time, the status of "Region" did not have the same meaning as it has nowadays.[7] On the other hand, from a macroeconomic perspective, I include the years of the Global Financial Crisis in order to examine earlier trends and assess whether its aftermath affected the three regions differently. Yet, I exclude the years of the COVID-19 pandemic, as the crisis introduced short-term distortions in labour markets that are not representative of underlying structural or cyclical dynamics.

Overall, the data employed consists of a merged repeated cross sectional data set counting 1,400,564 observations. Of this original sample, I only consider the working age population (people aged between 15 and 64), for a total of 960,769 observations.

The variable "Region" refers to the region of residence, which does not imply any assumptions about the localisation of labour markets. Hence, individuals can work in regions different from

[5]Persons carrying out obligatory military or community service are not included in the target group of the survey, as is also the case for persons in institutions/collective households. For a detailed description of the data set, see Eurostat (2020).

[6]A minor Sixth State Reform also followed in 2011.

[7]2004 is the first year, after the State Reforms, since which all variables are collected identically and follow the same the specifications and standards.

the one in which they live. On the one hand, this is in line with the actual evidence showing that a significant number of people in Belgium commute to work in regions different from the one where they reside, particularly to the Brussels Metropolitan Area. On the other hand, the validity of my results is not undermined, as my outcomes of interest (labour force participation and unemployment risk) do not rely on the assumption that individuals work in their region of residence and, therefore, are not biased by this possibility.

2.2 Summary Statistics

Table 1 presents descriptive statistics for the working age population, separately for each region and aggregated at national level. While the population in the Walloon and in the Flemish Regions present similar characteristics, that in Brussels differ on a number of aspects. Thus, tests of equal means are performed to examine if there are systematic differences between figures in the Brussels-Capital Region and in the other two regions. Differences in the variables' means for Wallonia with respect to Brussels and Flanders with respect to Brussels are in fact almost all statistically significant.

The working age population in Brussels is on average about two years younger than in Wallonia and Flanders. In these two regions, between 85% and almost 90% of the sample's individuals were born in Belgium; in Brussels, this is the case for only half of the population. The remaining half is mainly made up of people originally from other European countries (25% EU, and 3% non-EU) and from Africa (14% North Africa, and 5% Sub-Saharan Africa).

At the country level, the proportion of singles (including households with more than one adult not in a couple) and of couples is broadly similar. However, in the Brussels-Capital Region, almost a third of the individuals live alone, against 21.14% in the Walloon Region and only 14.86% in the Flemish Region. In contrast, the number of children per capita in Brussels is higher than in the other two regions.[8]

Education-wise, in Wallonia and Flanders about 60% of individuals attained up to secondary education, with low shares of people with either very low or very high levels of education. On the contrary, in Brussels the share of the population with at most secondary education is only 47%, with almost a fourth of the individuals holding a Master's or a higher diploma, but also 15% with at best primary education. Together with the country of birth, this is the most pronounced difference between the capital and the rest of the country, and it reflects the increasing job and social polarisation that Brussels, as most of large metropolises in the world, has been experiencing in the last three decades: new jobs are created for either highest or least

[8]This could be possibly related with the higher share of foreigners from African countries, who have on average more children (Kulu and González-Ferrer (2013)).

qualified workers, with clear backlashes on labour force composition (Biot et al. (2019)).

Finally, in terms of breakdown according to the labour status, in the Brussels-Capital Region, the shares of the employed (employment rate) and the unemployed are smaller and larger, respectively, with respect to both the Walloon and notably the Flemish Region. However, Brussels shows a share of people still in education (14%) that is larger than the one in both the other two regions.

Table 1: Descriptive statistics

	Brussels	**Wallonia**	**Flanders**	**Belgium**
Individual characteristics				
Male	49.71%	49.97%	50.51%***	50.25%
Age (average)	38.08	39.75***	40.19***	39.83
Married	42.52%	43.84%***	50.18%***	47.35%
Country of birth (BE)	51.42%	85.37%***	88.91%***	83.85%
Living in a city[a]	100%	34.26%***	34.27%***	41.14%
Household composition				
Single adult	30.00%	21.14%***	14.86%***	18.46%
Couple	43.96%	47.68%***	55.83%***	51.96%
More than one adult[b]	26.04%	31.18%***	29.31%***	29.57%
Number of children in the HH (average)	1.10	1.02***	0.96***	0.99
Education				
Up to primary education	14.83%	13.02%***	9.79%***	11.36%
Up to secondary	47.16%	57.12%***	56.82%***	55.90%
Bachelor's or equivalent	13.23%	17.91%***	19.91%***	18.57%
Master's (or higher)	24.78%	11.95%***	13.48%***	14.17%
Labour Status[c]				
Employed	54.88%	56.96%***	66.48%***	62.21%
Unemployed	10.75%	6.73%***	3.20%***	5.12%
Education	14.03%	13.36%***	11.98%***	12.64%
Disability	5.00%	5.47%***	5.04%	5.20%
Inactive	15.34%	17.48%***	13.3%***	14.83%
Observations				
Non-weighted absolute number	123,453	364,354	472,962	960,769
Weighted share	10.46%	32.13%	57.41%	100%

[a]Defined as by Eurostat (see Appendix B, Section B.1).

[b]Not a couple (*e.g.,* flatmates).

[c]Defined as by the International Labour Organisation (see Appendix B.2).

Notes. Test of equal means is performed Wallonia with respect to Brussels, and Flanders with respect to Brussels, respectively. Significance codes: ***p < 0.01; **p < 0.05; and *p < 0.1.

Source. Author's calculations based on the Labour Force Survey (Eurostat) for Belgium, years 2004-2019.

Figure 1 and Figure 2 show the average activity and unemployment rate, respectively, during the period 2004-2019, for four individual characteristics in all three regions: gender, age, education level and country of birth. The activity rate is defined as the share of the working-age population that is either employed or actively seeking work (labour force). The unemployment rate measures the share of the labour force that is without a job but actively seeking and available for work. A thorough analysis of trends is relegated to Appendix A.

For most categories, the activity rate is higher in Flanders than in the other two regions. However, the Flemish Region presents also the largest cross-category gaps. This is the case for the age groups, where a 41 percentage points difference holds between individuals aged 45-54 and those 55-64 (the same gap is 28 in Brussels and 37 in Wallonia). It also applies to education at the lower end of the distribution. That is, the difference in the activity rate between people with primary (or lower) education and those who attained up to secondary education is 24 percentage points (as in Wallonia, but much higher than in Brussels, where it is only 15).[9] This evidence suggests that, although Flanders presents a higher average labour force participation, it should not be exempt from targeted group-specific policies (*e.g.*, to retain older workers or improve their employability).

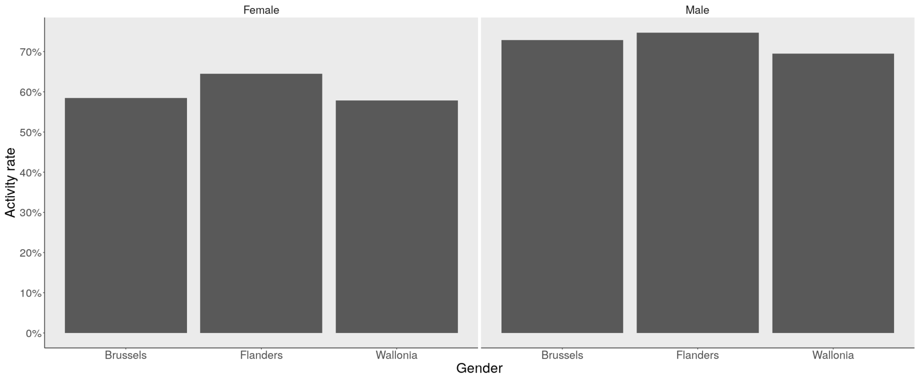
On the contrary, the most sizeable gaps in the unemployment rate between different groups of workers are in the Brussels-Capital Region,[10] where the unemployment rate is consistently higher than in Wallonia and Flanders for all categories. These gaps, in turn, seem to highlight some well-documented trends in Brussels labour market. Chiefly, the difference in the unemployment rate between individual aged 15-24 and those aged 55-64 is more than 20 percentage points. While youth unemployment is a particularly worrying problem in many other large cities in Europe (ESPON (2019)), in Brussels this is amplified by net negative workers' migration flows. In fact, as shown by Berns et al. (2022), young workers tend to leave the Capital Region once they have the economic means to do so. In addition, individuals with only primary education face an unemployment rate that is almost 5 times larger than the one for people holding a Master's (or higher) degree. Finally, workers originally from Africa face an unemployment rate that is about 15 percentage points larger than the one for those born in Belgium. However, it must be noted that this statistic does not account for informal work, where Arab and African workers are over represented (perspective.brussels (2020)).

[9]In the Brussels-Capital Region, larger gaps (compared to the other regions) are detectable rather at the upper end of the distribution according to education level. That is, in Wallonia and Flanders there is almost no difference in the activity rate of people holding a Bachelor's or a Master's. Contrarily, in Brussels, a 7 percentage points different stands between the two categories.

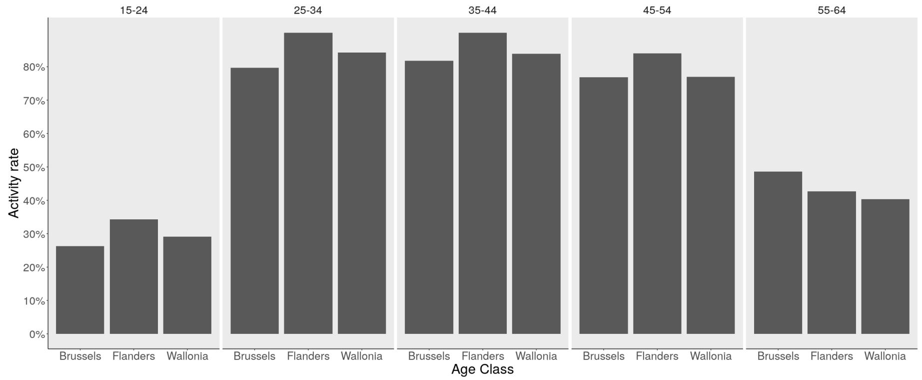
[10]Gaps in the unemployment rate also exist in the other two regions, but their size is much less remarkable.

Figure 1: Activity rate by population characteristics

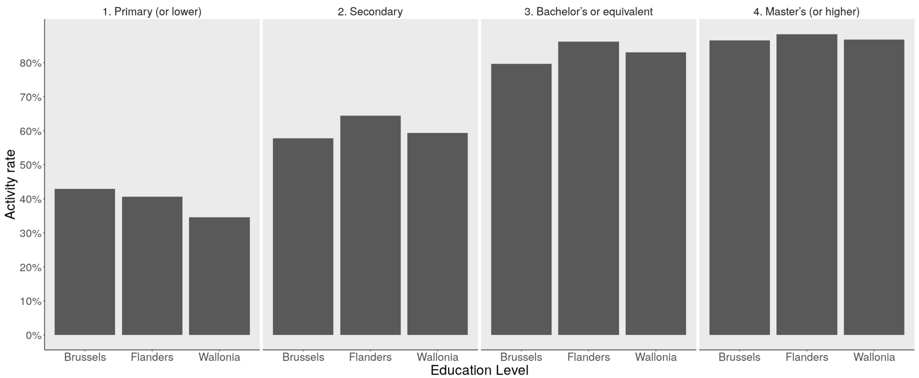
(a) Gender



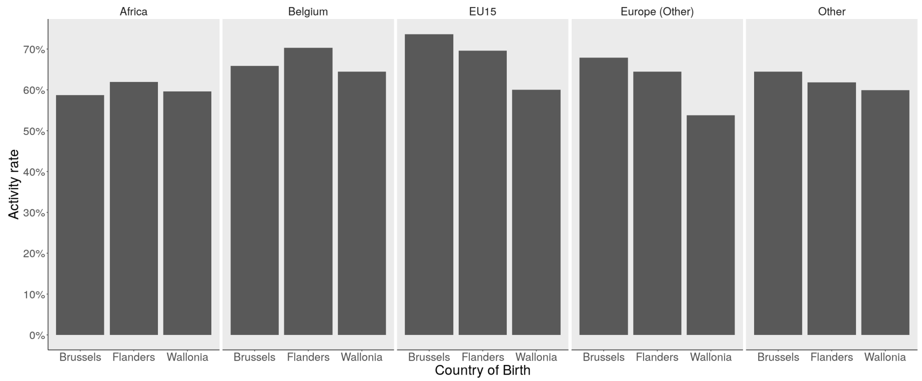
(b) Age



(c) Education level



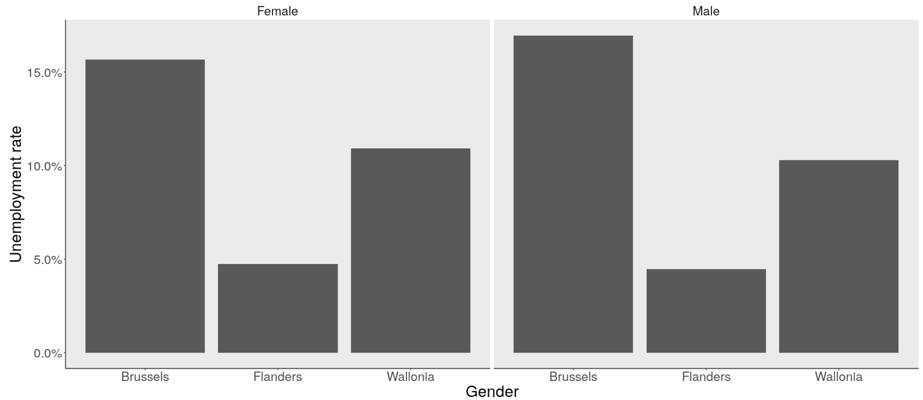
(d) Country of birth



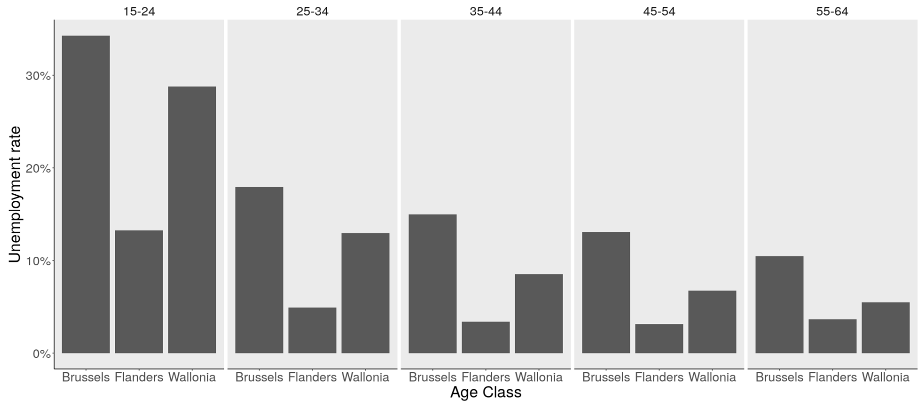
Source. Author's calculations based on the Labour Force Survey (Eurostat) for Belgium, years 2004-2019.

Figure 2: Unemployment rate by population characteristics

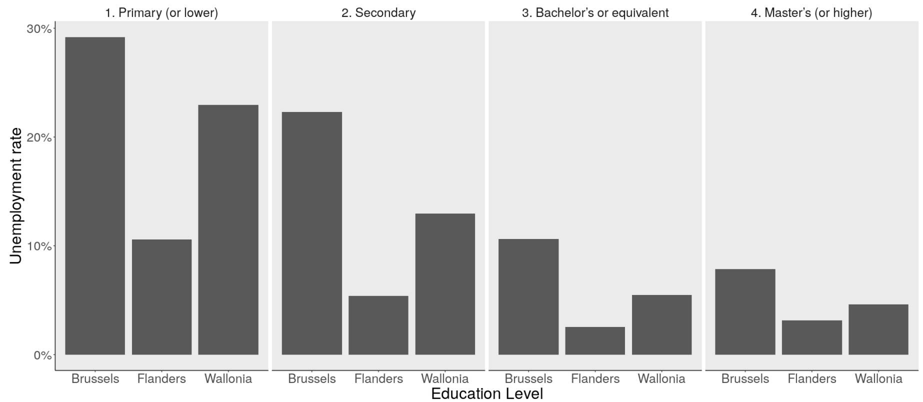
(a) Gender



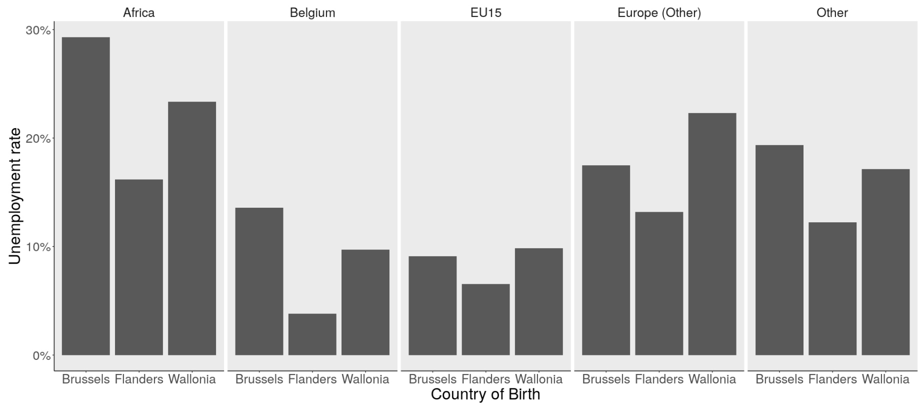
(b) Age



(c) Education level



(d) Country of birth



Source. Author's calculations based on the Labour Force Survey (Eurostat) for Belgium, years 2004-2019.

3 Methodology

From an econometric point of view, this paper employs the Heckman model to account for selection in labour force participation. In particular, the methodology employed in the present study is broadly consistent with that of Marelli and Vakulenko (2016), who use the Heckman model similarly in a joint focus on Italy and Russia.

The labour status "unemployed" is mutually exclusive with any other labour status. That is, every individual either is unemployed or is not. Therefore, it is defined by the following binary variable.

$$Y_i = \begin{cases} 1, & \text{if } \textit{unemployed} \\ 0, & \text{otherwise} \end{cases}$$

Given this specification, the probability for an individual to be unemployed $P(Y_i = 1|X = x_i)$ is determined by a binary choice model, where x_i is a vector of variables including individual characteristics and macroeconomic controls. Thus, the probability that an individual is unemployed is modelled using the following probit specification.

$$P(Y_i = 1|X = x_i) = \Phi(x_i'\beta) \tag{1}$$

where $\Phi(.)$ denotes the cumulative distribution function (CDF) of the standard normal distribution, x_i is the vector of explanatory variables, and β is the vector of parameters. Formally, the probit regression is estimated using the latent equation $y_i^* = x_i'\beta + \epsilon_i$, $\epsilon_i \sim N(0, 1)$, such that $Y_i = 1$ if $y_i^* > 0$. However, I can observe the outcome $Y_i = 1$ (*i.e.*, the individual is unemployed) only among the active population (labour force, which equals the total of both employed and unemployed individuals). Given that the choice of being active in the labour force is not necessarily random, such sub-sampling (from the original sample of the working age population to the sub-sample of the labour force) might raise sample selection concerns. In other words, the sub-sample of the active population might not be randomly selected, and its employment in the estimation of the model could provide biased estimates. To account for non-random selection in labour participation, I estimate Equation 1 with the correction for sample selection developed by Heckman (1976). Precisely, I develop the following Heckman selection model.

Similarly to unemployment, being active is a mutually exclusive labour status in the economy (*i.e.*, every individual either is active or is not).[11] Hence, I assume that the individuals' labour force participation is determined by the following selection equation.

$$y_i^{active} = \begin{cases} 1, & \text{if } z_i'\gamma + u_i > 0 \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where z_i is a vector of explanatory variables (including all variables in x_i plus at least one exclusion restriction), γ is a vector of coefficients, and u_i is the error term capturing unobserved factors affecting an individual's labour force participation (assumed to follow a standard normal distribution, $u_i \sim N(0,1)$). Therefore, I model the sample selection effect as follows. I assume that $Y_i = 1$ (*i.e.*, the individual is unemployed) is observed only when $y_i^{active} = 1$ (*i.e.*, the individual is active). The underlying hypothesis of non-random selection can be expressed by $corr(\epsilon_i, u_i) = \rho \neq 0$. That is, should ρ be equal to zero, I can reject non-random selection and I do not need to correct for selection.[12]

The estimation proceeds in two steps. First, I remove from my sample of the working-age population people who are still in education. The reason for excluding this category lies in the design of the selection criterion, which is activity in the labour market. In fact, while these people are inactive by definition, their inactivity is normally a temporary status: for most people, education only delays entrance to the labour market (with higher skills), rather than permanently avoiding it. When I sub-sample the active population, I want to distinguish it from the purely inactive (*i.e.*, people who are inactive as a permanent "choice", *e.g.*, retirees). Therefore, individuals who are inactive as a temporary status due to education are excluded.

Second, I estimate jointly Equations 1 (outcome) and 2 (selection) using the maximum likelihood method[13] both for Belgium (Section 4.1) and for each of its regions (Section 4.2). The explanatory variables are individual demographic characteristics (gender, age, country of birth, education level), household characteristics (composition), characteristics of the place of residence (urban area and regional unemployment rate) and time effects to control for macro conditions. Marelli and Vakulenko (2016), who perform a similar exercise, use as the unique variables for the selection equation (Equation 2) two determinants of inactivity: student status and disability. As I exclude student status from the selection process (see the first step), I use only disability as an exclusion restriction. This relies on the identifying assumption that, conditional on participation, disability does not directly affect the probability of unemployment. In

[11]Inactivity is defined as by the International Labour Organisation (see Appendix B.2).
[12]This is tested in the estimation of the model (See Table 2 and Table 3).
[13]The correction for selection bias is implicit in the joint bivariate normal structure of the probabilities.

other words, once an individual is active in the labour market, their likelihood of being unemployed is primarily determined by labour market conditions, skills, and other socio-demographic characteristics rather than the presence of a disability.

To quantify the impact of the explanatory variables of interest on labour force participation and unemployment risk, the average marginal effects (AMEs) of covariates are computed in both the outcome and selection equations. For continuous variables, AMEs are calculated as the average of the individual-level partial effects.

$$\begin{aligned} AME_C^{outcome} &= \frac{1}{N}\sum_{i=1}^{N}\phi(x_i'\hat{\beta})\cdot\hat{\beta}_j \\ AME_C^{selection} &= \frac{1}{N}\sum_{i=1}^{N}\phi(z_i'\hat{\gamma})\cdot\hat{\gamma}_k \end{aligned} \tag{3}$$

where $\phi(\cdot)$ is the standard normal probability density function (PDF), $x_i'\hat{\beta}$ $(z_i'\hat{\gamma})$ is the predictor for observation i, and $\hat{\beta}_j$ $(\hat{\gamma}_j)$ is the estimated coefficient of the continuous covariate of interest. For binary variables, the AME is computed as the average discrete change in predicted probability when the variable changes from 0 to 1. This effect is calculated as in the following expressions.[14]

$$\begin{aligned} AME_D^{outcome} &= \Phi(a^{outcome}+\hat{\beta})-\Phi(a^{outcome}) \\ AME_D^{selection} &= \Phi(a^{selection}+\hat{\gamma})-\Phi(a^{selection}) \end{aligned} \tag{4}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF), $\hat{\beta}$ $(\hat{\gamma})$ is the estimated coefficient of the binary variable, and $a = \Phi^{-1}(\bar{P})$ is the inverse CDF of the average predicted probability $\bar{P}$ of outcome (selection).

4 Results

4.1 Belgium

The econometric results for Belgium are presented in Table 2. Results for the activity (selection) equation (Equation 2) are presented in Column 1. Results for the unemployment (outcome) equation within the Heckman selection model (Equation 1) are presented in Column 2. Col-

[14]Technically, Equation 4 corresponds to a marginal effect at the mean (MEM), since the true average marginal effect (AME) of a binary covariate is defined as $AME_D^{outcome} = \frac{1}{N}\sum_{i=1}^{N}\left[\Phi(x_i'\hat{\beta}+\hat{\beta})-\Phi(x_i'\hat{\beta})\right]$ $(AME_D^{selection} = \frac{1}{N}\sum_{i=1}^{N}[\Phi(z_i'\hat{\gamma}+\hat{\gamma})-\Phi(z_i'\hat{\gamma})])$, which is the average of the individual discrete changes in predicted probability across all observations. In practice, when the distribution of covariates is not highly skewed, MEMs are numerically very close to AMEs. They are also easier to interpret as they show the effect evaluated at the mean index, providing a clear measure of the effect size.

umn 3 reports the estimation of a corresponding probit equation without selection. Since ρ is statistically significant and different from zero (-0.090), it is relevant to control for non-random selection. Precisely, a negative ρ means that unobserved factors that increase the likelihood of being active are associated with a lower likelihood of being unemployed once active. The majority of the coefficients are significant at 0.01 level for all specifications. The average marginal effects (AMEs) related to education variables (Table C.1) are calculated as in Equation 4 and are further discussed and interpreted in this section.

Labour force participation and the risk of unemployment increasingly rises and falls, respectively, with the level of education (reference category: up to primary education). That is, the probability of being active in the labour market soars by, on average, 13.7 percentage points when completing secondary education, 19.6 percentage points when attained up to tertiary education, and 20.5 percentage points when a Master's or even a higher level is reached. Similarly, the probability of being unemployed shrinks by, on average, 4.3 percentage points when completing secondary education, 6.8 percentage points when attained up to tertiary education, and 7 percentage points when a Master's or a higher level is reached. The residual positive impact of holding up to primary education is partially captured by the intercept.

Albeit these results might not seem *per se* surprising, the magnitude of these coefficients reveals the presence of diminishing marginal returns to education. This means that, while initial investments in education yield significant improvements in labour market outcomes, the incremental benefits tend to decrease with higher levels of educational attainment. In line with human capital theory (Becker (1964)), diminishing marginal returns to education have also been found in other contexts (*e.g.*, Psacharopoulos and Patrinos (2004)), but they remain broadly understudied. Here, I find that the shift from primary to secondary education results in a 13.7 percentage point increase in the probability of being active and a 4.3 percentage point decrease in that of being unemployed. From secondary to tertiary education, additional 6 and 2.5 percentage points are gained, respectively. Yet, further education on top of a Bachelor's (or equivalent) improves both probabilities by less one percentage point. This suggests that, beyond this type of degree, the marginal benefits of a Master's degree or higher are relatively smaller (in increasing labour force participation and reducing unemployment risk).

To take macroeconomic conditions into account, I also control for features such as the regional unemployment rate.[15] This variable is significant and has a sizeable effect in both equations. Precisely, it impacts negatively the probability of being active, whereas it presents a positive coefficient in the estimation of the unemployment equation. This confirms that regional

[15] This is measured with the standard ratio (Total unemployed/Total labour force) in each region separately.

labour market conditions affect severely the chances of being unemployed.

Time dummies (not reported in the table) are positive and statistically significant for all years after 2009 in the selection equation, whereas they are not statistically significant in the unemployment equation.

Table 2: Heckman selection model for Belgium, 2004-2019

	Activity equation	**Unemployment equation**	
		Heckman selection	**No selection**
	(1)	**(2)**	**(3)**
Intercept	-53.267***	14.204***	11.434***
	(1.513)	(3.179)	(0.294)
Individual characteristics			
Male	0.557***	-0.087**	-0.074***
	(0.028)	(0.043)	(0.005)
Log(Age)	31.098***	-7.791***	-6.202***
	(0.841)	(1.796)	(0.166)
Log(Age)^2	-4.488***	0.999***	0.769***
	(0.117)	(0.257)	(0.024)
Married	-0.102**	-0.147***	-0.142***
	(0.037)	(0.050)	(0.007)
Disability	-1.917***	\	\
	(0.046)		
Living in a city	-0.066**	0.126**	0.128***
	(0.030)	(0.041)	(0.006)
Country of birth (BE)	0.382***	-0.317***	-0.313***
	(0.082)	(0.108)	(0.016)
Country of birth (EU15)	0.303**	-0.213*	-0.235***
	(0.096)	(0.127)	(0.019)
Country of birth (Other Europe)	-0.047	0.098	0.115***
	(0.097)	(0.125)	(0.019)
Country of birth (North Africa and Middle East)	-0.133	0.357***	0.370***
	(0.100)	(0.131)	(0.019)
Country of birth (Sub-Saharan Africa)	0.227**	0.269*	0.267***
	(0.111)	(0.140)	(0.020)
Household composition			
Couple	0.015	-0.438***	-0.437***
	(0.044)	(0.056)	(0.008)
More than one adult	0.128***	-0.292***	-0.304***
	(0.041)	(0.050)	(0.007)
Number of children	-0.042***	0.046***	0.044***
	(0.012)	(0.017)	(0.002)
Education			
Up to secondary	0.557***	-0.346***	-0.314***
	(0.038)	(0.066)	(0.008)
Bachelor's or equivalent	1.000***	-0.737***	-0.687***
	(0.049)	(0.083)	(0.010)
Master's (or higher)	1.118***	-0.778***	-0.732***
	(0.053)	(0.087)	(0.010)
Macroeconomic controls			
Regional unemployment rate	-1.824***	4.430***	3.970***
	(0.373)	(0.482)	(0.061)
Time effects (dummies)	included	included	included
Observations	832,496		632,899
Uncensored observations		632,899	
ρ (rho)	-0.090**		
	(0.013)		
Log-likelihood	-52,466.43		-15,0405

Notes. Standard errors in parentheses. Significance codes: ***p $<$ 0.01; **p $<$ 0.05; and *p $<$ 0.1. Column 1 presents the results of the selection equation (Equation 2), which represents the probability of being active in the labour market. Column 2 presents the results for the outcome equation after selection (Equation 1), which represents the probability of being unemployed. Column 3 presents the results for the unemployment equation (Equation 1) without selection. All three equations are estimated via probit models. The reference categories are: rest of the world, for the country of birth; single adult, for household composition; up to primary education, for education level. Disability is the exclusion restriction and enters only the selection equation.

4.2 Brussels, Wallonia, and Flanders

The analysis by region is carried out by estimating a single Heckman selection model with interaction terms for Wallonia and Flanders, leaving Brussels as the reference category. Thus, the coefficients for Wallonia and Flanders capture the deviation from the baseline region, Brussels. Table 3 presents the estimation results for the three regions. Columns 1, 2, and 3 show the results for the Brussels-Capital Region; Columns 4, 5, and 6 for the Walloon Region; and Columns 7, 8, and 9 for the Flemish Region. For every region, the first column presents the results for the activity (selection) equation (Equation 2). Results for the unemployment equation with correction (Equation 1) are shown in the second column. The third column presents the estimation of a corresponding probit equation without selection. The ρ is different from zero (-0.057).[16] Therefore, it is important to control for non-random selection. The average marginal effects (AMEs) of the variables of interest (Table C.1) are calculated as in Equations 3 and 4.

The estimated coefficients in Brussels, which is the baseline region, broadly align with the results at country levels (Table 2). The statistical significance of the coefficients varies across the various regions and equations. One general observation is that, contrary to labour force participation (Columns 4 and 7), regional effects on unemployment risk are overall not statistically significant when selection is considered (Columns 5 and 8). Interestingly, several of these coefficients gain statistical significance when the unemployment equation is re-estimated without selection, underscoring the relevance of accounting for selection into the labour force.

In terms of labour force participation, major gaps (with respect to Brussels) in terms of magnitudes of the coefficients highlight important structural cross-regional differences. Age has a further positive impact on the probability of being active in both Wallonia and Flanders (negative in both regions when Log(Age)^2 is considered), which might relate to the fact that the labour force in the capital is, on average, younger. The country of birth matters more in the Flemish Region than in Brussels. In Flanders, the labour force participation premium for being Belgian is almost 10 percentage points larger than in Brussels. However, the estimated coefficients indicate that non-European immigrants face the largest penalty associated with their foreign status in the Brussels Region. These correlations can be particularly harmful considering that approximately half of Brussels' population consists of individuals who were not born in Belgium.

Besides age, in Brussels, the largest coefficients, both in the selection and in the outcome equations, are associated with education, which increases and decreases labour force participa-

[16] Since the model is estimated jointly, there is a single value of ρ and a single log-likelihood for all three regions.

tion and unemployment risk, respectively. These effects are even more pronounced in Wallonia, where the probability of being active is further positive for all levels of education, increasingly soaring with each additional attainment.[17] In Flanders, the coefficients for the education levels are all not statistically significant, suggesting no relevant difference of these variables with respect to Brussels. In fact, statistical significance is observed only in the re-estimated unemployment equation without selection, where the sign of the coefficients reveals a positive effect of high education on unemployment compared to Brussels (Column 9). Specifically, when selection is not accounted for, an individual holding a Master or higher diploma would face a probability of being unemployed that is about 3 percentage points higher in Flanders than in Brussels.

Reasons for differences in returns to education across Belgian regions can be manifold, including industrial specialisations and demographic profiles. Flanders has a more diversified economy with lower unemployment, which may limit the additional employment gains from higher education. Since the decline of its traditional industries (steel and coal), Wallonia has been gradually reinventing its economy by focusing on new sectors (e.g., ICT, digital technologies, and renewable energy) that require highly skilled workers. In Brussels, the working-age population is highly polarised in terms of education: almost 25% of individuals hold a Master's or higher diploma, while around 15% have attained at most primary education. Although the composition of the working-age population does not directly explain the effects of education on labour force participation and unemployment risk, it is likely to partly reflect dynamics in the labour demand, which in Brussels is increasingly skewed toward highly qualified workers. While some jobs remain accessible to the least qualified, the share of positions for the so-called "middle class" is steadily shrinking. In fact, secondary education, which is the average level of middle-class workers, decreases the probability of being unemployed by less than 5 percentage points, while holding a Master's or higher degree does so by more than 12.

To account for macroeconomic conditions, I control only for time fixed effects, since there is no remaining variation in the regional unemployment rate when including the three regions. The variable "Living in a city" is also excluded, as it has no variation within Brussels.

[17] Accordingly, they are negative for the probability of being unemployed. Yet, they are not statistically significant.

Table 3: Heckman selection model for Brussels, Wallonia and Flanders, 2004-2019

	Brussels			**Wallonia**			**Flanders**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	-42.382***	10.957	9.642***	-4.091**	0.169	-0.894	-19.154***	6.504	6.549***
	(1.917)	(7.290)	(0.735)	(2.190)	(7.985)	(0.862)	(2.121)	(7.635)	(0.871)
Individual characteristics									
Male	0.638***	-0.009	0.004	-0.095**	-0.097	-0.111***	-0.073**	-0.083	-0.090***
	(0.033)	(0.170)	(0.011)	(0.037)	(0.122)	(0.014)	(0.036)	(0.117)	(0.014)
Log(Age)	24.342***	-5.684***	-4.925***	2.640**	0.320	0.819*	11.547***	-3.961	-3.986***
	(1.066)	(0.231)	(0.412)	(1.218)	(4.484)	(0.485)	(1.180)	(4.284)	(0.489)
Log(Age)^2	-3.463***	0.695***	0.584***	-0.442***	-0.068	-0.128*	-1.724***	0.588	0.594***
	(0.148)	(0.032)	(0.058)	(0.169)	(0.628)	(0.068)	(0.164)	(0.599)	(0.069)
Married	-0.186***	-0.062	-0.048***	0.141***	-0.094	-0.121***	0.077	-0.119	-0.130***
	(0.044)	(0.140)	(0.016)	(0.050)	(0.164)	(0.019)	(0.049)	(0.157)	(0.019)
Disability	-1.878***	\	\	0.207***	\	\	-0.256***	\	\
	(0.070)			(0.079)			(0.076)		
Country of birth (BE)	0.233***	-0.190	-0.195***	0.059	-0.125	-0.155***	0.402***	-0.343	-0.326***
	(0.072)	(0.227)	(0.026)	(0.107)	(0.342)	(0.041)	(0.089)	(0.269)	(0.037)
Country of birth (EU15)	0.204**	-0.262	-0.267***	0.021	0.043	-0.005	0.292***	0.044	0.037
	(0.082)	(0.262)	(0.030)	(0.118)	(0.382)	(0.046)	(0.104)	(0.319)	(0.045)
Country of birth (Other Europe)	0.028	-0.038	-0.014	-0.271**	0.318	0.304***	-0.039	0.177	0.175***
	(0.080)	(0.250)	(0.029)	(0.124)	(0.395)	(0.049)	(0.102)	(0.302)	(0.043)
Country of birth (North Africa and Middle East)	-0.208***	0.364	0.356***	-0.001	0.099	0.105**	0.037	-0.014	0.035
	(0.078)	(0.249)	(0.028)	(0.124)	(0.391)	(0.049)	(0.104)	(0.313)	(0.045)
Country of birth (Sub-Saharan Africa)	0.046	0.318	0.306***	0.141	-0.028	-0.037	0.324***	-0.089	-0.076
	(0.090)	(0.274)	(0.031)	(0.133)	(0.407)	(0.050)	(0.122)	(0.342)	(0.049)
Household composition									
Couple	-0.058	-0.345**	-0.359***	0.144**	-0.227	-0.181***	0.039	-0.029	-0.024
	(0.049)	(0.153)	(0.017)	(0.057)	(0.179)	(0.021)	(0.056)	(0.172)	(0.022)
More than one adult	0.058	-0.285**	-0.305***	0.114**	-0.128	-0.092***	0.061	0.056	0.057***
	(0.046)	(0.140)	(0.015)	(0.053)	(0.161)	(0.019)	(0.053)	(0.158)	(0.019)
Number of children	-0.048***	0.058	0.059***	0.010	-0.013	-0.020***	0.001	-0.034	-0.033***
	(0.014)	(0.046)	(0.005)	(0.016)	(0.054)	(0.006)	(0.016)	(0.052)	(0.007)
Education									
Up to secondary	0.518***	-0.217	-0.206***	0.119**	-0.208	-0.181***	-0.042	-0.067	-0.079***
	(0.043)	(0.166)	(0.017)	(0.050)	(0.190)	(0.022)	(0.049)	(0.187)	(0.023)
Bachelor's or equivalent	0.878***	-0.635***	-0.626***	0.254***	-0.255	-0.187***	0.003	0.046	0.051*
	(0.059)	(0.207)	(0.022)	(0.068)	(0.236)	(0.027)	(0.066)	(0.230)	(0.028)
Master's (or higher)	1.012***	-0.757***	-0.742***	0.285***	-0.195	-0.119***	-0.072	0.259	0.257***
	(0.053)	(0.188)	(0.020)	(0.065)	(0.229)	(0.026)	(0.061)	(0.214)	(0.026)
Macroeconomic controls									
Time effects (dummies)	included	included	included	included	included	included	included	included	included
Observations	105,066		105,066	314,554		314,554	412,876		412,876
Uncensored observations		79,094			231,567			322,238	
ρ (rho)	-0.057*			-0.057*			-0.057*		
	(0.012)			(0.012)			(0.012)		
Log-likelihood	-51,907.86		-14,9493	-51,907.86		-14,9493	-51,907.86		-14,9493

Notes. Standard errors in parentheses. Significance codes: ***p < 0.01; **p < 0.05; and *p < 0.1. For each region (Brussels, Wallonia, and Flanders), the first column (Columns 1, 3, and 5) presents the results of the selection equation (Equation 2), which represents the probability of being active in the labour market. The second column (Columns 2, 4, and 6) presents the results for the outcome equation after selection (Equation 1), which represents the probability of being unemployed. The third column (Columns 3, 6, and 9) presents the results for the unemployment equation (Equation 1) without selection. All equations are estimated via probit models. The Heckman selection model is estimated jointly for the three regions with interaction terms (Brussels is the reference category). The reference categories are: rest of the world, for the country of birth; single adult, for household composition; up to primary education, for education level. Disability is the exclusion restriction and enters only the selection equation.

5 Conclusions

In this paper, I delve into the differences in the labour market among the three regions in Belgium: Brussels, Wallonia and Flanders. I use the Labour Force Survey data (for the years 2004-2019) to identify the main personal and family characteristics correlated with activity and unemployment, both at national level and for each of the three regions. To run this exercise, I use a Heckman selection model to minimise the sample selection bias in the estimation of the unemployment risk.

I find a substantial impact of education in increasing and decreasing labour force participation and unemployment risk, respectively. At the regional level, higher education plays a crucial role in individuals' employment in Brussels, and even more in Wallonia. Statistically significant differences are not found between the Brussels-Capital Region and the Flemish Region. While a negative correlation of education with unemployment is not *per se* surprising, the structure of the labour demand can sharply amplify its scale. In Brussels, for example, such an enormous unemployment gap between the tails of the education distribution is likely to be related to the labour demand, whose sharp rise in skill requirements is leaving the least educated with less and less job opportunities. In general, Brussels reports the highest unemployment rates, the most variegated composition of workers and the most peculiar trends (see Appendix A) when looking at characteristics affecting the probability of being active and the probability of being unemployed. Consequently, in the capital, several categories of workers, including the low-to-middle educated and people with foreign origins, face much higher and persistent unemployment than other kinds of workers.

Albeit a labour demand oriented towards higher and higher qualified workers certainly certifies progress, a good functioning of the economy linked to low-skilled jobs also contributes to prosperity and social stability. Employment problems among the least qualified cannot be addressed solely through spillovers from growth in the dominant sectors. On the contrary, the emergence of a dual-structured economy, where gains are concentrated among higher-skilled or higher-income groups, may exacerbate existing socioeconomic disparities.

While the coefficients resulting from the estimations in Section 4 provide the sign and size of the effects of the explanatory variables at the level of correlation, they do not allow for any causal inference. Additionally, whether these effects originate from the demand or the supply side cannot be derived from the data available for this study. Despite these limitations, the present study offers a valuable framework to guide future analyses in uncovering causal mechanisms and the pathways through which individual characteristics influence labour force participation and unemployment risk.

References

Becker, G. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Chicago Press.

Berns, H., Lenel, E., Schaut, C., & Van Hamme, G. (2022). Towards a paradigm shift in the residential appeal policy of the Brussels-Capital Region. *Brussels Studies, Collection générale*, (172).

Billiet, J., Maddens, B., & Frognier, A. (2006). Does Belgium (still) exist? Differences in political culture between Flemings and Walloons. *West European Politics*, *29*(5).

Biot, V., Van Hamme, G., & Wertz, I. (2019). Youth unemployment: Territorial trends and regional resilience. *Synthesis Report*.

Brooks, B. (2005). Chronic unemployment: A statistical profile. *Statistics Canada*, (Analysis in Brief 31).

Brügger, B., Lalive, R., & Zweimüller, J. (2009). Does culture affect unemployment? Evidence from the Röstigraben. *IZA Discussion Paper*, (4283).

Buysse, K., & Saks, Y. (2020). Drivers of labour force participation in Belgium. *NBB Economic Review*.

ESPON. (2019). Youth unemployment: Territorial trends and regional resilience. *Synthesis Report*.

Estevào, M. (2003). Regional labor market disparities in Belgium. *Reflets et perspectives de la vie économique, De Boeck Université*, (1), 95–114.

Eurostat, E. C. (2020). EU Labour Force Survey Database - User Guide. *Unit F-3: Labour market and lifelong learning*.

Feld, L. P., & Savioz, M. R. (2000). Cantonal and regional unemployment in Switzerland: A dynamic macroeconomic panel analysis. *Schweiz. Zeitschrift für Volkswirtschaft und Statistik*, *136*(3), 463–483.

Gérard, M., Jayet, H., & Paty, S. (2010). Tax interactions among Belgian municipalities: Do interregional differences matter? *Regional Science and Urban Economics*, *40*(5).

Heckman, J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement*, (5/4).

Hoorelbeke, D. (2013). Regional disparities in GDP and disposable income in Belgium. *EcoMod*, (5327).

IMF. (2002). Regional labor markets in Belgium. *IMF Staff Country Reports*, (43).

Kostoris, F., & Lupi, C. (2002). Family income and wealth, youth unemployment and active labour market policies, international review of applied economics. *International Review of Applied Economics*, *16*(4), 407–416.

Kulu, H., & González-Ferrer, A. (2013). Family dynamics among immigrants and their descendants in Europe: Current research and opportunities. *Families and Society Working paper Series*, (3).

Marelli, E., Patuelli, R., & Signorelli, M. (2012). Regional unemployment in the EU before and after the global crisis. *Post-Communist Economies*, (2).

Marelli, E., & Vakulenko, E. (2016). Youth unemployment in Italy and Russia: Aggregate trends and individual determinants. *The Economic and Labour Relations Review*, *27*(3), 387–405.

Meunier, O., & Mignolet, M. (2005). Regional employment disparities in Belgium: Some empirical results. *Centre de Recherche sur l'Economie Wallonne, University of Namur, Belgium.*

Mincer, J. (1974). Schooling, Experience and Earnings. *Columbia University Press, New York.*

OECD. (2016). *OECD regions at a glance 2016* (https://doi.org/10.1787/reg_glance-2016-26-en). OECD Publishing.

OECD. (2020). *OECD regions and cities at a glance 2020* (https://doi.org/10.1787/959d5ba0-en). OECD Publishing.

OECD. (2022). *OECD regions and cities at a glance 2022* (https://doi.org/10.1787/14108660-en). OECD Publishing.

perspective.brussels. (2020). *Redéploiement socio-économique, territorial et environnemental suite à la crise du Covid-19*. IBSA.

Plasman, R., Rusinek, M., & Tojerow, I. (2008). The regional dimension of collective wage bargaining: The case of Belgium. *IZA Discussion Papers*, (3864).

Psacharopoulos, G., & Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education Economics*, *12*(2).

Schmillen, A., & Möller, J. (2010). Determinants of lifetime unemployment: A micro data analysis with censored quantile regressions. *IZA Discussion Paper*, (4751).

Tandé, A. (2018). Fight against discrimination or promotion of diversity in employment? A public policy in the Brussels-Capital Region (1997-2012). *Brussels Studies, General collection*, (120).

Van Oyen, H., Tafforeau, J., & Roelands, M. (1996). Regional inequities in health expectancy in Belgium. *Social Science & Medicine*, *43*(11).

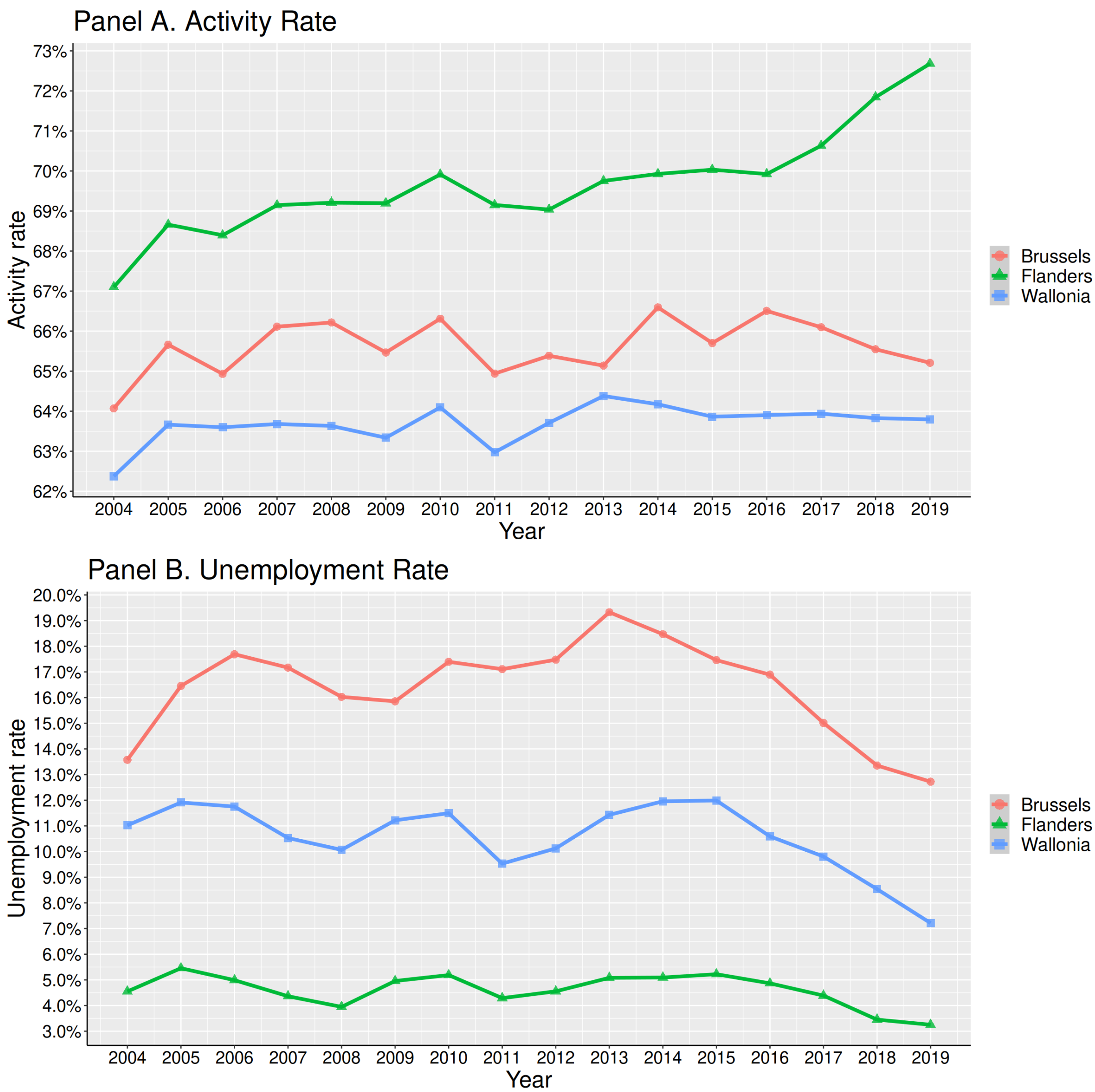
Vandenberghe, V. (2010). Inter-regional educational discrepancies in Belgium. how to combat them? *Re-Bel initiative*.

Zimmer, H. (2012). Labour market mismatches. *NBB Economic Review*.

Appendices

A Trends in the Activity and Unemployment Rate

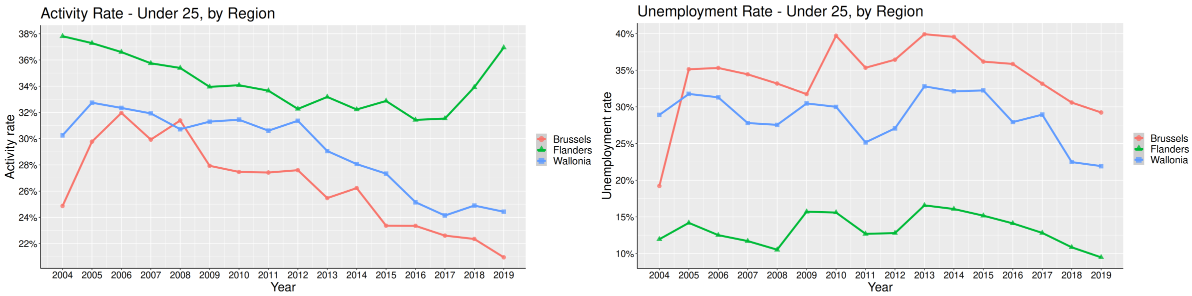
Figure A.1: Activity and unemployment rates in the Belgian regions



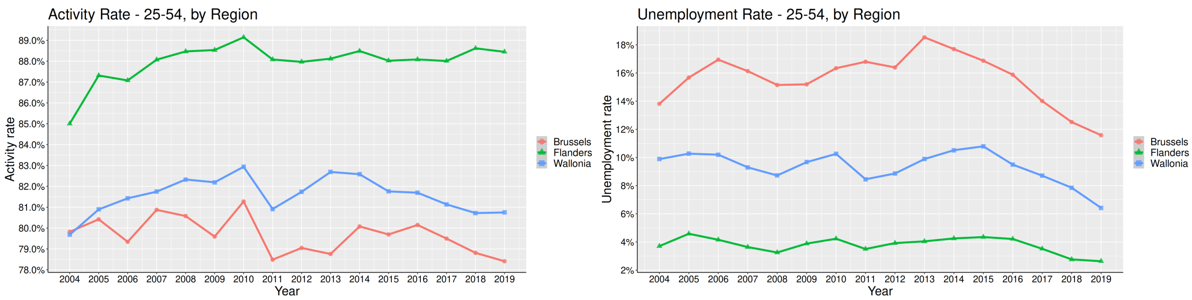
Source. Author's calculations based on the Labour Force Survey (Eurostat) for Belgium, years 2004-2019.

Figure A.2: Trends in the activity and unemployment rate, by age group

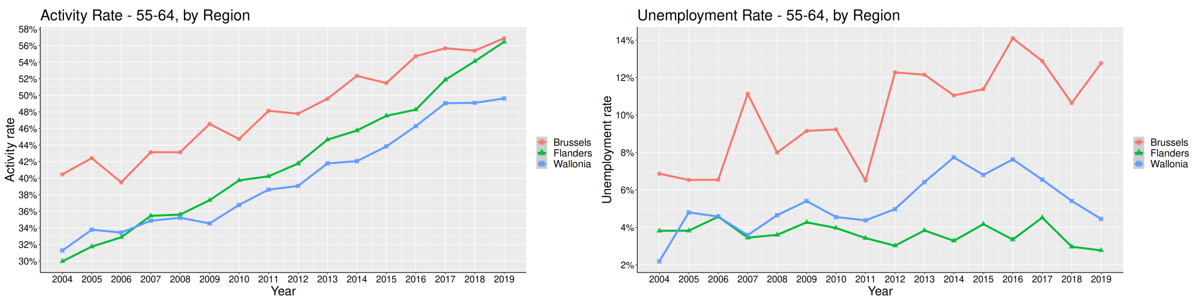
(a) Under 25



(b) 25-54



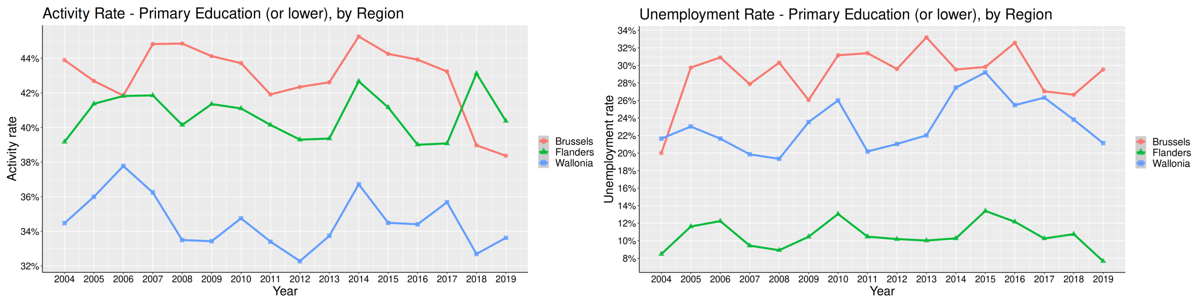
(c) 55-64



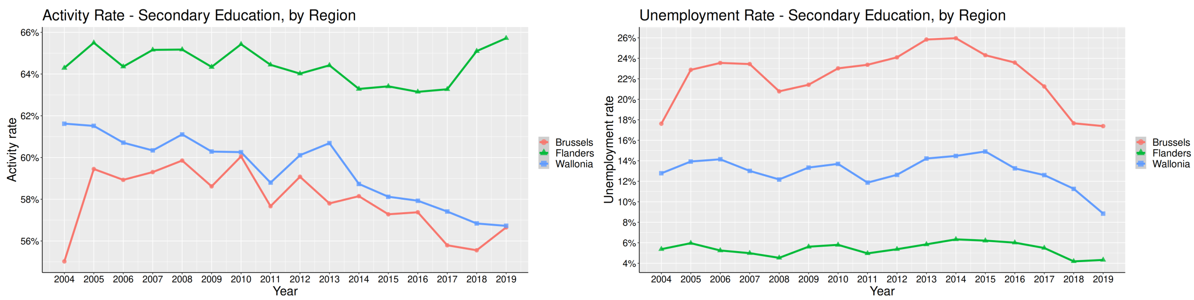
Source. Author's calculations based on the Labour Force Survey (Eurostat) for Belgium, years 2004-2019.

Figure A.3: Trends in the activity and unemployment rate, by education level

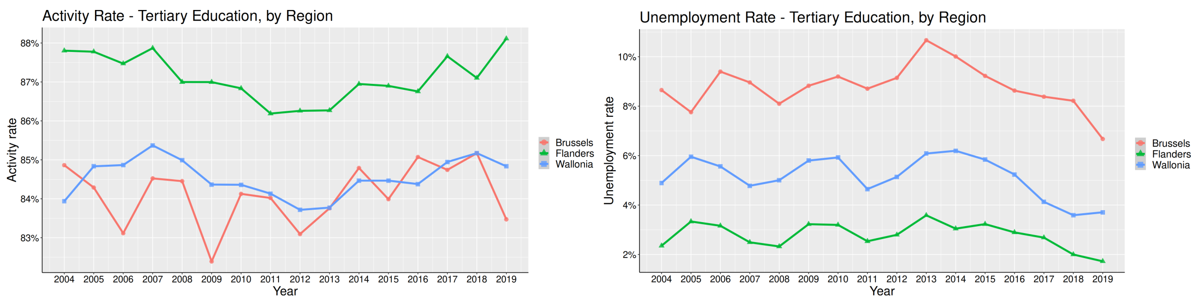
(a) Primary Education (or lower)



(b) Secondary Education



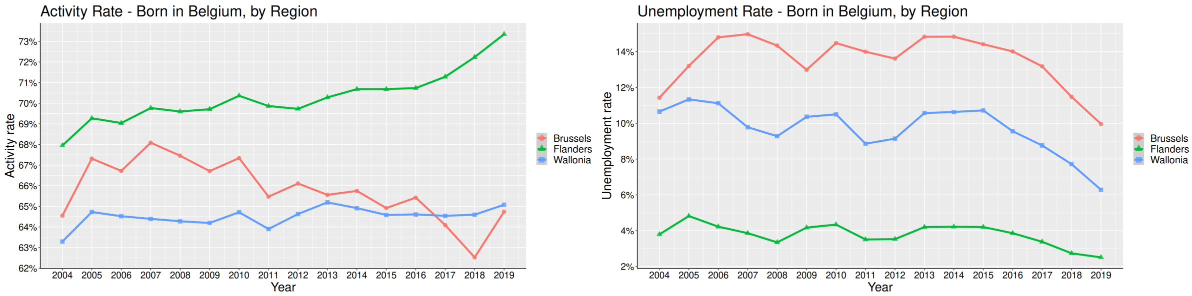
(c) Tertiary Education



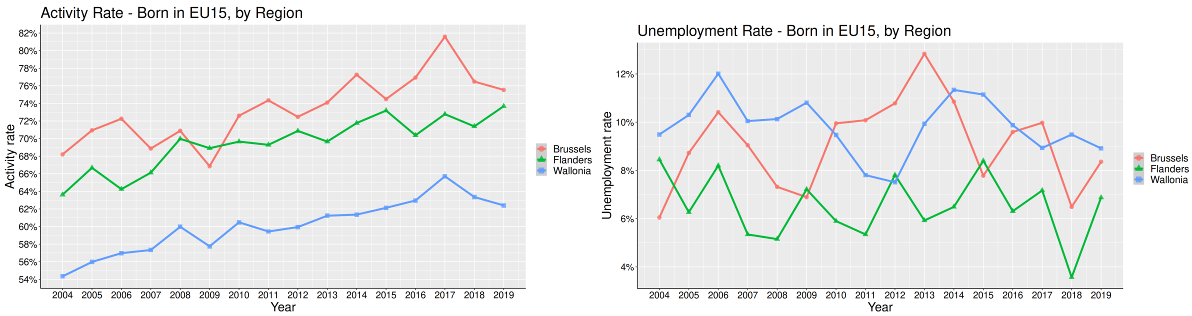
Source. Author's calculations based on the Labour Force Survey (Eurostat) for Belgium, years 2004-2019.

Figure A.4: Trends in the activity and unemployment rate, by country of birth

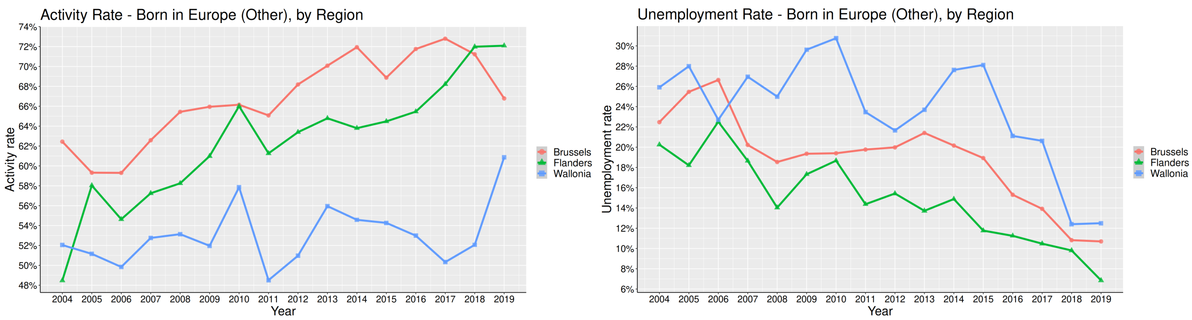
(a) Belgium



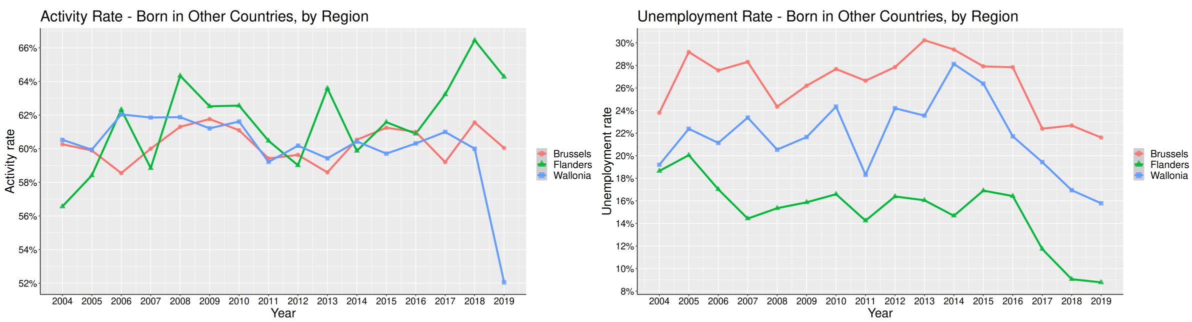
(b) EU15



(c) Europe (other)



(d) Other



Source. Author's calculations based on the Labour Force Survey (Eurostat) for Belgium, years 2004-2019.

B Definitions

B.1 Degree of Urbanisation Classification (Source: Eurostat)

The degree of urbanisation of a certain area in the LFS data is defined by Eurostat in two consecutive steps.

First step: Definition of clusters.

- High-density cluster/urban centre: contiguous grid cells of 1 km^2 with a density of at least 1,500 inhabitants per km^2 and a minimum population of 50,000;
- Urban cluster: cluster of contiguous grid cells of 1 km^2 with a density of at least 300 inhabitants per km^2 and a minimum population of 5,000;
- Rural grid cell: grid cell outside high-density clusters and urban clusters.

Second step: Definition of areas.

- Cities or large urban areas: at least 50% lives in high-density clusters; in addition, each high-density cluster should have at least 75% of its population in a large urban area; this also ensures that all high-density clusters are represented by at least one large urban area, even when this cluster represents less than 50% of the population of that area;
- Towns and suburbs or small urban areas: less than 50% of the population lives in rural grid cells and less than 50% live in high-density clusters;
- Rural areas: more than 50% of the population lives in rural grid cells.

B.2 Categorisation of the Labour Status (Source: ILO)

The labour status classification in the LFS data follows the International Labour Organisation's definition, according to which three mutually exclusive labour statuses exist.

- Employed: the individual worked 1+ hours during the reference week *or* the individual did not work but had a job or business from which he/she was absent during the reference week;
- Inactive (it includes people in education and early retirement): the individual is not looking for a job *or* the individual has found a job that will start after 3+ months *or* the individual is seeking a job but he/she is not available to start within the following 2 weeks;

– Education: the individual received some education (formal) or training (non-formal) during the previous four weeks;

- Unemployed: the individual is seeking a job and he/she is available to start within the following 2 weeks *or* the individual has found a job that will start within 3 months.

C Average Marginal Effects

Table C.1: Average marginal effects for binary education variables

	Belgium			**Brussels**			**Wallonia**			**Flanders**		
	(i)	(ii)	(iii)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Up to secondary	0.137	-0.043	-0.039	0.132	-0.049	-0.046	0.038	-0.030	-0.027	-0.012	-0.006	-0.007
Bachelor's or equivalent	0.196	-0.068	-0.066	0.188	-0.111	-0.110	0.076	-0.036	-0.028	0.001	0.004	0.005
Master's (or higher)	0.205	-0.071	-0.068	0.202	-0.122	-0.121	0.085	-0.029	-0.018	-0.021	0.031	0.031

Notes. For Belgium and each of its regions (Brussels, Wallonia, and Flanders), the first column (Columns i, 1, 3, and 5) presents the average marginal effects (AMEs) of the selection equation (Equation 2), which represents the probability of being active in the labour market. The second column (Columns ii, 2, 4, and 6) presents the AMEs for the outcome equation after selection (Equation 1), which represents the probability of being unemployed. The third column (Columns iii, 3, 6, and 9) presents the AMEs for the unemployment equation (Equation 1) without selection. AMEs for binary education variables are calculated as in Equation 4 based on the estimated coefficients in Tables 2 and 3. They quantify the percentage point change in the corresponding probability (labour force participation or unemployment risk) with respect to the reference category: up to primary education, for Belgium and Brussels, and Brussels for Wallonia and Flanders.