DISCUSSION PAPER SERIES

IZA

IZA DP No. 9954

Faces of Joblessness: Characterising Employment Barriers to Inform Policy

Rodrigo Fernandez
Herwig Immervoll
Daniele Pacifico
Céline Thévenot

May 2016

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Faces of Joblessness: Characterising Employment Barriers to Inform Policy

Rodrigo Fernandez
OECD

Herwig Immervoll
OECD and IZA

Daniele Pacifico
OECD

Céline Thévenot
OECD

Discussion Paper No. 9954
May 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 9954
May 2016

ABSTRACT

Faces of Joblessness: Characterising Employment Barriers to Inform Policy*

This paper proposes a novel method for identifying and visualising key employment obstacles that may prevent individuals from participating fully in the labour market. The approach is intended to complement existing sources of information that governments use when designing and implementing activation and employment-support policies. In particular, it aims to provide individual and household perspectives on employment problems, which may be missed when relying on common labour-force statistics or on administrative data, but which are relevant for targeting and tailoring support programmes and related policy interventions. A first step describes a series of employment-barrier indicators at the micro level, comprising three domains: work-related capabilities, financial incentives and employment opportunities. For each domain, a selected set of concrete employment barriers are quantified using the EU-SILC multi-purpose household survey. In a second step, a statistical clustering method (latent class analysis), is used to establish profiles and patterns of employment barriers among individuals with no or weak labour-market attachment. A detailed illustration for two countries (Estonia and Spain) shows that "short-hand" groupings that are often highlighted in the policy debate, such as "youth" or "older workers", are in fact composed of multiple distinct sub-groups that face very different combinations of employment barriers and likely require different policy approaches. Results also indicate that individuals typically face two or more simultaneous employment obstacles suggesting that addressing one barrier at a time may not have the intended effect on employment levels. From a policy perspective, the results support calls for carefully sequencing activation and employment support measures, and for coordinating them across policy domains and institutions.

JEL Classification: C38, J08, H31, J21, J22, J68, J82

Keywords: unemployment, employment barrier, activation, targeting, latent class, active labour market programmes

Corresponding author:

Herwig Immervoll
OECD
2, rue André Pascal
75016 Paris
France
E-mail: herwig.immervoll@oecd.org

* This document was produced with the financial assistance of the European Union Programme for Employment and Social Innovation "EaSI" (2014-2020, EC OECD grant agreement VS/2016/0005, DI150038), as part of a joint project between the European Commission (EC), the OECD and the World Bank. The authors gratefully acknowledge feedback on earlier drafts from a number of EC and OECD colleagues and, in particular, multiple rounds of discussions with the World Bank (Aylin Isik-Dikmelik, Sandor Karacsony, Natalia Millan, Mirey Ovadiya, Frieda Vandeninden and Michele Davide Zini), and the extensive in-depth comments they provided on the empirical results, the underlying concepts and Stata programs. The paper also accounts for feedback received during and after a project kick-off seminar, held at OECD in Paris on 3 March 2016 with the participation of the EC, World Bank and representatives from twelve countries, and during a seminar held at the European Commission, DG EMPL, on 17 February 2016. All results and opinions are the sole responsibility of the authors. In particular, the opinions expressed and arguments employed herein do not necessarily reflect the official views of the OECD member countries, of the European Union, the World Bank, or of the governments of countries referenced in the text or results.

TABLE OF CONTENTS

Tables

Figures

Boxes

FACES OF JOBLESSNESS: CHARACTERISING EMPLOYMENT BARRIERS TO INFORM POLICY

Rodrigo Fernandez (OECD)
Herwig Immervoll (OECD, IZA)
Daniele Pacifico (OECD)
Céline Thévenot (OECD)

1. INTRODUCTION

The literature on activation and employment-support policies (AESPs), and on social protection systems more generally, commonly emphasises targeting and tailoring of policy interventions to individual circumstances as crucial factors for policy success (OECD, 2013a, 2013b, 2014a, 2015a; Immervoll and Scarpetta, 2012; Arias et al., 2014; World Bank, 2013; European Commission 2015; Eurofound, 2012). Yet, relatively little is known about what these individual circumstances look like or how they may translate into employment barriers that policies aim to address. The approach outlined in this paper aims to contribute to the debate of employment, activation and social inclusion policies through an improved understanding of the characteristics and labour-market barriers of out-of work and "low-work-intensity" individuals. The paper proposes and applies a simple statistical method for guiding the design and targeting of AESPs. It is intended as a step-by-step guide for researchers and practitioners who wish to apply and adapt it to different country contexts.

In the absence of comprehensive information on employment barriers, policy discussions frequently refer to broad groupings of individuals as a short-hand for the specific difficulties facing jobless individuals or those whose employment is, in some sense, precarious. "Youth", "older workers", "recipients of disability benefits", or "lone parents" are examples for proxy groupings that are frequently used in the policy debate. An implicit assumption is that these groupings are useful for describing different sets of employment barriers that may inform policy design and implementation. This may or may not be true; for instance being young is, in and of itself, not an employment barrier.

There has been little attempt to systematically assess the specific barriers of policy clients whom AESPs are intended to help. The lack of suitable "profiles" of individuals with no or weak labour-market attachment is particularly limiting in a comparative perspective. For instance, standard tabulations from commonly available international labour-market statistics provide two- or three-dimensional breakdowns of employment status that are valuable for many purposes. But, for a number of reasons, they are insufficient as basis for discussing gaps in existing AESPs, for shedding light on specific targeting or design issues, and for guiding the design of tailored policy interventions.

First, standard labour-market statistics provide information about correlates of employment barriers (e.g., age or the number of children), but not about employment barriers as such (e.g., limited work experience or care responsibilities). They also provide little to no information on the incidence of multiple *simultaneous* barriers, which can be a major challenge for policy effectiveness. Second, available comparative statistics typically adopt an individual perspective and lack data on the socio-economic household context which shapes individual employment opportunities and incentives. Third, to facilitate comparability, standard labour-market statistics tend to use identical categories across countries. As a result, the format and content of available information may not be equally well-suited for documenting labour-market challenges in different national contexts.

At the national level, several countries have developed powerful profiling tools that provide policy makers with customised and timely views on the circumstances of specific groups of policy clients (e.g., Bimrose and Barnes, 2011; Konle-Seidl, 2011). Results rely on administrative data which provide very rich information on the dimensions they cover. But they are often limited to specific sub-groups (such as the registered unemployed), while others with no or weak labour market attachment remain out of scope. As a result, profiling tools are geared towards refining employment-support and reintegration processes at the individual level, in the institution that develops and uses them (e.g., a public employment service, PES). They may be less useful at providing input into broader policy-design questions that require a "birds-eye" view on people's employment barriers irrespective of the institution they are (or are not) registered with. Examples of such broader policy-design questions are priorities for linking up services between relevant institutions, or identifying groups that could benefit from policy support but who are currently not easily reachable by institutions providing such support (e.g., because they are not registered as unemployed).

Based on the premise that AESPs seek to alleviate specific employment barriers, this paper advocates a bottom-up approach to policy analysis, starting from a careful and country-specific assessment of these barriers. We then use a statistical clustering method to separate the highly heterogeneous population of individuals with no or weak attachment into groups ("clusters") that are homogeneous with respect to the types of employment barriers that they face.

The resulting profiles of employment barriers are intended to facilitate discussions of the strengths and limitations of different policy interventions for concrete groups of policy clients. They can also be used to help inform decisions on whether to channel additional efforts towards specific priority groups. The approach is related to existing studies that describe characteristics of different groups of individuals with labour-market difficulties (European Commission, 2012; Ferré *et al.,* 2013; Immervoll, 2013; Sundaram *et al.*, 2014). Compared with those earlier exercises, the present study considers a broader set of labour-market problems and adopts an explicit conceptual framework that facilitates its use by practitioners and provides a reference for future methodological extensions and country-specific applications.

The paper first outlines three main domains of employment barriers that potentially drive poor labour-market outcomes: *(i)* a lack of work-related capabilities, *(ii)* a lack of work incentives and *(iii)* a lack of employment opportunities. It then suggests a set of workable indicators in each domain that can be implemented for selected EU countries using EU-SILC data at the individual or family level. In a final step, we identify groups of individuals facing similar combinations of employment barriers. An illustration of the clustering approach using data for two countries (Estonia and Spain) suggests the following:[1]

- A **large number of different employment barrier profiles** characterize the population out of work or 'underemployed'.

- Proxy groupings that are commonly referred to in the policy debate, such as "youth", "NEETs", "older workers", etc., are not homogeneous but include **distinct sub-groups with very different**

1 . Results shed some light on the patterns of labour-market difficulties, employment barriers and policy priorities in seven selected EU countries. They are, however, meant to illustrate the proposed methodological approaches, rather than provide succinct policy conclusions. For in-depth analyses that consider country-specific contexts and data quality in more detail, readers are referred to a series of forthcoming country studies undertaken by EC, OECD and World Bank as part of a joint project (see, e.g., http://www.oecd.org/social/faces-of-joblessness.htm).

employment barrier profiles. This finding highlights the need for a careful examination of these groups as a basis for suitably targeted and tailored AESPs.

- **Individuals** with labour market difficulties frequently **face multiple overlapping employment barriers**. Policy approaches that address only some of these obstacles may then not be enough to facilitate returns to employment as long as other barriers remain.

The rest of the paper is structured as follows. Section 2 identifies the sample of interest: individuals with potential labour-market difficulties. Section 3 proposes empirically feasible methods for deriving indicators of different labour market barriers using the information available in EU-SILC data. Section 4 presents the clustering exercise and discusses results. Section 5 outlines preliminary conclusions and considers possible next steps. A technical annex sets out the conceptual background and the statistical properties of the latent-class method used for the clustering exercise.

2. JOBLESSNESS AND UNDEREMPLOYMENT: CAPTURING A BROAD SET OF POTENTIAL LABOUR-MARKET DIFFICULTIES

Individuals with labour market difficulties frequently move between non-employment and different states of "precarious" employment. As a result, limiting attention to only unemployed or other non-employed individuals may not capture the true extent of labour market difficulties or the need for policy intervention. In line with the typical scope of activation and employment support policies (AESPs), the target population of the analysis in this paper therefore includes working-age individuals who are entirely out of work (either actively searching for a job or inactive) or whose labour-market attachment is "weak" (Figure 1). "Weak" labour-market attachment can include individuals with unstable jobs working only sporadically, those working persistently with restricted working hours, and those with very low earnings (due to, for example, being partially unpaid, or working informally). The resulting **target population** is a sub-set of the reference population of working-age adults relevant for AESPs. This **reference population**, in turn, is defined as working-age individuals but excluding groups that are normally outside the scope of AESPs: full-time students and individuals in compulsory military service.

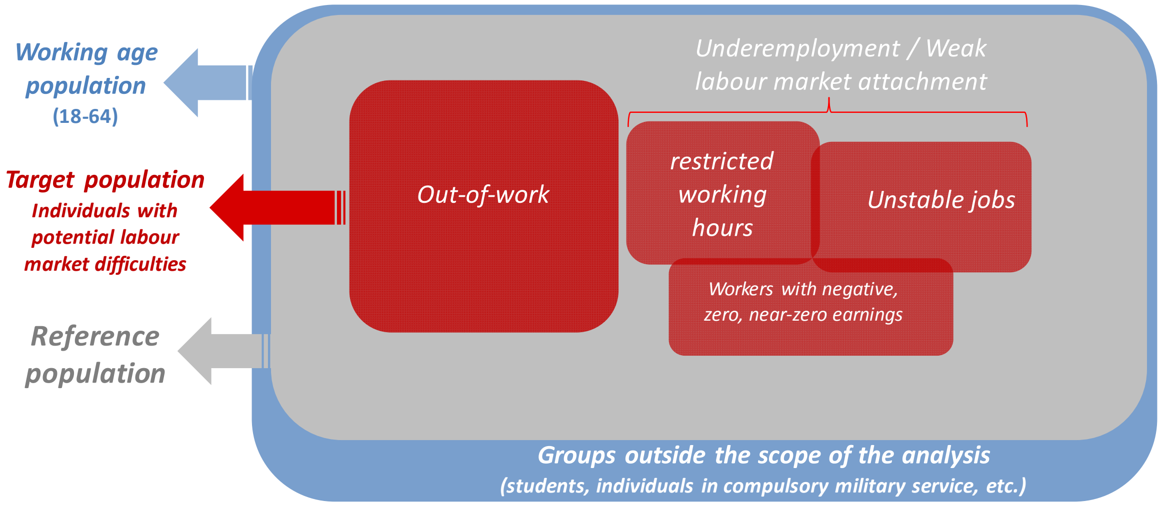
We do not attempt to distinguish between voluntary and involuntary joblessness or reduced work intensity. Individuals can of course choose to be out of work, or in part-time or part-year employment, voluntarily, and some surveys ask respondents whether they "want to work". However, those saying they do not want employment, or prefer to work part-time or part-year, may do so as a result of employment barriers they face, such as care obligations or weak financial incentives, and which policy could address. If extended voluntary labour-market inactivity or underemployment creates or exacerbate certain types of employment barriers, it may subsequently give rise to involuntary labour-market detachment or partial employment in later periods.[2] For these reasons, this paper is descriptive in its approach and takes no position on whether policy intervention is justified for specific groups. It identifies empirical combinations of employment

2. There are further measurement issues, for instance employment preferences may follow own work habits or those of an individual's reference group.

barriers for a broad group of individuals with *potential* labour-market difficulties.[3] Based on the results, policymakers can decide which groups should and should not be targeted by AESPs.

Table 1 shows distributions of main activity status in the reference population for seven selected EU countries with very different labour-market situations and institutions. Information on activity status is derived using longitudinal information provided by EU-SILC data. Purpose, definitions and data sources differ from other common tabulations, such as those based on labour force survey "snapshots" relating to a specific point in time. In particular, cross-sectional EU-SILC data contain information on activity status for multiple points in time during a **reference period** of at least 12 months. This is useful for the purpose of the present paper, as it enables us to capture employment difficulties over an extended period. The categories shown in Figure 1 utilise all available information: typically 12 consecutive monthly observations corresponding to the calendar year (January-December of year T-1) plus one additional observation at the moment of the interview (in year T). Box 1 explains the procedure for identifying *main activity status* in more detail.

Figure 1. Individuals with potential labour market difficulties



3 . Although this information is not considered when defining the target population in this paper, data on involuntary part-time employment is used as an input for deriving an indicator of limited employment opportunities (labour demand) in Section 3.3.

Table 1. Working-age and reference populations by main activity status

In number of observations and as a share of the working-age (ages 18 to 64) sample, EU-SILC 2013

Sample size

	ESP	EST	HUN	IRL	ITA	LTU	PRT	All countries
Working-age population, *of those*	19,722	9,245	16,695	6,988	27,399	7,128	9,772	96,949
Reference population	18,070	8,334	15,260	6,468	25,338	6,520	9,131	89,121

Share of working-age population

	ESP	EST	HUN	IRL (1)	ITA	LTU	PRT	All countries (2)
Working-age population, *of those*	100	100	100	100	100	100	100	100
Reference population, *of those*	**93**	**92**	**91**	**93**	**92**	**91**	**93**	**92**
FT employee	40	58	49	36	39	56	51	41
PT employee	8	5	2	13	7	3	3	7
Self-employed	9	6	6	7	13	6	7	10
Unemployed	22	7	10	14	11	10	16	15
Retired	3	4	11	3	6	6	7	5
Unfit to work	3	6	5	5	2	6	2	3
Care duties	8	6	3	12	13	2	5	10
Other inactive	1	1	4	2	2	2	2	2
Full-time students	7	8	9	9	8	9	7	8
Compulsory military service	0	0	0	0	0	0	0	0

Notes: Shares computed with survey weights. “Working-age”: 16-64 age range. Records with jointly-missing information on activity status, educational achievements and health conditions are excluded from the sample. To account for miscoded categories for some less frequent activity status in Ireland, a correction based on information at the moment of the interview was applied. “All countries”: weighted average of individual country shares.

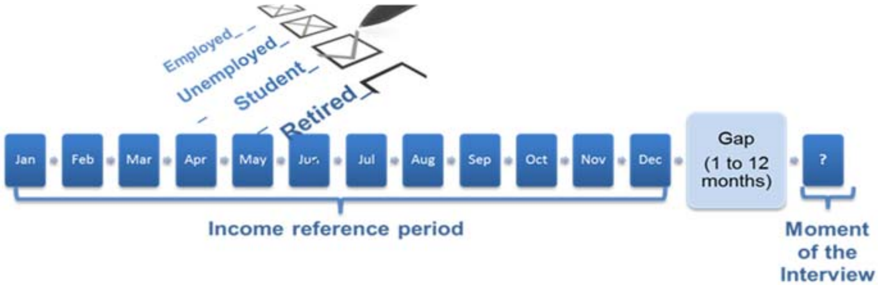
Source: Authors’ calculations based on EU-SILC 2013.

The data presented in Table 1 refer to a year with particularly difficult labour-market situations and very high unemployment in crisis-hit countries. On average, the just under half of individuals are classified as employees, mainly working full time. Ireland has a relatively high share of part-time employees (13%, versus about 5% for all countries), whereas Italy has a large share of self-employed (13% versus 11% for all countries). In the aftermath of the financial and economic crisis, Spain shows a very large share of self-reported unemployment during the reference period (22%). Italy and Ireland have the highest shares of individuals who say they do not work because of care responsibilities (13% and 12%, respectively). The share of non-working individuals who are “permanently disabled or unfit to work” is relatively high in Estonia and Ireland (between 5% and 6%), while the incidence of retired individuals is especially high in Hungary (11%). Full-time students represent, on average, about the 8% of the population of working age. The share of individuals in compulsory military service is very small (less than 0.5% in all countries). The total sample size ranges from around 7 000 in Ireland to more than 27 000 in Italy.

Box 1. Main activity status: Exploiting all information available in EU-SILC

The data collection of the main labour-market status in the EU-SILC questionnaire consists of 13 identical questions. Twelve of them refer to the self-assessed status in each month of the income reference period (the calendar year before the interview) and an additional question refers to the moment of the interview (Figure).

EU-SILC information about the self-reported main activity status



The EU-SILC guidelines define the main individual activity status as the prevailing status during the 12 months of the income reference period (variable *px050*), i.e. the status reported in more than 50% of the 12 activity observations for each individual. Similarly, the main activity status as used in this paper is defined as the status reported in more than 50% of the activity observations, but including also the activity status declared at the moment of the interview. Considering the 13th data point has the advantage of enriching the information on the main occupation during the whole observational period, limiting the number of ambiguous cases (e.g., due to an equal number of different activity status during the year) that would be otherwise coded as missing.

The information content of the 13th data point is greatest when the gap between the last month of the income reference year and the moment of the interview is relatively short. In this paper the additional data point is used only when the interviews take place in the first two quarters.

The table below shows that in five of the seven countries nearly all the interviews take place during the first half of the year. Interviews in Ireland and Italy were mostly conducted during the second half.

Table - Timing of survey interviews in EU-SILC

% of total country sample

	ESP	EST	HUN	IRL	ITA	LTU	PRT
First quarter	11	50	0	22	0	19	0
Second quarter	**80**	**50**	**100**	**25**	**25**	**81**	**100**
Third quarter	8	0	0	22	33	0	0
Fourth quarter	0	0	0	32	42	0	0
Total	100	100	100	100	100	100	100

Source: Author calculations based on EU-SILC 2013

2.1. Out-of-work individuals

In addition to self-reported activity status, EU-SILC also includes detailed income information. Possible approaches for identifying the out-of-work population therefore include the following:

- Individuals whose main self-reported status is "out of work" in every month of the reference period.
- Individuals with no labour incomes during the reference period.

For a number of reasons, the two criteria do not necessarily overlap entirely. Some may consider their main activity status "not at work", even if they have received some labour incomes during the period (they may, for example, devote most of their time to care duties). Similarly, someone without labour incomes during the income reference period could consider herself as mainly working (e.g. a self-employed individual).

Table 2 shows for the seven countries under analysis the contingency table for ‘positive earnings’ and ‘market activity’ during the reference period. As expected, the overlap is large, but not complete. On average, about 4% of the reference population in the seven countries report positive earnings but no work activity. In countries where this phenomenon is significant (mainly in Italy and Spain), the individuals concerned are mostly report unemployed or retired as their main activity status (they may have worked a small number of days/weeks in some months, some of them could have received delayed wage payments after their job ended). 2% of the reference population reporting no market incomes but a positive work activity; most of them are self-employed.[4]

Table 2. Out of work: correspondence between information on work activity and labour income

% of reference population, total for seven countries
(ESP, EST, HUN, IRL, ITA, LTU, PRT)

	any work activity	no work activity	Total
any (positive) labour income	**57**	4	61
no labour income	2	**37**	39
Total	69	31	100

Source: Authors’ calculations based on EU-SILC 2013.

In the rest of the paper, the “out-of-work” population is defined as individuals reporting no employment activity. Individuals declaring some employment activity but no labour incomes are considered to have “weak” labour-market attachment of one type or another, as discussed below. Table 3 shows the distribution of main activity status in the resulting out-of-work population, and the overall size of this group relative to the reference population. Implied activity rates range between 66% of the reference population in Ireland and 81% in Estonia.

The composition of the out-of-work population can be interpreted as a first indication of the potential role for different AESPs, and results suggest that these vary substantially across countries: In Lithuania, Portugal and Spain, unemployed jobseekers were the biggest category of individuals without any work activity during the (post-crisis) reference period. In Hungary, the number of (early) retired is almost twice the number of unemployed, whereas domestic tasks (Ireland, Italy) and health-related inactivity (Estonia) are the most sizeable categories in the remaining countries.

Table 3. Out of work: incidence and group composition

	ESP	EST	HUN	IRL	ITA	LTU	PRT	All
out-of-work incidence, % of reference population	30	19	30	34	32	23	30	31
out-of-work composition, % of out-of-work								
Unemployed	52	23	22	36	27	32	45	37
Retired	9	23	39	7	19	25	24	17
Unfit to work	10	32	18	16	5	28	7	9
Domestic tasks	26	21	11	37	44	10	18	33
Other inactive	3	1	10	4	5	5	5	5
Total	100	100	100	100	100	100	100	100

Source: Authors’ calculations based on EU-SILC2013.

4 . Individuals with positive market income but no work activity account for about 6% of the reference population in Italy and Spain – they mainly report being unemployed (38% in Italy, 68% in Spain), retired (23% in Italy, 15% in Spain) or performing domestic tasks (26% in Italy, 9% in Spain).

2.2. Individuals in unstable employment

A useful reference for identifying individuals with unstable jobs is Eurostat's methodology for deriving the Europe 2020 indicator "household work intensity". This indicator measures the number of full-time equivalent months that working-age household members worked during the income reference year, as a proportion of the total number of months that household members could potentially have worked. The indicator adopted in this paper follows the Eurostat methodology but it differs from the Europe 2020 indicator in two respects:

- In line with the reference population in this paper, it refers to months worked at the *individual* rather than the household level;
- It is calculated for the reference population (i.e., ages 18-64 excluding full-time students and individuals in military service), rather than the 18-60 group in the case of the Eurostat indicator.

The threshold to identify individuals with "unstable jobs" is equivalent to Eurostat's low-work-intensity measure: Above zero but no more than 45% of potential working time in the income reference year.[5] To reconcile information reported for the income reference period and at the moment of the interview (see Box 1) the following individuals are also considered in this group:

- Workers who report no employment or self-employment activity during the income reference period but who report being employed or self-employed at the moment of the interview;
- Workers with between 45% and 50% of work activity during the income reference period who do not report any work activity in either the last month of the income reference period *or* at the moment of the interview.

Table 4 shows the resulting shares of individuals with unstable jobs in the reference population. Spain and Ireland have the highest shares (10.7% and 9.3% respectively); shares in Italy and Portugal are lowest (5.0%).

Table 4. Unstable jobs: incidence across countries
% of reference population

ESP	EST	HUN	IRL	ITA	LTU	PRT	All
10.7	7.8	6.9	9.3	5.0	6.3	5.0	7.3

Note: The indicator is based on the self-reported calendar activities during the income reference period and takes into account a correction for non-response and the difference between full-time and part-time work activities. Individuals in this group report one of the following: a) work intensity in the interval 0.01-0.45 during the income reference period and no work activity at the moment of the interview; b) work intensity equal to 0 during the income reference period but a positive work activity at the moment of the interview; c) work intensity equal to 0.5 during the income reference period and any work activity in either the last month of the income reference period or at the moment of the interview.

Source: Authors' calculation based on EU-SILC 2013.

2.3. Individuals with restricted working hours

We characterize as workers with restricted working hours individuals who spent most or all of the reference period working *20 hours or less* a week for one of the following reasons: *illness or disability*, *housework or care duties*, *absence of other job opportunities, voluntary part-time, other reasons*. Individuals working 20 or less hours due to undergoing education or training or because the number of working hours is

5. The Eurostat thresholds are [0; 0.2[for "very low" and [0.2; 0.45[for "low" work intensity.

considered already a full-time job are excluded as they are unlikely to have unused work capacity. The 20-hours threshold is approximately in-line with the 45% threshold that identifies the group with unstable jobs.[6]

An important limitation is that EU-SILC collects working-hours information only for the current job *at the moment of the interview*. The main activity status reported in each month of the income reference period distinguishes between full-time and part-time activities but without imposing an explicit threshold to distinguish between the two.[7] Considering these limitations, we include individuals in the "restricted working hours" category only if they are working 20 hours or less a week at the moment of the interview *and* if the spent at least 6 months of the income reference period working in part-time activities.[8]

Table 5 shows that individuals with restricted working hours represent a relevant share of the reference population in Ireland (8%) and to a lesser extent in Spain (3.3%) and Italy (2.5%). The most frequent reason for working with restricted working hours is what can be regarded as *involuntary* part-time work (i.e., individuals would prefer a full-time work but feel they are constrained by labour-demand factors). A range of other reasons, such as health problems or care responsibilities, do not easily correspond to a "voluntary" versus "involuntary" dichotomy, as they can relate to either preferences or labour-demand constraints. Care or other domestic responsibilities are quantitatively important in Italy, and health issues are an important driver of restricted working hours in Estonia and especially in Hungary.[9]

Table 5 Restricted working hours: Incidence and reasons

	ESP	EST	HUN	IRL	ITA	LTU	PRT	All
part-time work incidence, % of reference population	3.3	1.7	1.1	8.2	2.5	1.9	1.9	2.8
part-time work composition, % of part-timers								
Illness /disability	3	18	33	3	4	25	9	5
Absence of other job opportunities (involuntary)	67	41	5	48	50	42	73	56
Do not want to work more hours (voluntary)	4	26	0	14	13	11	3	9
Housework or care duties	13	16	20	19	26	7	6	18
Other reasons	14	0	43	15	8	16	8	12
Total	100	100	100	100	100	100	100	100

Note: Labels are adapted based on EU-SILC interview guidelines. See text for details on definitions and limitations.

Source: Authors' calculations based on EU-SILC 2013.

2.3. Others with near-zero or negative labour incomes (residual category)

Identifying joblessness or weak labour-market attachment on the basis of self-reported activity status can be subject to measurement/classification errors. As a result, some categories of individuals with potential

6. For a 40-hours working week in a full-time job, 45% of full-time would correspond to 18 hours a week. However, in EU-SILC, the distribution of working hours in the main job shows a high degree of bunching at 10, 15, 20 and 25 hours a week. For this reason, we chose to round to 20, the closest multiple of 5.

7. According to EU-SILC guidelines, full-time / part-time status is self-assessed, as applying a unified threshold is not useful in view of variations in working hours between Member States, sector, collective agreement, etc.

8. The implicit assumption is that the reason for working 20 hours or less applies also to the previous months as long as individuals have not changed their main activity status between the moment of the interview and the income reference period.

9. In Hungary the share of individuals declaring "other reasons" as the main reason for working with restricted working hours is unusually high (43%) while the share of "involuntary" part-timers is unusually low (5%). This may depend on how the question was structured in the Hungarian questionnaire or on miscoding errors in the current EU-SILC release.

labour-market difficulties may not be captured in categories described above. For instance, individuals declaring zero or near-zero earnings may define themselves as full-time workers for most of the year. In addition to possible classification error, these situations could signal potential labour market difficulties, such as underpayment and/or informal activities. In view of the considerable effort that went into ensuring good-quality income information in EU-SILC, individuals reporting some work activity but negative, zero or near-zero earnings over the same period are included in the target population.[10]

Table 6 shows the size of groups with negative, zero or near-zero earnings in the reference population. Portugal shows the highest share (3.4%), followed by Spain (2.0%). Additional tabulations (not reported) show that, when not already included in other groups of the target population, individuals in this group are largely from one of the following groups: full-time self-employed reporting negative or zero earnings, as well as full-time employees with extremely low earnings despite being employed during most of the year.[11]

Table 6 Workers with near-zero or negative labour incomes

% of reference population

	ESP	EST	HUN	IRL	ITA	LTU	PRT	All
negative, zero or near zero income workers, *of those*	**2.0**	**1.1**	**0.6**	**2.0**	**0.4**	**3.3**	**3.4**	**1.3**
negative earnings	0.6	0.1	0.0	0.0	0.1	0.0	0.0	0.0
zero earnings	0.5	0.7	0.2	1.6	0.0	2.1	3.0	3.0
positive but near-zero earnings	0.9	0.4	0.4	0.4	0.3	1.2	0.4	0.4

Note: Labour income is sum of the gross employee cash or near-cash earnings and gross cash income from self-employment. The median annual loss for workers reporting negative earnings is: €4 244 in Spain, €8 851 in Estonia, €35 in Hungary, €1,450 in Italy. The median annual loss for workers reporting negative earnings is: €4 244 in Spain, €8 851 in Estonia, €34 in Hungary and €1 450 in Italy. For the group with near-zero monthly earnings, annual earnings are divided by the number of months spent in paid work during the income reference year. The income thresholds of €120/month in PPP12 are: €116 for Spain, €93 for Estonia, €145 for Ireland, €126 for Italy, €101 for Portugal, €79 for Lithuania (reference = USD) and €67 for Hungary (reference = EU28).

Source: Authors' calculations based in EU-SILC2013.

2.4. Wrapping up: Incidence of different types of potential labour market difficulties

Figure 2 illustrates the sizes of different categories of individuals with potential labour market difficulties, and the extent of overlap between them. As explained at the beginning of Section 2, the label "potential labour-market difficulties", highlights the fact that not all individuals without a job or with weak labour-market attachment will consider themselves in difficulties. Across the seven countries, out-of-work individuals accounted for just under a third of the reference population and represent the majority of the target population for the remainder of this paper. But the categories of individuals with some form of low work intensity, considered together, are sizeable as well, summing to about one third of the out-of-work group (some 10% of the reference population). Those with negative, zero or near-zero earnings are a relatively small group that overlaps with the two other categories of low work intensity. Restricted hours and unstable employment frequently also occur in combination.

Figure 3 reports group sizes per country. Given the overlaps, any one-dimensional grouping depends on how one ranks the overlapping sub-groups in creating each of the categories. The results in Figure 3 are

10. For simplicity, we adopt a common (and arbitrary) low-earnings threshold for all countries: EUR 120 / month in purchasing power parities, which corresponds to approximately the 5th earnings percentile across the selected countries. The 120 EUR in PPP translates into monthly cash values of: EUR 116 in Spain, 93 in Estonia, 145 in Ireland, 126 in Italy, 101 in Portugal, 79 in Lithuania and 67 in Hungary. All values are well below applicable statutory minimum wages in countries where these exist. Other thresholds are possible (such as a given fraction of the minimum wage) but tests performed by the authors showed that these do not produce significantly different results in practice.

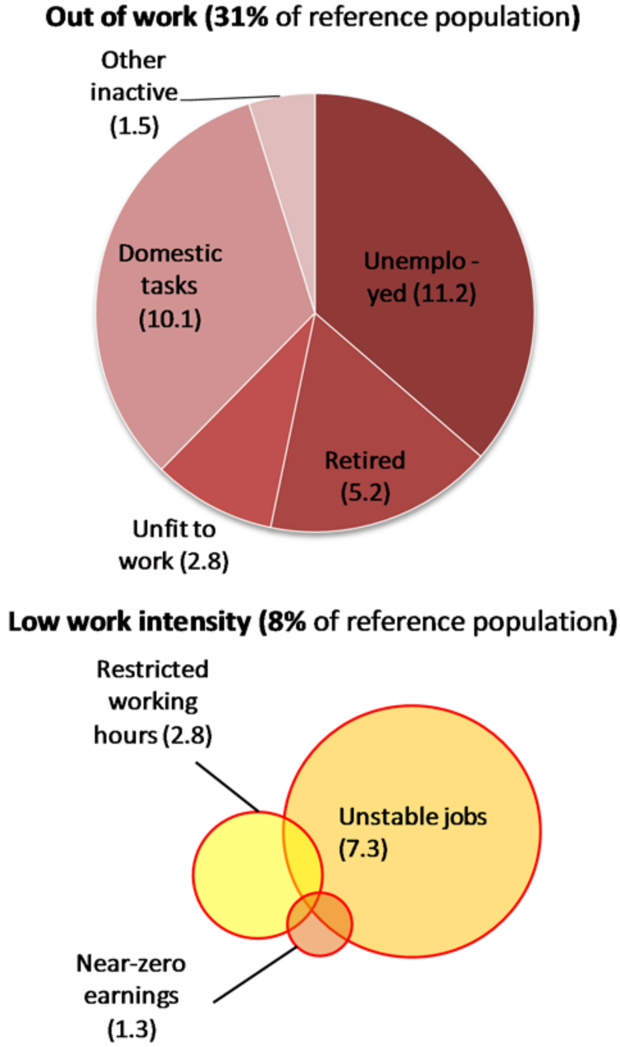
11. The highest share of self-employed reporting near zero income (in Portugal) was around 80%.

based on the following ordering of individuals falling into more than one category: (1) out-of-work, (2) unstable jobs, (3) restricted working hours, (4) negative, zero and near-zero earnings.

The complement of these groups corresponds to workers with what might be loosely termed a 'good job' in terms of the characteristics considered here: employees and self-employed working mostly full-time and with significant earnings during most of the reference period. According to this measure, Estonia performs best across the selected countries, with the largest proportion of individuals without major labour-market difficulties (71%). Estonia also has the lowest share of out-of-work population (19%) contrasting with the highest shares in Ireland (34%) and Italy (32%). Ireland and Spain have the highest proportions of groups in partial employment: Spain has the highest share of low-work intensity workers (11%), whereas Ireland has a relatively high share of individuals on restricted working hours (7%).

Figure 2. Potential labour-market difficulties: Incidence and overlaps

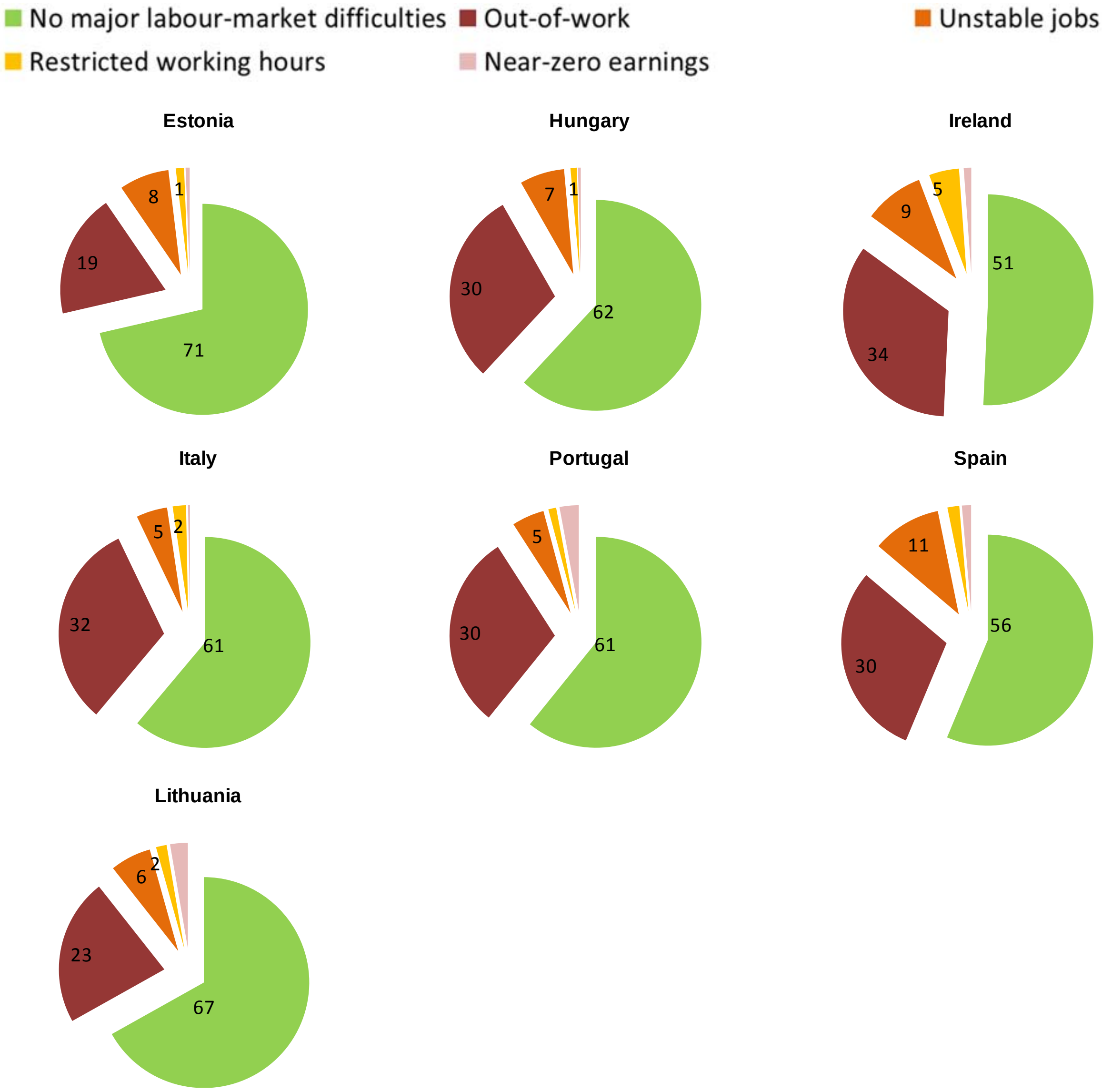
% of reference population, total for seven countries
(ESP, EST, HUN, IRL, ITA, LTU, PRT)



Source: Authors' calculations based on EU-SILC 2013.

Figure 3. Individuals with potential labour market difficulties by country

% of reference population



Source: Authors' calculations based on EU-SILC 2013.

3. INDICATORS OF EMPLOYMENT BARRIERS: CONCEPTS AND MEASUREMENT

Working age individuals with no or weak labour-market attachment may face a number of employment barriers that prevent them from fully engaging in labour market activities. A thorough understanding of these barriers is a pre-requisite for designing and implementing policy interventions in a way that is well-targeted and suitably adapted to the circumstances of different policy clients. To be as effective as possible, activation and employment-support measures should closely correspond to the specific drivers behind people's labour-market difficulties.

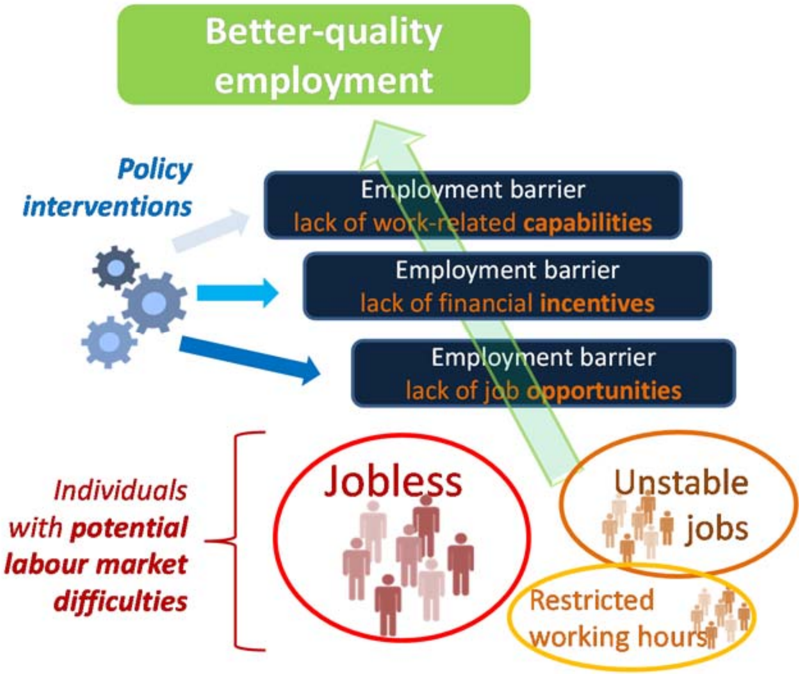
As a first step in operationalising the concept of employment barriers for empirical work, this paper adopts the following three categories of barriers, as proposed by Immervoll and Scarpetta (2012) and used in OECD (2015b) and illustrated in Figure 4 below:

- Insufficient **work-related *capabilities*** comprise a broad range of different factors that may limit individuals' capacity for performing specific tasks. Examples are a lack of education, skills or work experience, care responsibilities, or health-related limitations.
- Weak ***incentives* to look for or accept a 'good' job**, e.g., because of low potential pay, relatively generous out-of-work benefits, or high standards independently of own work effort; and
- Scarce **employment *opportunities***, e.g., a small number of vacancies in the relevant labour-market segment, friction in the labour market due to information asymmetries, skills mismatch, or discrimination in the workplace.

Barriers in any of these categories can result in no or weak labour-market attachment, but each barrier generally calls for different policy approaches. For instance, weak job-search incentives, poor access to job offers (e.g. lack of information about vacancies, lack of capacity to effectively apply for a vacancy etc.), or obsolete skills, may be amenable to being tackled through activation, job-search assistance and training measures. Other barriers, such as depressed labour demand in a given region, health limitations or care responsibilities, require different approaches and may also signal a need for structural reforms that go beyond the scope of AESPs (e.g., reducing non-wage labour costs, encouraging firms to adapt work environments or introduce work flexibility, or social policy reforms that strengthening institutions for child and elderly care or make health services more accessible). If multiple barriers exist simultaneously, successful interventions are likely to require an appropriate combination, coordination and sequencing of policy measures.

The remainder of this section uses EU-SILC data to derive a set of empirical indicators in each category of employment barrier outlined above. These should be understood as illustrations rather than an attempt at a comprehensive list of all indicators that are useful or can be derived using this or other available data sources. In each category, the proposed indicators are a basis for discussion that should be refined further for in-depth country-specific analyses. (For instance, richer indicators of people's work-related skills might be derived using information on people's present or past occupation, and information on people's job-search activity could be used to identify additional employment barriers in the "incentives" category.) Throughout the section, we refer to data limitations, their implications, and possible ways to address them in the context of this paper. Methodological choices that cut across several or all of the proposed indicators are discussed in Box 2.

Figure 4. Employment Barriers – Conceptual framework



Source: adapted from Immervoll and Scarpetta (2012).

Box 2. Selecting indicators and thresholds

Employment-barrier indicators can have various purposes. In this paper, a primary consideration is whether they provide a suitable input for identifying target groups for AESPs. Binary indicators have a number of advantages for the statistical clustering approach (Latent Class Analysis, LCA) as employed in Section 4. They greatly simplify the statistical model for the LCA. A dichotomy of "barrier" versus "no barrier" also facilitates the interpretation of the resulting groupings.

The construction of binary (or any ordinal) indicator involves establishing thresholds for the relevant categories. As the choice of thresholds is essentially arbitrary, it is desirable to make it as transparent and consistent across indicators as possible. In this paper, continuous variables are generally discretized into binary "barrier" versus ""no barrier" indicators using thresholds that are defined as fixed proportions of the median in the reference population. The approach is, for instance, equivalent to the Eurostat and the OECD approaches for segmenting populations into "income poor" and "not income poor".

In some cases, depending on the underlying distribution of the continuous variable, the discretization of the indicator based on the fixed proportions from the median may not work in practice, as it can produce a very low (or very high) share of individuals facing the barrier. This can happen when the underlying continuous variable is bounded into a small interval, e.g. a probability or a share in the interval 0-1, or when the variable is highly concentrated around specific values (e.g. the mean). In these cases other ad-hoc thresholds are discussed in the respective indicator section.

3.1. Insufficient work-related capabilities

Individuals who would like to work may be unable to provide the type or quantity of labour that is demanded by employers. The resulting mismatch reduces their chances of finding a job, and their productivity while in employment. This section considers the following types of potential capability barriers: skills and education, work experience, health limitations and care responsibilities.

Skills and education

The role of low skills and low education as drivers of poor employment outcomes has been extensively documented. The type of skills acquired, and proficiency in these skills, affect both the probability of finding a job and levels of pay when in work (OECD, 2014b). Skilled workers outperform low-skilled peers in terms of wages, employment stability and upward mobility in income (OECD, 2015b; OECD, forthcoming). Individuals with inadequate skill levels face higher risks of labour-market marginalisation and longer unemployment spells, and they are more likely to depend on social benefits as a main source of income (OECD, 2012a).

Skills encompass a wide range of dimensions. Education attainment – certified skills acquired in initial education – is one of them. Although a great deal of skill acquisition happens on the job (along with some skill obsolescence), educational attainment remains strongly linked with productivity and labour market outcomes (OECD 2014b). Adults with higher education levels are more likely to develop better general, numerical and problem-solving skills that translate into better labour-market outcomes.

An ideal indicator of skills-related employment barriers would capture both work-related skills and educational attainment. However, while this information is available in specialised surveys, notably the OECD Adult Skills Survey (PIAAC), information in EU-SILC is limited to the highest attained level of education and the type occupation (according to ISCO standards).

The highest educational attainment constitutes the preferred "skills" indicator in the context of this paper. We classify individuals who have achieved less than upper secondary education (according to ISCED 2011 standards) as having low skills, and those with complete upper secondary and above as having medium to high skills. Individuals in the low-skills category are hence considered as facing a skills-related employment barrier.[12]

Table 7 shows population shares for "low skills" in the target population. They range from only 19 percent in Estonia, to 76 percent in Portugal. This very wide range highlights the importance of country context when interpreting results, and the potential adaptation of indicator content that may be needed for country-specific analyses: Individuals with upper secondary education may be considered "medium skilled" in Portugal and Spain but can be expected to be firmly in the "low skilled" category in Estonia.

Table 7. **Share of individuals facing skills-related employment barriers**

% of target population

	ESP	EST	HUN	IRL	ITA	LTU	PRT	All
Education: "low"	61	19	31	37	55	21	76	56

Source: Authors' calculations based on EU-SILC 2013.

Work experience

Work experience constitutes both human and social capital (Becker, 1993; and Lin, 2001), enhances and maintains work-related skills, and plays an important role in explaining different labour market outcomes among individuals with the same educational attainment (OECD 2014b). Both technical and social (or 'soft')

12 . Occupation could be a useful additional proxy of skills level. EU-SILC contains occupation information for all individuals with current or previous work experience. Following the ILO guidelines (ILO, 2010) the ten ISCO-08 major groups of occupations could be organized into four skills levels ranging from elementary occupations ("low" skills) to managers and senior officials ("high" skills). Exploring this option further is left to future work.

skills, such as ability to work with others or to meet deadlines, are typically developed and enhanced on the job. For employers, work experience can serve as a valuable signal for unobservable skills or traits. With more work experience, individuals typically also extend their work-related social networks, which can be instrumental in maintaining employment, achieving career progression and securing re-employment after job loss (Marsden and Gorman, 2001; Fernandez et al., 2000; Mouw, 2003; Mc Donald, 2011; Contini, 2010).

Effects of work experience on employment outcomes can be expected to be cumulative to some extent. But due to a depreciation of skills and social capital, recent work experience will generally be a more important driver than experience acquired in the more distant past. A reasonable requirement for an indicator of experience-related employment barriers is therefore that it should capture both total and recent work experience. Multiple or lengthy career interruptions may also erode the value of total experience. An ability to account for career breaks might therefore be a second desirable property of a work-experience indicator, in order to distinguish between individuals with the same total experience but different patterns of gaps in their careers.

The information available in EU SILC raises several challenges in this context. The survey provides data on: (1) the number of years spent in paid work, (2) the year when the highest level of education was attained and (3) the activity status during each month over the survey reference period. None of this information in isolation fully captures the potential employment barriers outlined above. For instance, the number of years spent in paid work does not provide the desirable distinction between more and less recent career breaks. Graduation year lacks information about any type of career breaks, whereas the monthly activity status provides information about recent work experience the reference period but possibly misses very recent career breaks (between the end of the reference year and the time of the interview).

Accounting for all relevant in one single indicator is challenging, and we therefore propose two distinct indicators as follows. The first seeks to captures the overall stock of work experience relative to the potential work experience since graduation, while a second indicator relates to work experience accumulated more recently:

- *Relative total* work experience is given by the *ratio* of total reported work experience and the *potential* total work experience that would have been achieved in absence of any career break since graduation. Since potential total work experience is not observable precisely, we proxy it as the difference between age and the typical graduation age for his or her highest completed education level. The indicator can take one of *three* values. It is set to 1 for individuals who have never worked; to 2 for those with some positive actual work experience and a ratio between actual and potential work experience *below 60%,* and to 3 for the remaining of the individuals.

- *Recent* work experience is constructed from the number of months spent in work during the reference period (the past 12-18 months, depending on the time of interview, see Box 1). The indicator identifies the individuals with some weak labour market attachment as defined in Section 2, i.e. individuals with unstable jobs, restricted working hours or very-low earnings. By definition, these individuals have some work experience during the reference period, while the out-of-work population has no work experience during the reference period.

Table 8 shows the cross tabulation of the two indicators. As expected, “recent” work experience and “relative” work experience are correlated. However, they do not capture the same situations. Calculated over the target population of all seven countries, the share of individuals with low relative work experience who have some recent work experience is about 9%, whereas those with no recent work experience but who have worked in the past are quite evenly distributed across the different categories of relative work experience:

24% of individuals with no recent work experience have low total work experience whereas 31% are in the "medium" to "high" category.

Table 8. Work experience indicators are correlated but capture different situations

% of target population, total for seven countries
(ESP, EST, HUN, IRL, ITA, LTU, PRT)

	Relative work experience			
	Have never worked	"Low"	"Medium" to "high"	*Total*
Recent work experience	0	9	16	*25*
No recent work experience	20	24	31	*75*
Total	*20*	*32*	*47*	*100*

Source: Authors' calculations based on EU-SILC 2013.

Table 9 shows breakdowns by country. Italy has the highest share of individuals that do not have any work experience at all (30%) while Estonia, Hungary and Lithuania show the lowest shares (9-10%).

Table 9. Work experience indicators: incidence in the target population

	ESP	EST	HUN	IRL	ITA	LTU	PRT	All
Relative work experience								
None (has never worked)	14	9	9	13	30	10	11	20
"Low"	37	32	…	32	29	…	23	32
No recent work experience	68	67	78	70	82	68	77	75

Note: Data for Hungary and Lithuania do not provide information on the number of years spent in paid work. For these two countries an alternative indicator could, instead, be based on a binary variable capturing whether the individual has ever worked.

Source: Authors' calculations based on EU-SILC 2013.

Health limitations

Self-reported sickness or disability affect large parts of the working population in EU and OECD countries. There are different views on the extent to which an observed rise in reported health problems (notably mental health issues) reflects an objective increase, or whether it is due to greater awareness or better diagnostics. Irrespective of the drivers behind observed trends, employment among individuals with a disability is relatively low in many EU and OECD countries (e.g., just over 40% in the average OECD country, compared with 75% for people without disability; OECD, 2010). In a large number of countries, including those with relatively low unemployment rates, disability benefits are at least as common a form of out-of-work support as unemployment benefits (OECD SOCR database). In quantitative terms, the situation of people suffering (partial) disabilities is a particularly important driver of overall labour-market performance and living conditions.

The relationship between health and work is complex, involving a two-way causal link, with effects running not only from health to work, but also from work to health. Various dimensions of work such as employment and working conditions, including employment status, working hours, job decision latitude, job demand and job strains) have an impact on physical and mental health (Barnay, 2015, Devaux and Sassi, 2015). Poor health and health-related behaviours that increase people's risk of developing chronic diseases may also cause adverse labour market outcomes. Chronic diseases and lifestyle risk factors have an impact on labour market engagement in terms of employment opportunities, wages, productivity, sick leave, early retirement and receipt of disability benefits (Devaux and Sassi, 2015).

An ideal indicator of health-related employment barriers should describe individuals' physical and mental abilities to, and capacity for, work. EU-SILC contains three directly relevant variables: self-perceived health status; chronic health conditions (e.g., relating to a long-standing illness); and activity limitations due to physical and mental health conditions (self-perceived long-standing limitations in usual activities due to health issues). These variables are likely to address quite different aspects of poor health and an analysis of the relationship between health and labour market outcomes may therefore be sensitive to the measure that is adopted.

In this paper, health-related employment barriers are operationalised using the EU-SILC variable on limitations in usual activities due to long-lasting physical or mental health conditions.[13] Specifically, following Knudsen *et al.* (2010), individuals who report *some* or *severe* limitations in usual activities are characterized as having a reduced work capacity due to health issues, while individuals reporting being "not limited" in this regard are presumed to face no relevant employment barriers.[14]

Table 10 shows the resulting incidence of health-related employment barriers in the target populations of the seven countries. Large shares of the target populations in Estonia (44%), Hungary (37%) and Lithuania (35%) report some or severe limitations in their daily activities. Health barriers appear to be less frequent in the three Southern European countries and in Ireland.

Table 10. Health-related employment barrier: incidence in the target population

	ESP	EST	HUN	IRL	ITA	LTU	PRT	All
Health: Some or "high" limitations	25	44	37	23	26	35	29	27

Source: Authors' calculations based on EU-SILC 2013.

Care responsibilities

Care responsibilities can be primary drivers of individuals' inability to participate in the labour market, particularly among women. Unpaid work, including childcare or care for incapacitated family members, is time consuming and reduces the amount of time that can be spent in paid work (OECD, 2011). High-intensity care-giving, in particular, is associated with low labour supply among family carers (OECD, 2012b).

An informative indicator of care-related employment barriers can be constructed using EU-SILC data on *i)* the family members who face some *unmet care need*, such as young children, incapacitated family members or elderly relatives, and *ii)* the availability of alternative care arrangements (use of formal care services and availability of *potential care-givers* other than the person whose employment barrier is being evaluated). Combining the two types of information, it is possible to quantify the extent of unmet care needs that each adult household member may be responsible for. A general limitation of such an indicator is that care responsibilities are considered only within a household: Any care responsibilities for family members residing in other households (e.g., an elderly mother living

13 The variable "chronic morbidity" is not suitable for measuring an employment barrier: Following the EU-SILC guidelines an individual with, say, hay fever would be classified in this category.

14 Similar to the skills/education indicator, the most suitable threshold may again differ depending on the country context. For instance, in some countries individuals reporting "some" limitations may not be considered restricted in their work capacity.

alone) are therefore missed. In this sense, the proposed EU-SILC based indicator might usefully be considered a lower bound for the true extend of care-related employment barriers.

Specifically, the steps for deriving an indicator for care responsibilities can be summarised as follows:

1. Identify the family members who potentially require care from other family members.
2. Identify potential care-givers in the household.
3. Identify the household members that are most likely to face care-related employment barriers, by assigning unmet care needs to each potential care giver.

The identification of individuals who could *potentially* require care from other family members distinguishes between children and elderly / incapacitated adults:

- For **children,** the EU-SILC information on the weekly hours of non-parental childcare allows identifying the children who, most likely, face some unmet child care needs. Specifically, following the Eurostat indicator for measuring progresses towards the Barcelona targets, a young child (below 13 years) receiving 30 or less hours of non-parental childcare a week necessitates additional childcare.

- For elderly or incapacitated **adult family members**, EU-SILC data do not provide information on the number of hours of care provided by professional or other informal carers. One of the most relevant variables is the level of limitations in usual activities due to health issues, which, in combination with age and information on the main activity status, can help identify family members who are likely to require care. Specifically, working-age family members are likely to need care if: 1) they report *severe* long-lasting limitations in activities due to health problems and, 2) report a permanent disability as the main reason of inactivity. Similarly, elderly family members are classified as requiring care if condition (1) holds and if they report to be inactive during each month of SILC reference period.

Potential care-givers are individuals with a potentially-significant capacity to provide care to other household members. They are adults aged 18-75 with no severe health-related limitations and observed in one of the following main activities during the SILC reference period: part-time work, unemployment, retirement, domestic responsibilities and other types of inactivity excluding a permanent disability. Individuals reporting full-time activities, i.e., full-time workers, full-time students and individuals in compulsory military service, are expected to have no or little residual capacity to provide care to other household members and are not considered potential care-givers.

Family members with unmet care needs are assumed to represent a significant care-related employment barrier for a potential care-giver if the following conditions apply:

1. There is *only one* potential care-giver in the household.
2. There is more than one potential care-giver, but only one of them reports to be inactive or working part-time *because of care responsibilities*.

The idea behind the identification of care-related employment barriers as outlined above is that the presence of a *plurality* of potential care-givers within the same household automatically reduces the related employment barrier for the other potential care-givers. For instance, if there are two family members who are staying at home to look after a young child, who can share responsibilities, their resulting employment barriers would be less binding. When several potential care-givers are available in the household, we use information on self-reported care responsibilities to discriminate who faces a significant barrier; in this case only individuals who report being actually engaged in full-time care duties face a significant care-related employment barrier.

To combine the two types of care barriers, we construct a dichotomous variable that takes the value 1 if an individual faces *at least one* care-related employment barrier and 0 otherwise. Table 11 shows the share of individuals in the target population facing significant care barriers and the related breakdown into the two sub indicators.

Table 11. Care barriers indicator: incidence and composition

	ESP	EST	HUN	IRL	ITA	LTU	PRT	All
Significant care responsibilities incidence, % of target population	15	21	15	25	16	13	8	15
Significant care responsibilities composition in %								
Care responsibilities, children	83	84	74	88	81	72	54	81
Care responsibilities, incapacitated family	15	14	23	9	18	24	44	18
Both	2	2	3	3	1	4	1	1
Total	100	100	100	100	100	100	100	100

Source: Authors' calculations based on EU-SILC 2013.

3.2. Weak incentives to look for or accept a 'good' job

Weak work incentives can arise when income gain from taking up a job or working more is limited, because net wages are low or because generous out-of-work benefits are withdrawn as people start to work (this driver of employment behaviour is referred to as "substitution effect" in standard micro-economic economic models). In addition, people may decide to limit their work effort because they have access to other income sources and can therefore "afford more leisure" (economic models refer to this as the "income effect"). Even though results are not available for all OECD countries, there exists a relatively broad consensus among labour economists on the responsiveness of people's employment decisions to financial work incentives, such as the net income gain of working one hour more or of working at all. Among the main findings are the following (see, e.g., the overview in Immervoll, 2012 and the references therein):

- The substitution effect is a more powerful driver of employment behaviour on average, but the income effect can be relevant for some groups (e.g., spouses of well-paid principal earners).
- Financial incentives affect overall labour supply mainly through their influence on labour force participation (i.e., the decision on whether or not to work, the so-called "extensive margin" of labour supply decisions), while changes in the number of hours worked (the "intensive margin") are quantitatively less important.
- Low-income groups and lone parents react more strongly to financial incentives; and
- Labour supply is more responsive (or "elastic") for women than for men.
- These results are important when considering the potential economic cost of reforming out-of-work support programs. They also provide essential guidance for targeting make-work-pay policies (e.g., for a given amount spent on in-work benefits, targeting these resources on women and low-income groups, especially when children are present, is likely to create the biggest payoff in terms of stronger employment and higher earnings).

Despite these stylised results emerging from the international evidence, it is notable that there are widespread differences across countries. One important reason for large country differences is that financial incentives may have limited effects on observed employment outcomes if other barriers prevent people from adjusting their labour-market status or working hours. For instance, when involuntary unemployment is high during a downturn, many individuals who want to work cannot find a job, or they cannot work as many hours as they would like. Frictions in the labour market (e.g., due to poorly functioning public employment services) can have similar effects.

Although work incentives are sometimes a principal focus of the activation policy discourse, these considerations point to the importance of analysing possible work-incentives issues alongside other employment barriers.

Work disincentives: Income available independently of own work effort

A proxy for the "income effect" can be derived straightforwardly from the EU-SILC variable 'gross household income' (which includes pre-tax income from labour and capital plus government transfers) *minus* own income that depends on the person's own work efforts (*i.e*., employment income and earnings-replacement benefits, such as unemployment benefits, of the person of interest) and *minus* a share, depending on the number of adults in the household, of social transfers awarded at household level (for instance, social assistance or rent allowances). We derive this indicator in equivalised terms to account for differences in household size. The indicator is then discretised into two categories following the procedure discussed in Box 3.

Box 3. Measuring income from sources other than own employment

A consistent measure for incomes from sources other than own employment can be derived straightforwardly by subtracting all income components that *are* linked to own employment from total gross household income. Incomes that depend on own employment are:

- Earnings: employee cash or near cash income, employer's monetary contribution for the company car and cash benefits from self-employment.

- Own individual earnings-replacement benefits: unemployment, survivor, sickness and disability benefits, education-related allowances, and early-retirement pensions.[1]

- The individual's share of earnings-replacement benefits that are paid at the household level, measured as the sum of these benefits divided by the number of adults in the household.

The specific choice of income components classified as "incomes from sources other than own employment" depend on country-specific policy settings and, in particular, may differ between types of benefits that cannot be distinguished in harmonized EU-SILC data provided by Eurostat. Country-specific versions of SILC or other household surveys may be better suited for identifying relevant benefit categories more precisely.

The resulting measure is divided by the "modified OECD" equivalence scale and discretized into 2 categories using a threshold of 1.6 times the median value in the reference population: Individuals with a value above this threshold are those with "high" incomes from sources other than own employment. The table below shows results by country.

Determinants of financial work disincentives (income effect)

All income measures are shown in equivalised terms

	ESP	EST	HUN	IRL	ITA	LTU	PRT	ALL
"High" income from sources other than own employment, % of target population	23	23	19	21	24	19	20	23
Income available independently of own work effort (€/year)	24503	11776	8284	36874	32169	8508	17692	26586
Income dependent on own work effort (€/year)	8675	4072	3581	13465	13980	2314	7028	10659
Reminder of the target population, %	77	77	81	79	76	81	80	77
Income available independently of own work effort (€/year)	6125	2835	2641	8284	8896	1896	4167	6842
Income dependent on own work effort (€/year)	7038	2848	2545	12681	8117	2299	4229	7053

Note: Results expressed in EUR 2013, income measures correspond to average values.
Source: Authors' calculations based on EU-SILC 2013.

1. We assume that old-age pensions received after the statutory retirement age not dependent on own current employment, i.e., a pensioner can take up employment without losing his or her pension entitlement.

Work disincentives: Limited income gain from own work effort

There are well-established indicators of the financial work disincentives resulting from the operation of the tax-benefit system, i.e., the combined effects of the taxation of in-work incomes and the withdrawal of benefits. Participation Tax Rates (PTR) and Marginal Effective Tax Rates (METR) measure which part of any additional earnings is "taxed away" as individuals take up employment (PTRs) or increase their earnings (METRs).

PTRs and METRs could be computed for specific labour-market and family situations using tax-benefit models. The OECD's tax-benefit models are available for all or most OECD and EU countries. As model-household calculations, they are however not well-suited for determining tax burdens and benefit entitlements for a representative set of households as included in EU-SILC or other household surveys.

Alternatively, the cross-country fiscal microsimulation model EUROMOD, which uses EU-SILC as input data, could be used (Immervoll et al., 1999; Sutherland, 2001; Sutherland and Figari, 2013). However, matching EUROMOD results to the most recent wave of SILC is not always possible. In particular, at the time of writing, the latest EUROMOD input data refer to 2012, whereas this paper employs EU-SILC 2013.

In view of these limitations, an alternative, and simpler, option is to construct an approximate indicator of the PTR based on the extent of benefit reductions that an individual is likely to experience when taking up full-time employment. The focus on PTRs is in line with the empirical finding that the extensive margin of labour supply is quantitatively more important than the intensive margin. The restriction to the effects of benefits also appears reasonable since benefits are typically found to be a much more powerful driver of PTRs than the taxation of in-work earnings, especially for group of interest in this paper: individuals at the margin of the labour market with lower earnings potential.

To proxy the PTR resulting from the withdrawal of benefits as an individual takes up employment, we use the ratio of the amount of earnings-replacement benefits received at the individual level in the numerator, and own potential (shadow) wage. The resulting variable is discretised into a binary "disincentive" indicator that takes value 1 for ratios equal or higher than 60% (i.e., 60% or more of potential in-work earnings are "taxed away" as the individual takes up employment). Box 4 outlines the steps used to construct the indicator.

Box 4. Measuring financial gain from own work effort

Earning-replacement benefits are a main determinant of financial work disincentives at low wage levels. Since they are typically lost upon entering employment, an indicator of financial gain from the own work effort is the ratio of earnings-replacement benefits and own shadow labour income.

The amount of cash-benefits received at the individual level is constructed as the sum of individual earning-replacement benefits plus the individual's part in any earnings-replacement benefits received at household level (see Box 3 for a listing of benefits counted in these categories).

Shadow labour incomes are estimated using a Mincer-type earnings equation employing the Heckman (1974) procedure to correct for endogenous sample selection. The covariates that enter the (log-) earnings equation are the same in all countries except for the regional effects, which are substituted with the degree of urbanization in countries without regional information (Estonia, Ireland, Portugal and Lithuania): education, age, age squared, gender, health limitations and regions / degree of urbanization. The participation equation uses a number of additional variables for identification: family structure (household type), composition (number of children) and household incomes that are independent of the respective individual's labour supply.

The estimated parameters are used to derive estimates of shadow labour incomes. The amount of earnings-replacement benefits is then compared with the estimated shadow labour incomes to identify the individuals with a low financial gain from own work activities: individuals whose earnings-replacement benefits represent at least 60% of the shadow labour incomes are considered facing financial disincentives.

The table below shows estimated potential earnings and earnings-replacement benefits for those with "high" earnings replacement benefits (60% of potential earnings or above) and for the remainder of the target population across the seven countries.

Determinants of financial work disincentives (substitution effect)

	ESP	EST	HUN	IRL	ITA	LTU	PRT	ALL
"High" earnings replacement benefits, % of target population	11	6	14	11	10	6	11	11
Amount of cash benefits (€/year)	21023	6461	5154	29921	28024	5707	16284	22003
Shadow labour income (€/year)	19777	8134	6376	33444	28471	5924	15974	21891
Reminder of the target population, %	89	94	86	89	90	94	89	89
Amount of cash benefits (€/year)	2306	1522	1067	6573	1797	1042	1840	2107
Shadow labour income (€/year)	17062	8883	5970	34472	22971	6577	12372	18900

Source: Authors' calculations based on EU-SILC 2013.

3.3. Scarce employment opportunities

A lack of employment opportunities in general relates to demand-related constraints in the respective labour market segment. While a number of indicators of labour demand exist at the aggregate or semi-aggregate level, capturing the scarcity of job opportunities at the micro-level can be challenging. Essentially, it is necessary to describe the availability of vacancies in labour-market segments that are relevant for each individual given their skills set, location, etc.

In the present paper, this is operationalised by estimating a risk of demand-side constraints in labour-market segments described by age, gender and education. When regional breakdowns are not available (notably in small countries where the EU-SILC sample consists of only one region), the degree of urbanisation is used instead to capture geographical differences within the country. Scores of demand-side constraints are derived using EU-SILC data on variables relating to labour-market tightness: long-lasting unemployment (corresponding to the 12 months of the income reference period *and* the moment of the interview)[15] or persistent *involuntary* part-time activity. In order to better isolate demand-side constraints, we only include those long-term unemployed who also report active job search and being available to take up employment.[16]

The probability of facing labour market constrains is estimated by means of a simple regression model for the entire reference population. The estimated parameters are then used to predict the risk of being demand constrained (i.e., long-term unemployed or working part-time involuntarily) for each individual in the reference population, given their characteristics and circumstances. The estimated risk depends on the empirically observed relation between covariates included in the regression model and the variable describing labour-market tightness.

Based on the estimates, individuals are grouped into two risk categories following the general approach described in Box 2: those with a risk higher than 1.6 times the median in the reference population are considered facing “scarce job opportunities”.

Table 12 shows the incidence of the scarce job opportunities in the seven countries. The shares are highest in Spain and Ireland, two of the countries that were severely affected by the economic crisis (shares of 51% and 47% of the target population respectively).

Table 12. Scarce job opportunities: incidence in the target population

Shares of individuals with a high risk of long-term unemployment or involuntary part-time employment

	ESP	EST	HUN	IRL	ITA	LTU	PRT	All
Hiring / labour demand: "low"	51	24	41	47	43	33	36	45

Source: Authors’ calculations based on EU-SILC 2013

4. STATISTICAL PROFILES OF EMPLOYMENT BARRIERS

This section describes the results of a statistical method for *segmenting* the target population (jobless and underemployed) into groups that are meaningful for designing, tailoring and targeting activation and employment support policies (AESPs). Building on the employment-barrier indicators developed in the previous section, the objective is to obtain groups (or ‘classes’) of individuals with combinations of employment barriers that are as similar as possible *within* groups, and as different as possible *between* groups.

The objectives of analysing employment barriers in this way are different from a traditional regression analysis. Regression models would, e.g., show how each barrier in isolation affects the risk of facing potential

15. Long-term unemployment, as opposed to unemployment of any duration is taken into account because short-term unemployment could also occur due to a natural and temporary situation (e.g. between 2 jobs) and not to a structural weakness on the demand side.

16 . EU-SILC employment status information for the reference year is not fully in line with the ILO definition of unemployment, but survey respondents are asked to report job search activity and availability for employment at the time of the interview.

labour market difficulties *while holding all other barriers constant*. By contrast, the proposed approach focuses on the interrelations between employment barriers and how they *jointly* determine observed labour-market outcomes. The focus on joint patterns of employment barriers is relevant as the success of AESPs typically depends on their ability to address real-world combinations of different labour-market obstacles.

The segmentation approach focuses explicitly on employment barriers, rather than on other characteristics that are commonly used when breaking down labour-market statistics (e.g. gender and age). As highlighted in the introduction, the reason is that sub-groups that policy debates commonly consider at risk of labour-market marginalization (e.g. 'youth', 'old workers', 'women'), are in fact highly heterogeneous in terms of their employment obstacles. Tailoring policies to only the most prominent real or assumed barriers facing these groups may therefore not be sufficient for increasing their employment chances.

The statistical segmentation method presented in this paper is *Latent Class Analysis* (LCA). This method exploits the interrelations of an array of indicators through a fully-specified (i.e. parametric) statistical model for organizing the target population into homogeneous groups. In the present framework, the indicators represent employment barriers and the statistical algorithm therefore identifies population sub-groups sharing similar barriers to employment, e.g. "low skills *and* limited labour demand" for group 1; "low work experience *and* low financial work incentives" for group 2, etc.

LCA has three main advantages relative to other common segmentation (or 'clustering') methods:

1. Formal statistical tests guide the selection of the optimal number of groups and other model's features;
2. LCA does not allocate individuals into specific groups in a deterministic way but, instead, provides *probabilities* of group membership, thus reducing possible classification errors in any post-estimation analysis;
3. LCA deals easily with common data-related issues such as missing data and complex survey designs.

Sub-sections 4.1 and 4.2 below illustrate the approach using data for **Estonia** and **Spain**. In each case, a first part presents the main barrier profiles characterising each resulting group and provides descriptive statistics using a set of individual and family characteristics (e.g. age and gender). This is followed by a discussion of the extent of policy-relevant overlaps between employment barriers (incidence of multiple simultaneous barriers), including an assessment of the main combinations of barriers using simple Venn-diagram representations.

Two annexes provide further technical details that may be useful for practitioners and others interested in replicating the approach or adapting it to specific analytical needs, data or policy contexts. Annex I outlines the main concepts underlying LCA using a text-book example applied to the context of the employment barriers. Annex II details a step-by-step guide to the model selection process using Estonia and Spain as case studies. To summarize, the model selection process starts with the definition of a *baseline* model that is repeatedly estimated with an *increasing number of latent classes*. The *baseline* model includes the set of indicators derived in Section 3 and takes the form of a *standard* latent class model.[17] The choice of the *optimal* number of classes is primarily based on goodness-of-fit and error-classification statistics and on misspecification tests.

17. A *standard* latent class model means that no additional information enters the model's (log-) likelihood function (in the form of *active* covariates) and that the likelihood function is derived under the so-called Local Independence Assumption (LIA). See Annex II for details.

4.1 Results for Estonia

Using the 2013 SILC data for Estonia, the model selection process outlined above leads to a model with **6 latent classes** (groups). Table 13 shows the estimated parameters, i.e. the *share* of individuals facing the employment barriers in each latent group and the related *group size* in the target population (first row). Groups are ordered by size; colour shadings are used to highlight barriers with higher (dark blue) and lower (light blue) frequencies in each group.

Table 13. Latent class estimates for Estonia

	Class name	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Target Population
	Class size (% of target population)	26	25	21	14	9	6	
	INDICATOR	Share of individuals facing each barrier, by class						
1 -	Low skills	7	22	18	44	17	0	19
2 -	Care responsabilities	13	6	7	14	82	99	21
3 -	Health limitations	32	66	66	32	9	7	44
4 -	No recent work experience (WE)	24	84	99	63	80	52	67
5 -	Have never worked in the past	0	0	0	51	17	0	9
	Positive but low overall relative WE	3	60	10	48	75	11	32
6 -	High non-labour Income	25	7	19	25	41	85	24
7 -	High earnings replacement (benefits)	7	0	10	4	1	31	6
8 -	Low job opportunities (limited hirings)	33	33	5	35	11	0	24

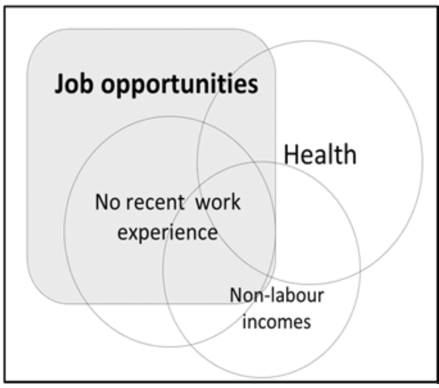
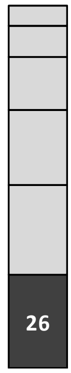
Notes: Section 3 describes the indicators and applicable thresholds. Class sizes refer to the target population as defined in Section 2. Colour shadings identify categories with high (dark blue) and lower (light blue) frequencies. Complementary categories (e.g. 'high' skills) are omitted. Additional information on model selection and model specification is provided in Annex II.

Source: Authors' calculations based on EU-SILC 2013

The following paragraphs describe each group in detail. At the end of each paragraph a box reports a Venn diagram showing extent and degree of overlap of the main barriers characterising the group, as well as a list of selected individual and household characteristics with a high probability of occurrence. Together, this information can help in attaching labels ("*faces*") to group members. Table 14 at the end of this sub-section reports a more complete list of individual and household characteristics.

Group 1 (**26%** of the target population): The main barriers characterizing this group are: scarce job opportunities (33% of group members), health limitations (32%), the possibility to draw on other incomes independently of own work efforts (25%) and no recent work experience (24%). Many individuals in this group are "close" to the labour market (Table 14): 30% are employed with restricted working hours, 25% are short-term unemployed, and 97% have considerable past work experience. Most have also at least an upper-secondary degree (only 7% below) and, although 32% of group members report health limitations, only 8% reports *severe* limitations. The incidence of multiple simultaneous barriers is limited in this group (Figure 5 further below). Individuals with two simultaneous barriers represent about 27% and those with three or more account for less than 16%. Among those with multiple barriers, the most common combination is low job opportunities and no recent employment activity (Box 5).

Box 5. Estonia, Group 1: barriers and characteristics

Main barriers	Selected characteristics	% Target pop.
	- Prime age (average: 45 years) - Mostly women - Couple without young children - 22 years of paid work (average) - Upper-secondary/Tertiary degree (14.2 average years of schooling) - Underemployed or short-term unemployed actively looking and available for work - Urban areas - Average equivalent income: €6126 (corresponds to 3rd quintile)	

Notes:

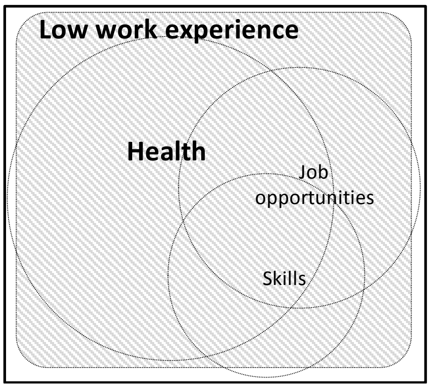
"Main barriers": Surface areas of shapes in the diagram are proportional to the number of group members facing the related barrier ("Venn Diagrams"). The outer square represents the group size (100%). The shaded squared figure represents the most frequent barrier. To facilitate the graphical representation of multiple barriers, the two work-experience indicators ("recent" and "overall" work experience) are combined and shown as one element rather than two. The label attached to the work-experience element highlights the more frequent of the two ("low work experience" if both are equally frequent).

"Selected characteristics": Characteristics that distinguish this group from other groups, i.e., categories that have a high probability of occurring in the group. Table 14 reports individual and household characteristics in more detail. * Income quintiles are calculated for the entire national population.

Source: Authors' calculations based on EU-SILC 2013.

Group 2 (**25%** of the target population): A majority in this group face health issues (66%) combined with no recent work experience (84%). Many of them also have no or limited overall past work experience (61%). 33% are classified in the "scarce job opportunities" category, and 22% have low education. Many group members report to be unfit to work (48%), with 23% facing *severe* health limitations (Table 14). This group has the biggest share of individuals at risk of poverty (61%), most of them are men (64%) aged under 55. In this group, about 46% face 2 simultaneous employment obstacles and a significant 36% face 3 or more simultaneous barriers (Figure 5). The main overlap is between health and low work experience (recent or low overall work experience); these two major barriers often cumulate also with low skills and labour-demand constraints.

Box 6. Estonia, Group 2: barriers and characteristics

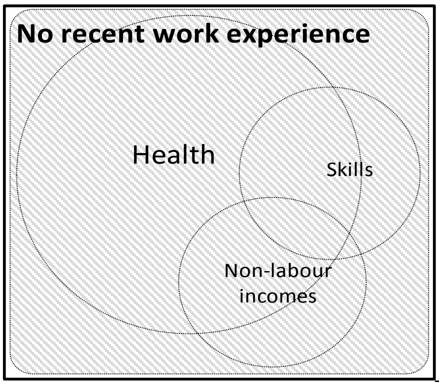
Main barriers	Selected characteristics	% Target pop.
	- Prime age (47 years - average) - Mostly men - 2-adult family without children - 16 years of paid work (average) - Lower-secondary/Upper-secondary (12.7 years of schooling - average) - Disabled / unfit to work - Average eq. income: €3776 (corresponds to bottom *quintile*)	

Notes and sources: See Box 5.

Group 3 (**21%** of the target population): This group is also characterised by health issues and a lack of recent work experience. But, in contrast to Group 2, fewer individuals (19%) report *severe* health issues and group members have more overall work experience (Indicator 5 – Table 13). Many group members are older and have already retired from the labour market in recent years (Table 14). As one might expect in the case of retirees, many of them face possible work-incentive issues (indicators 6 and 7) as a result of out-of-work benefits they may receive. Nevertheless, about 40% of the individuals in this group are at risk of poverty. In this group, 82% have two or more barriers and about 34% have three or more barriers (Figure 5). Excluding the generalized lack of recent work experience characterizing all the individuals in this group, the second most common barrier, health limitations, often cumulates with low skills or low work incentives (Box 7).

Box 7. Estonia, Group 3: barriers and characteristics

Main barriers



Selected characteristics

- Old age (61 years - average)
- Mostly women
- 2-adult family without children
- 33 years of paid work (average)
- Upper secondary/tertiary degree (14 years of schooling - average)
- Retired
- Average equivalent income: €5205 (corresponds to 2nd quintile)

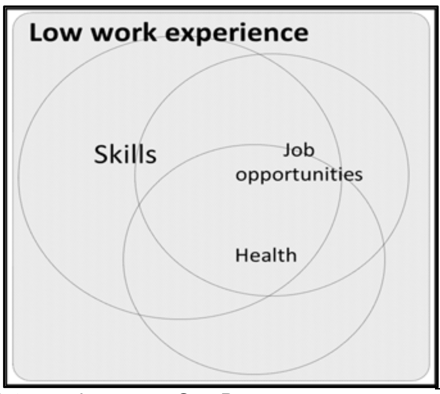
% Target pop.

Notes and sources: See Box 5.

Group 4 (**14%** of the target population): Individuals in this group are characterised by limited work experiences, with half of them having no work experience *at all*. Many of them report low educational attainment (44%) and/or are likely to face labour-demand issues (35%). Other barriers are important as well (health issues, incentive issues due to high non-labour income). Results in Table 14 show that a large majority are youth (77%) and many of them are unemployed (40%), unfit to work (21%) or inactive fulfilling domestic tasks (19%). Figure 5 shows that these group members deal with several simultaneous barriers: 46% have at least 3 barriers and about 20% have four or more barriers. At the top of the low overall work experience characterizing all the individuals in this group, low skills often cumulates with low job opportunities and, to a less extent, to health limitations (Box 8).

Box 8. Estonia, Group 4: barriers and characteristics

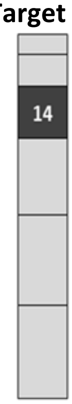
Main barriers



Selected characteristics

- Youth (27 years - average)
- 2+ adult household
- Three years of paid work (average)
- Lower-secondary/Upper-secondary degree (12.2 years of schooling - average)
- Unemployed or inactive
- Average equivalent income: €5438 (corresponds to 2nd quintile)

% Target pop.



Notes and sources: See Box 5.

Group 5 (**9%** of the target population): Most individuals in this group have care responsibilities (82%) and low overall work experience (92%). About 41% can draw on other incomes independently from the own employment effort (e.g., earnings of a spouse) and 63% do not have a recent employment record. This group consists entirely of women with children. This group has the second highest average number of simultaneous barriers per individual (Figure 5) but the second lowest poverty rate.

Box 9. Estonia, Group 5: barriers and characteristics

Main barriers

Selected characteristics

- Youth or prime age (32 years - average)
- Mostly women
- Couple with children
- Six years of paid work (average)
- Upper secondary /Tertiary degree (13.5 years of schooling - average)
- Housework and care responsibilities
- Rural areas
- Average equivalent income: €6700 (corresponds to 3rd quintile)

% Target pop.

Notes and sources: See Box 5.

Group 6 (**6%** of the target population): Care responsibilities are again by far the most frequent employment obstacle facing individuals in this group (99%). However, unlike in Group 5, a vast majority have significant work experience. Financial disincentives can represent a serious issue for them, as 85% can draw on other incomes that are independent of their own employment effort and 31% receive significant earnings replacement benefits. As for Group 5, this group consists mainly of women with children with a very low risk of poverty (only 3%). This group has the highest average number of simultaneous barriers per individuals. On the top of care responsibilities, most of the individuals in this group face simultaneously both work-incentive obstacles and have also no recent work experience (Box 10).

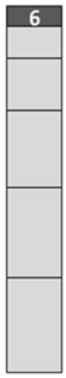
Box 10. Estonia, Group 6: barriers and characteristics

Main barriers

Main characteristics

- Youth and prime age (32 years - average)
- Mostly women
- Couple with children
- Nine years of paid work (average)
- Tertiary degree (17 years of schooling - average)
- Care responsibilities
- Urban areas
- Average eq. income: €12503 (corresponds to top quintile)

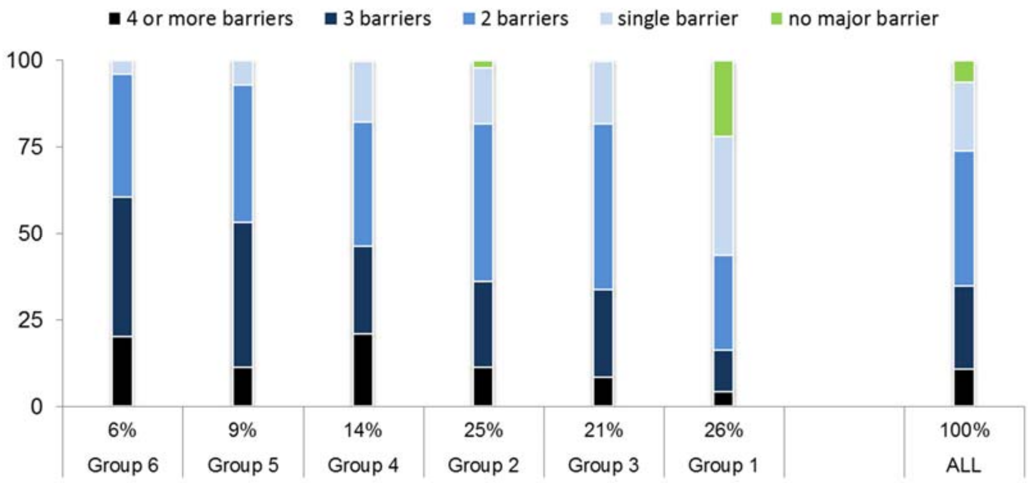
% Target pop.



Notes and sources: See Box 5.

Figure 5. Estonia: Shares of individuals facing multiple employment barriers

In descending order of shares facing at least 3 barriers



Note: Group sizes are reported on the horizontal axis. In counting the number of simultaneous barriers, the two work experience indicators are merged into a single indicator of overall lack of work experience (see Box 5).

Source: Authors' calculations based on EU-SILC 2013

Table 14. Characterization of the latent groups in Estonia

Percentage of individuals with selected characteristics, by group

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Target Population
Live with young children (<=12) (%)*	34	18	6	35	100	100	34
Number of adults (%)							
1 adult	17	25	24	13	7	1	18
2 adults	58	53	55	46	73	88	57
2+ adults	25	22	21	42	20	11	25
Women (%)*	61	36	61	48	100	99	59
Age group (%)*							
Youth (18-29)	15	0	1	77	36	32	20
Old age (60-64)	27	16	98	1	1	0	32
Average age	45	47	61	27	32	32	44
Average years of paid work experience †	22	16	33	3	6	9	20
Highest educational attainment (ISCED)							
Upper secondary	59	64	54	49	59	26	56
Tertiary	35	14	27	7	24	74	25
Average years of schooling †	14	13	14	12	14	17	14
Migrant (%)	20	16	31	5	6	4	17
Severe health limitations (%)	8	23	19	15	1	1	14
Main activity status (%)							
Employed	30	3	0	8	2	5	10
ST Unemployed (<=12 months)	25	12	2	20	10	3	14
LT Unemployed (>12 months)	12	21	4	19	9	0	13
Retired	7	5	64	0	0	0	17
Unfit to work	10	48	28	21	2	0	23
Domestic tasks	13	11	2	19	75	92	21
Other inactive	3	1	0	12	1	0	3
Actively looking and available for work (% out of work)	49	30	7	45	19	5	26
Live in rural area (%)	36	51	46	49	55	35	45
At risk of poverty (%)	31	61	38	38	28	3	39
Income distribution							
1st quintile (Bottom 20%)	32	63	40	39	28	3	41
2nd quintile	23	21	28	22	20	5	22
3rd quintile	18	10	16	19	22	12	16
4th quintile	17	5	11	12	17	27	12
5th quintile (Top 20%)	10	1	5	8	13	53	9
Average eq. disposable hh. income (€/year)	6126	3776	5205	5438	6700	12503	5661
Average number of barriers per individual	*1.4*	*2.3*	*2.2*	*2.5*	*2.6*	*2.8*	*2.1*

Notes: Colour shadings identify categories with high (darker) frequencies. Average numbers of barriers per individual are computed after grouping the two work-experience indicators (recent and low work experience), into a single "low work experience" indicator. Income quintiles refer to the entire population. Poverty risks are calculated following Eurostat methodology.

* "active covariate", i.e. variables that enter directly the probability of group membership during the estimation of the latent class model, as explained in Annex II.

† average across observations with strictly positive values.

Source: Authors' calculations based on EU-SILC 2013.

4.2 Results for Spain

Using Spanish SILC data for 2013, the model selection process resulted in a **7-class solution**.[18] Table 15 presents results for Spain following the same format as used in Table 13 for Estonia above. The following paragraphs describe each group in further detail, based on categories that stand out in terms of their incidence.

Table 15. Latent class estimates for Spain

	Class name	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Target Population
	Class size (% of target population)	39	24	14	7	6	5	4	
	INDICATOR	Share of individuals facing each barrier, by class							
1 -	Low skills	74	50	74	67	24	50	0	61
2 -	Care responsabilities	2	3	6	98	85	3	1	15
3 -	Health limitations	18	43	29	10	15	25	9	25
4 -	No recent work experience (WE)	72	52	90	77	62	89	20	68
5 -	Have never worked in the past	14	6	33	19	11	1	16	14
	Positive but low overall relative WE	40	27	57	34	44	1	41	37
6 -	High non-labour Income	13	17	53	15	31	29	54	23
7 -	High earnings replacement (benefits)	6	14	4	7	5	75	1	11
8 -	Low job opportunities (limited hirings)	100	5	7	99	1	4	59	51

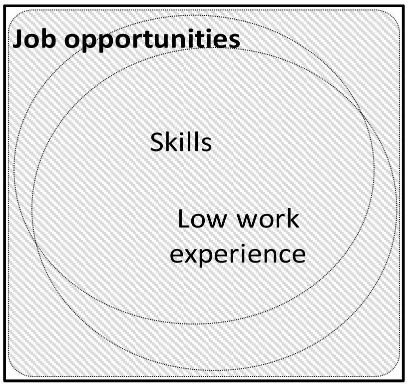
Notes: Section 3 describes the indicators and applicable thresholds. Class sizes refer to the target population as defined in Section 2. Colour shadings identify categories with high (dark blue) and lower (light blue) frequencies. Complementary categories (e.g. 'high' skills) are omitted. Additional information on model selection and model specification is provided in Annex II.

Source: Authors' calculations based on EU-SILC 2013

Group 1 (**39%** of the target population): In line with the still very difficult labour-market situation in Spain during 2012, this largest group consists entirely of people facing high risks of scarce job opportunities resulting from low labour demand and hiring. In addition, 78% have no recent work experience. A lower but still majority share (54%) have worked little or not at all after finishing education. Moreover, three fourths (74%) have less than upper secondary education. About 77% of the individuals in this group are unemployed (60% for more than 12 months, and the majority is actively looking and ready for a job. Older individuals have a low probability of being in this group: Youth and prime-age adults represent 93% of the group and migrants are a significant minority (26%). Poverty risks are the highest across the 7 groups, 50% being below the Eurostat poverty line. Multiple employment barriers are common (Figure 7), with lack of work experience (both recent and low overall) cumulating extensively with low skills on the top of low job opportunities (Box 11).

[18] The baseline model has the same features as for Estonia. See Annex II for details.

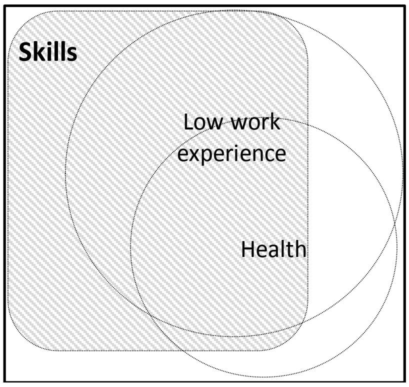
Box 11. Spain, Group 1: barriers and characteristics

Main barriers	Selected characteristics	% Target pop.
	- Youth and prime age (38 years - average) - 2+ adult household - 14 years of paid work (average) - Primary/lower secondary degree (10 years of schooling - average) - Long-term underemployed actively looking and available for work - Urban areas - Average eq. income: €9055 (corresponds to 2nd quintile)	

Notes and sources: See Box 5.

Group 2 (**24%** of the target population): This group is characterised by a large share of people with health limitations (43%), 10% reporting *severe* health issues (Table 16). Health limitations frequently overlap with skills and work-experience barriers (Box 12). Almost half of the individuals in this group have some recent employment activity, and can thus be seen as reasonably close to the labour market. Men are in the majority of this group (57%), most with 55 years or less. Individuals in this group are mainly from the middle class, with an average equivalent disposable income amounting to €12,780/year.

Box 12. Spain, Group 2: barriers and characteristics

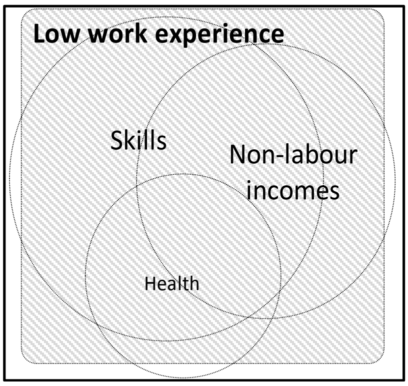
Main barriers	Selected characteristics	% Target pop.
	- Prime age (49 years - average) - Mostly men - 2-adult household - 23 years of paid work (average) - Lower secondary/Upper secondary degree (12 years of schooling - average) - Short-term unemployed or inactive with health issues - Average eq. income: €12780 (corresponds to 3rd quintile)	

Notes and sources: See Box 5.

Group 3 (**14%** of the target population): Almost all the individuals in this group have limited overall work experience combined with no recent employment. The majority have also low education levels (74%) and about 53% have also access to incomes that are independent of own employment efforts. This group consists entirely of old-age women (Table 16) and reporting domestic tasks as their main status (63%). The average number of barriers per individuals is the one of the highest (Table 16, last row), with 57% of the individuals facing 3 or more simultaneous barriers (Box 13).

Box 13. Spain, Group 3: barriers and characteristics

Main barriers



Selected characteristics

- Old age (55 years - average)
- Mostly women
- 2+ adult household without children
- 14 years of paid work (average)
- Primary / Lower secondary degree (10 years of schooling - average)
- Fulfilling domestic tasks
- Average eq. income: €15989 (corresponds to 4th quintile)

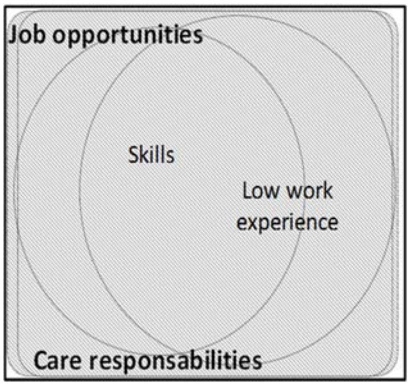
% Target pop.

Notes and sources: See Box 5.

Group 4 (**7%** of the target population): All individuals in this group face at least two simultaneous employment obstacles: care responsibilities and a high risk of low labour demand / scarce job opportunities. Educational attainment is low for a large majority (67%) and most have limited past work experience overall (53%). Table 16 shows that all group members have children and 85% are women. About half are either long-term unemployed or engaged in domestic tasks. They mainly live in urban areas and many of them are migrants (40%). Poverty risks are the second-highest (38%) among the seven groups. This group has the highest average number of barriers per individuals (Table 16, last row), with low skills and work experience obstacles cumulating extensively on the top of low job opportunities and care responsibilities (Box 14).

Box 14. Spain, Group 4: barriers and characteristics

Main barriers



Selected characteristics

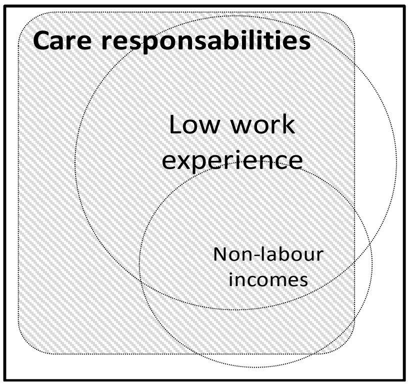
- Youth and prime age (36 years - average)
- Mostly women
- 2-adult couple with children
- 13 years of paid work (average)
- Primary/lower secondary degree (11 years of schooling - average)
- Inactive or long-term unemployed
- Average eq. income: €9780 (corresponds to 2nd quintile)

% Target pop.

Notes and sources: See Box 5.

Group 5 (**6%** of the target population): Like Group 4, care responsibilities are the most common potential employment barrier, all group members have children, and nearly 90% are women. However, unlike Group 4, risks of scarce job opportunities and low educational attainment are not a significant issue, while potential work disincentives (access to incomes that are independent of own employment effort) are relatively common. 62% have not had any recent employment and 55% have limited overall work experience.

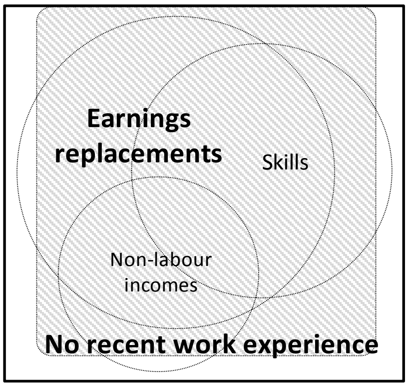
Box 15. Spain, Group 5: barriers and characteristics

Main barriers	Selected characteristics	% Target pop.
	- Prime age (41 years - average) - Mostly women - Couple with young children - 12 years of paid work (average) - Upper secondary /Tertiary degree (14 years of schooling - average) - Short-term unemployed or inactive - Not actively looking or not ready for a job - Average eq. income: €12601 (corresponds to 3rd quintile)	

Notes and sources: See Box 5.

Group 6 (**5%** of the target population): Most have no recent work experience (88%) and face potential work disincentives: 75% receive significant earnings-replacement benefits and many have also access to incomes that do not depend on own employment effort. About half of this group has less than upper secondary education, a barrier that often cumulates with no recent work experience and potential work disincentives (Box 21). This group consists mainly of old-age individuals, typically men (74%) reporting to be permanently retired (64%). Most of the individuals in this group live in wealthy households from urban areas (82%).

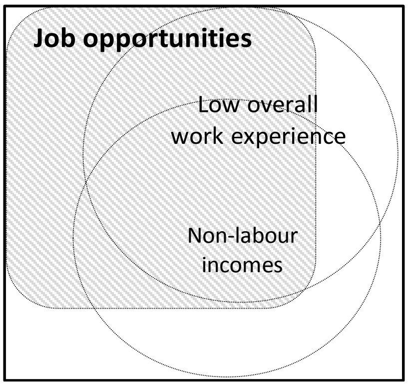
Box 16. Spain, Group 6: barriers and characteristics

Main barriers	Selected characteristics	% Target pop.
	- Old age (61 years - average) - Mostly men - 2+ adult household without children - 39 years of paid work (average) - Lower secondary /Upper secondary degree (12 years of schooling - average) - Retired - Urban areas - Average eq. income: €23469 (corresponds to top quintile)	

Notes and sources: See Box 5.

Group 7 (**4%** of the target population): This group includes mostly youth (99%) who are typically under-employed (33%) or short-term unemployed (43%). The main barriers in this case are low overall work experience (57%) and limited job opportunities (59%). Box 22 shows that for many individuals the two barriers often cumulate also with the possibility to draw on incomes that do not depend on work effort (54%), as most of them still live at home with their parents (Table 16). The youth in this group live mainly in urban areas (76%), are more often women (61%) with a tertiary degree (63%) and live in relatively wealth households (Table 16).

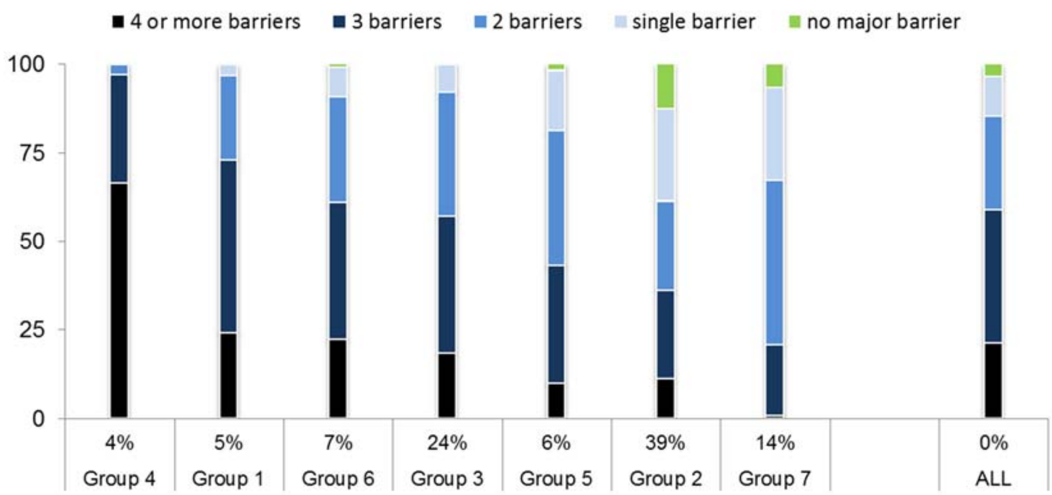
Box 17. Spain, Group 7: barriers and characteristics

Main barriers	Selected characteristics	% Target pop.
	- Youth (26 years - average) - Mostly women - Living with parents - Four years of paid work (average) - Tertiary degree (15 years of schooling - average) - Short-term unemployed actively looking for a job - Average eq. income: €16031 (corresponds to 4th quintile)	

Notes and sources: See Box 5.

Figure 6. Spain: Shares of individuals facing multiple employment barriers

In descending order of shares facing at least 3 barriers



Note: Group sizes are reported on the horizontal axis. In counting the number of simultaneous barriers, the two work experience indicators are merged into a single indicator of overall lack of work experience (see Box 5).

Source: Authors calculations based on EU-SILC 2013.

Table 16. Characterization of the latent groups in Spain

Percentage of individuals with selected characteristics, by group

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Target Population
Live with children (%)*	28	15	4	100	100	1	1	29
Number of adults (%)								
1 adult	9	14	5	5	3	15	6	9
2 adults	44	52	51	82	83	52	27	52
2+ adults	46	35	44	14	14	33	68	39
Women (%)*	49	43	100	85	88	26	61	59
Age group (%)*								
Youth	28	0	0	21	2	0	99	16
Old age	7	29	59	0	5	100	1	24
Average age	38	49	55	36	41	61	26	44
Average years of paid work experience †	14	23	14	13	12	39	4	17
Highest educational attainment (ISCED)								
Upper secondary	16	24	15	17	34	20	37	20
Tertiary	9	26	10	16	42	29	63	19
Average years of schooling †	10	12	10	11	14	12	15	11
Migrant (%)	26	14	7	39	16	5	9	19
Severe health limitations (%)	4	10	6	1	2	5	1	5
Main activity status (%)								
Employed	12	15	4	10	11	3	33	12
ST Unemployed (<=12 months)	17	32	8	12	25	8	43	20
LT Unemployed (>12 months)	60	13	11	48	14	9	14	33
Retired	0	9	5	0	0	64	0	6
Unfit to work	4	18	5	0	1	13	1	7
Domestic tasks	6	10	63	28	45	2	1	19
Other inactive	2	4	4	1	3	1	8	3
Actively looking and available for work (% out	79	26	8	64	21	6	87	47
Live in rural area (%)	30	33	29	26	30	18	24	30
At risk of poverty (%)	50	32	18	38	25	10	19	35
Income distribution								
1st quintile (Bottom 20%)	49	32	18	37	24	9	19	35
2nd quintile	24	21	18	35	27	7	15	22
3rd quintile	16	20	22	15	26	14	20	18
4th quintile	7	15	24	10	14	25	23	14
5th quintile (Top 20%)	4	11	18	2	8	45	23	11
Average eq. disposable hh. income (€/year)	9055	12780	15989	9798	12601	23469	16031	12219
Average number of barriers per individual	3.0	2.0	2.7	3.8	2.3	2.8	1.8	2.7

Notes: Colour shadings identify categories with high (darker) frequencies. Average numbers of barriers per individual are computed after grouping the two work-experience indicators (recent and low work experience), into a single "low work experience" indicator. Income quintiles refer to the entire population. Poverty risks are calculated following Eurostat methodology.

* "active covariate", i.e. variables that enter directly the probability of group membership during the estimation of the latent class model, as explained in Annex II.

† average across observations with strictly positive values.

Source: Authors' calculations based on EU-SILC 2013.

5. CONCLUSIONS

This paper proposes a novel method for identifying, analysing and visualising employment barriers that can prevent working-age individuals from participating fully in the labour market. The underlying premise is that out-of-work individuals (unemployed and inactive) and workers with weak labour market attachment face a number of possible employment obstacles, and each of them may call for different policy responses. The success of activation and employment-support policies, and of social protection measures more generally, is therefor expected to hinge on effective strategies to target and tailor policy interventions to individual circumstances. The paper argues that a systematic and regular assessment of employment barriers, their incidence among different groups, and their possible overlaps, is desirable as a basis for evidence-based policy design, and for adjusting existing policy interventions to changing household circumstances and labour-market situations.

As a first step, the paper derives quantifiable indicators for employment barriers at the individual level using a typology involving three categories: *(i)* a lack of work-related *capabilities* (e.g., a lack of education, skills or work experience); *(ii)* poor *financial incentives* to look for or accept employment (e.g., because of generous out-of-work benefits); and *(iii)* scarce employment *opportunities* (e.g., a shortage of vacancies in the relevant labour-market segment). To operationalise these concepts with available data, the paper proposes a series of feasible indicators and presents results for seven EU countries with very different labour-market institutions and outcomes: Estonia, Italy, Hungary, Ireland, Lithuania, Portugal and Spain.

In a second step, a latent class analysis, a statistical clustering method, is used to do establish the most prevalent *profiles* of employment barriers faced by individuals with no or weak labour-market attachment in two countries with very different labour markets, demographics and policies: Estonia and Spain.

The empirical analysis relies on a multi-purpose household survey (EU-SILC). As with any data source, the choice involves a trade-off between limitations and strengths. For the purpose of this paper, advantages include cross-country comparability, a longer reference period than alternative sources (such as labour force surveys) and information on individual and family characteristics that are key for understanding employment barriers (such as health limitations, care responsibilities, benefit receipt and incomes of other household members). Limitations are the relatively long time-lag between collection and availability of the data, as well as less detailed and reliable information on labour-force status than is available in labour-force surveys. In principle, the approach outlined here can be implemented with alternative sources of survey and administrative data that are characterised by different strengths and limitations.

The main results are as follows:

- On average across the seven countries approximately 60% of the working-age population have employment that is both relatively stable and characterised by significant hours. The remaining 40% may face some form of labour-market difficulties and represent the main target population considered in this paper: they are either unemployed, labour-market inactive, or in restricted-hours employment during all or most of the year.

- In principle, there is therefore enormous scope for activation and employment support policies to address potential employment barriers and strengthen labour-market attachment. However, a more detailed breakdown of the target population also indicates the wide range of circumstances and the resulting need to tailor and target policy interventions. For instance, the typical client group of most active labour market policies – those reporting “unemployed” as their principal employment status – account for no more than one fourth of the target population, while a majority are inactive, in unstable or marginal employment, or cycle into and out of work during the year.

- Indicators of potential employment barriers highlight concrete country differences in terms of key policy challenges. About half the target population in Spain, but only one fourth in Estonia, face high risks of job-opportunity barriers resulting from low labour demand. Low education levels affect 75% of individuals without jobs or in "low-intensity" employment in Portugal, but only 20% in Lithuania. Limited recent work experience is a widespread problem everywhere, but shares of individuals with no prior work experience at all ranges from 10% or less in Estonia, Hungary and Lithuania, to 30% in Italy. Care responsibilities are a frequent employment barrier in Ireland and Estonia, while health-related limitations appear to be especially prevalent in Estonia, Hungary and Lithuania. Financial work incentives are weak for between one fourth and one third of the target group, but generous earnings replacement benefits appear to play a relatively limited role as a driver of poor incentives.

- Sorting ("clustering") individuals according to their employment barriers uncovers patterns that can provide concrete guidance for policy design and targeting strategies. An illustration for Estonia and Spain shows that "short-hand" groupings that are often referred to in the policy debate, such as "youth", "women", "older workers", are far from homogeneous. Instead, they include several distinct sub-groups with very different combinations of employment barriers. This finding highlights that assumptions of relatively uniform patterns of employment obstacles in groups such as "youth" are inappropriate and may distract attention from the specific employment obstacles that policies seek to address.

- In both Estonia and Spain, a large majority of individuals face more than one employment barrier simultaneously. As a result, addressing one type of employment obstacle (such as demand-side constraints in the still-depressed Spanish labour market) may not be enough for boosting employment levels significantly. From a policy perspective, the results point to an urgent need for carefully sequencing different activation and employment support measures, and coordinating them across policy domains and institutions.

A systematic assessment of employment barriers and individual circumstances can help to adapt policy design to the needs of different groups. In a context of limited resources, the results presented in this paper may also highlight priority groups for policy interventions. For instance, very high poverty risks, a large number of youth or a strong over-representation of women in some groups may signal a need to review whether existing targeting strategies meet governments' social-cohesion objectives. Likewise, information on the intensity and number of barriers faced by individuals can inform difficult policy decisions involving trade-offs between equity and efficiency. Such trade-offs can arise, for instance, when considering whether resources should be channelled primarily to those who have the greatest need for support (e.g., those with severe or multiple barriers who are, in some sense, furthest from obtaining or holding a stable job), or to groups with moderate employment difficulties, for whom policy interventions may have a greater probability of success. Finally, information on income levels, poverty status or material deprivation can be suggestive of specific policy directions. For instance, a high poverty risk combined with weak work incentives may call for caution in applying benefit sanctions. By contrast, groups with relatively high incomes and financial disincentives (such as Group No. 6 in Estonia) may indicate scope for targeted benefit reductions or for strengthening mutual obligations principles.

REFERENCES

Arias, O. and C. Sánchez-Páramo (2014), "Back to Work: Growing with Jobs in Europe and Central Asia" World Bank, Washington DC.

Barnay T. (2015) "Health, work and working conditions: a review of the European economic literature" Eur J Health Econ, DOI 10.1007/s10198-015-0715-8.

Becker, G. (1993), "Human Capital: a Theoretical and Empirical Analysis with Special Reference to Education", University of Chicago Press, Chicago.

Bimrose, J. and S.A. Barnes (2011), "Profiling systems for effective labour market integration: New approaches to profiling and holistic assessment", Mutual Learning Programme for Public Employment Services, European Commission (DG Employment, Social Affairs and Inclusion).

Collins, L.M. and S.T. Lanza, (2010), "Latent class and Latent Transition Analysis".

Contini, B. (2010), "Youth employment in Europe: institutions and social capital explain better than mainstream economics (No. 4718). IZA Discussion Papers.

Devaux, M. and F. Sassi (2015), "The Labour Market Impacts of Obesity, Smoking, Alcohol Use and Related Chronic Diseases", OECD Health Working Papers, No. 86, OECD Publishing, Paris, http://dx.doi.org/10.1787/5jrqcn5fpv0v-en.

Eurofound (2012), "NEETS-Young people not in employment, education or training: characteristics, costs and policy responses in Europe.", Publications Office of the European Union, Luxembourg.

European Commission (EC) (2015). "Upskilling unemployed adults (aged 25 to 64): The organisation, profiling and targeting of training provision.", Publications Office of the European Union, Luxembourg.

European Commission (EC) (2015b). "The 2015 Ageing Report, Economic and budgetary projections for the 28 EU Member States (2013-2060)", Publications Office of the European Union, Luxembourg

European Commission (2012), "Social trends and dynamics of poverty and exclusion", Chapter 2 in Employment and Social Developments in Europe, Publications Office of the European Union, Luxembourg.

Ferré, C., H. Immervoll and E. Sinnott (2013), "Profiling of People with No or Limited Labor-market Attachment", Latvia: Who is Unemployed, Inactive of Needy? Assessing Post-crisis Policy Options, World Bank, Washington, DC.

Fernandez, R.M., Castilla, E.J. and Moore, P. (2000), "Social capital at work: Networks and employment at a phone center", American journal of sociology, No. 1288-1356.

Heckman, J. (1974), "Shadow prices, market wages, and labor supply", Econometrica: journal of the econometric society, pp. 679-694.

ILO (2010), Conceptual model to be used for ISCO-08. http://www.ilo.org/public/english/bureau/stat/isco/docs/annex1.doc

Immervoll H, C. O"Donoghue and H. Sutherland (1999), "An introduction to EUROMOD", EUROMOD Working Paper 0/99, Department of Applied Economics, University of Cambridge.

Immervoll, H. (2012), "Reforming the Benefit System to 'Make Work Pay': Options and Priorities in a Weak Labour Market", IZA Policy Paper No. 50, Institute for the Study of Labor, Bonn, http://ftp.iza.org/pp50.pdf.

Immervoll, H. (2013), "Profiles of vulnerable people in Turkey", in: Finn, D., Grun, R., Herrera-Sosa, K., Immervoll, H., Ridao-Cano, C., Uysal, G., & Yener, A. L., Activating Vulnerable People into Good Jobs in Turkey, World Bank, Washington, D.C.

Immervoll, H., & Scarpetta, S. (2012). Activation and employment support policies in OECD countries. An overview of current approaches. IZA Journal of Labor Policy, 1(1), 1-20.

Konle-Seidl, R. (2011), "Profiling systems for effective labour market integration: Use of profiling for resource allocation, action planning and matching", Mutual Learning Programme for Public Employment Services, European Commission (DG Employment, Social Affairs and Inclusion).

Knudsen, J. B., Hansen, M. O., Vandeweghe, B., Vanhoren, I., Rückert, E. and Schulte, B. (2010), "The Mobility and Integration of People with Disabilities into the Labour Market", DG Employment, Social Affairs and Inclusion.

Lazarsfeld, P.F. and Henry, N.W. (1968), Latent Structure Analysis, Boston: Houghton Mifflin.

Lin, N., Cook, K.S. and Burt, R.S. (eds.) (2001), "Social capital: Theory and research", Transaction Publishers.

Marsden, P. V. and Gorman, E.H. (2001), "Social networks, job changes, and recruitment, Sourcebook of labor markets, pp. 467-502). Springer US.

McDonald, S. (2011). "What You Know or Who You Know? Occupation-specific work experience and job matching through social networks", Social Science Research, Vol. 40, No. 6, pp. 1664-1675.

Mouw, T. (2003), "Social capital and finding a job: Do contacts matter? ", American sociological review, pp. 868-898.

OECD (forthcoming) "Getting Skills Right", OECD Publishing.

OECD (2015a), "Activation policies for more inclusive labour markets",Employment Outlook, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/empl_outlook-2015-en.

OECD (2015b), OECD Employment Outlook 2015, OECD Publishing, Paris, http://dx.doi.org/10.1787/empl_outlook-2015-en.

OECD (2014a), "The crisis and its aftermath: A "stress test" for societies and for social policies", Society at a Glance: OECD Indicators, OECD Publishing, Paris, www.oecd.org/els/societyataglance.htm.

OECD (2014b) Employment Outlook 2014, OECD Publishing, Paris, http://dx.doi.org/10.1787/empl_outlook-2014-8-en.

OECD (2013a), "Activation Strategies for Stronger and More Inclusive Labour Markets in G20 Countries: Key Policy Challenges and Good Practices", G20Task Force on Employment, Report prepared for the G20 Summit in St. Petersburg, July, www.oecd.org/g20.

OECD (2013b), "Activating jobseekers: Lessons from seven OECD countries", Employment Outlook, OECD Publishing. http://dx.doi.org/10.1787/empl_outlook-2013-en.

OECD (2013c), Employment Outlook 2013, OECD Publishing. http://dx.doi.org/10.1787/empl_outlook-2013-en.

OECD (2012a), OECD Employment Outlook 2012, OECD Publishing, Paris, http://dx.doi.org/10.1787/empl_outlook-2012-en.

OECD (2012b), "Closing the Gender Gap: Act Now, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264179370-en.

OECD (2011), "Doing Better for Families", OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264098732-en.

OECD (2010), "Sickness, Disability and Work: Breaking the Barriers: A Synthesis of Findings across OECD Countries", OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264088856-en.

OECD Social Benefit Recipients Database (SOCR), http://www.oecd.org/social/recipients.htm.

Sundaram, R., U. Hoerning, N. De Andrade Falcao, N. Millan, C. Tokman and M. Zini (2014), "Portraits of Labour Market Exclusion", World Bank, Washington, D.C..

Sutherland, H. (2001), "Final report EUROMOD: an integrated European benefit-tax model", EUROMOD Working Paper No. EM9/01.

Sutherland, H. and F. Figari (2013), "EUROMOD: the European Union tax-benefit microsimulation model." EUROMOD Working Paper No. EM8/13.

Vermunt J. K., Magidson J. (2016). Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax. Belmont, MA: Statistical Innovations Inc."

Vermunt J. K., Magidson J. (2005). Structural equation models: mixture models, in Encyclopedia of Statistics in Behavioral Science, eds Everitt B., Howell D., editors. (Chichester: Wiley;), 1922–1927

Vermunt, Jeroen K. (2010) "Latent class modeling with covariates: Two improved three-step approaches." Political analysis 18.4 (2010): 450-469.

World Bank (2013), World Development Report: Jobs, World Bank, Washington DC.

Wurpts, I. C., & Geiser, C. (2014). Is adding more indicators to a latent class analysis beneficial or detrimental? Results of a Monte-Carlo study. Frontiers in psychology, 5: 920.

ANNEX I: INTRODUCTION TO LATENT CLASS ANALYSIS

Let us consider 4 *binary* employment barrier indicators: 1) low education, 2) bad health conditions, 3) low financial incentives and 4) scarce job opportunities. Each indicator takes value 1 when the individual faces the barrier and 0 otherwise. Let us consider a sample of N individuals for whom data are collected. Each individual has a specific ***response pattern***, y, depending on the values taken by the four items. Table A.1 shows the first 5 individuals of a hypothetical sample.

Table A.1. Example of barriers faced by hypothetical individuals

Personal ID	Low education	Health limitations	Low financial work incentives	Scarce job opportunities
1	1	1	1	0
2	1	1	0	1
3	0	0	0	0
4	0	0	0	1
5	1	1	1	1

The response pattern of individual 2 is $y = (1, 1, 0,1)$ whereas the response patter of individual 4 is $y = (0,0,0,1)$. Given the presence of 4 indicators with 2 categories each, the *number of possible response patterns* (i.e. the possible combinations of indicator values) is equal to $2x2x2x2 = 16$.

The LCA algorithm searches for the most frequent and similar response patterns in the data and group them into homogenous clusters. For instance, in Table A.1 individuals 1, 2 and 5 have low education, bad health conditions and low financial incentives. Hence, a frequent response pattern could be $y = (1,1,1,.)$, where the dot for the final item (scarce job opportunities) leaves opened the possibility for further differentiations when looking at the overall sample. Another frequent response patter could be $y = (0,0,0,.)$, which characterises the 3rd and the 4th individual in Table A.1.

Let us assume that the two most frequent response patterns *in the whole sample* were $y = (1,1,1,0)$ and $y = (0,0,0,1)$. In such case the LCA algorithm would identify *at least* two **latent classes**: class 1 with $y = (1,1,1,0)$ and class 2 with $y = (0,0,0,1)$. If these were the only two classes then individuals with other response patterns would be allocated *to both* classes *with different probabilities*, depending on the overlap with the most frequent response patterns. For instance, the response pattern $y = (1,1,1,1)$ will have a much higher probability of being part of class 1, whereas $y = (0,0,0,0)$ would be closer to class 2; similarly, $y = (0,1,0,1)$ would be closer to class 2 and $y = (1,0,1,0)$ to class 1. The LCA algorithm, in other words, identifies the most frequent response patterns in the data and then estimates ***probabilities of class membership*** based on the similarity of the response patterns with the frequent ones.

Importantly, the probabilities are *model-based*: they depend on the model specification and the estimated parameters. For this reason, such probabilities are called ***posterior probabilities***, because they can be computed *only* once the model has been defined and estimated. When the posterior probabilities are close to 1 or 0 the model has a high ***predictive power***, as identifies *with precision* the latent-class membership of different response patters.

Two standard ***outputs*** characterise the results of a Latent Class Analysis: 1) the ***marginal distributions*** of the indicators in each latent class (1), and the ***class sizes*** as a share of the total population (2). Table A.2

shows hypothetical estimates for the example outlined above.[19] Class 1 represents 30% of the population (first row) and has the following characteristics: 90% of individuals have low education, 70% have bad health conditions, 76% low financial incentives and 17% scarce job opportunities. For class 2 the situation is different; it accounts for 70% of the population and the only significant employment barrier is the scarcity of job opportunities, which concerns 72% of the individuals in this class.

Table A.2. Hypothetical estimates of the latent class model

		Class 1	Class 2
Class shares		30%	70%
Education	1 (Low)	90%	15%
	0 (High)	10%	85%
Health	1 (Low)	70%	10%
	0 (High)	30%	90%
Financial work incentives	1 (Low)	76%	12%
	0 (High)	24%	88%
Job opportunities	1 (Low)	17%	72%
	0 (High)	83%	28%

Two concepts characterise the results of a Latent Class Analysis:

- The level of ***homogeneity*** within latent classes: the extent to which a unique response pattern *stands out* as having a much larger probability of occurrence than any other;
- The level of ***separation*** between latent classes: the extent to which each class has a peculiar response pattern that makes it different from the others.

In the example above both classes have a quite high level of homogeneity: the response patter $y = (1,1,1,0)$ stands out as having a much larger probability of occurrence than any other in class 1 whereas in class 2 the most probable outcome is $y = (0,0,0,1)$. Similarly, the model has a good level of separation: the likelihood of occurrence of response pattern $y = (1,1,1,0)$ is very low in class 2.[20]

LCA models with a high level of *homogeneity* and class *separation* are easy to interpret, so that **attaching labels** to each class becomes relatively straightforward. Continuing with the example, the first latent class identifies individuals with a *risk of labour-market difficulties related to the area of capabilities and incentives*, whereas individuals in class 2 have *risk of labour-market difficulties mainly related to the scarcity of job opportunities*.

19 For binary indicators the estimates for the residual categories (e.g. high education, high financial work incentives, etc.) are the complement of the other category. Although these are not estimates in a technical sense they are included in the table to show how to interpret results.

20 Separation implies homogeneity but the opposite is not true.

ANNEX II: MODEL SELECTION FOR ESTONIA AND SPAIN

A latent class model does not automatically provide an estimate of the *optimal* number of latent classes. Instead, models with different number of classes are estimated sequentially and the optimal model is chosen based on a series of statistical criteria. The selection process starts with the *definition of a baseline model* (**Step 1**), which is then estimated for an *increasing number of classes*. The choice of the *optimal* number of classes is primarily based on the goodness-of-fit statistics and classification-error statistics (**Step 2**) and then on the analysis of potential misspecification issues (**Step 3**). The final model can be further refined with the inclusion of the so-called *active covariate* (**Step 4**). The following sections describe the step-by-step implementation of each of the above steps.

Step 1: Choice of the baseline model

The *baseline* model has three main characteristics:

- The specification includes *the entire set of indicators* derived in Section 3. Although this characterise the choice of the baseline models for Estonia and Spain it may not be the case for other countries, either because one or more indicators poorly capture the underlying barrier (e.g. due to data limitations) or because the barrier itself does not represent a significant obstacle for that country (e.g. due to successful policies).

- *The employment barrier indicators are the only information provided to the baseline model.* Additional information typically enter the model in the form of *active covariates* (Vermunt and Magdison, 2005), i.e. variables that *actively* contribute to the model's ability to explain latent-class memberships. For instance, in a latent group describing a specific employment barrier profile there could be relevant age and gender differences (e.g. a higher percentage of women). Active covariates can be plugged-in to *test* such differences and, when significant, they can also reduce the classification errors of certain individuals. Although the main role of active covariates is to *describe* the latent classes they can also interfere with the actual *definition* of the latent groups driven by the employment barrier indicators. For this reason, following Collins and Lanza (2010), the *baseline* model used in this paper does not include active covariates. This is important to ensure that the indicators *alone* are able to produce an adequate representation of the target population with clearly identifiable latent classes.[21]

- The indicators are assumed to be pairwise independent *within latent classes*. This assumption, called *Local Independence Assumption* (LIA) originates from the *causal* foundation of the Latent Class Analysis. Box 5 further discusses the LIA and its main implications. Common model misspecification issues originate from the empirical violation of the LIA. It is therefore important to run diagnostic tests on the final model so to ensure that this assumption holds (the *third* step defined above).

---

[21] The association between group memberships and external variables such as age and gender can be investigated ex-post with the method described in Vermunt (2010), which involves computing descriptive measures for the association between covariates and the latent groups after estimating a standard latent class model.

Box 18. The Local Independence Assumption (LIA)

- The traditional theoretical framework of a latent class analysis postulates of the existence of a categorical unobserved (i.e. latent) variable. This latent variable is considered as fully responsible of the variations and co-variations of the observed indicators. In other words, the indicators are caused by the unobserved variable or, which is the same, are related to each other only through the latent variable. Hence, if it was possible to control for the latent variable there would be no residual association between the indicators. To see what the LIA means in practice we report a text-book example from Lazarsfeld and Henry (1968). Suppose that a group of 1000 individuals are asked whether they have read magazines A and B. Their responses are collected in the following matrix:

	Read A	Did not read A	Total
Read B	260	240	500
Did not read B	140	360	500
Total	400	600	1000

- The two variables are strongly related: Readers of A tend to read B more often (65%, i.e. 260/400) than do non-readers (40%, i.e. 240/600). Suppose now that we have information on the respondents' educational levels, dichotomized as high and low. When the 1000 individuals are divided into these two groups the readership are as follows:

	High education		Low education		
	Read A	Did not read A	Read A	Did not read A	Total
Read B	240	60	20	80	500
Did not read B	160	40	80	320	500
Total	400	100	100	400	1000

- It can be easily checked that in each of the 2-by-2 tables in which education is held constant there is no association between the two magazines: Readers of A tend to read B with exactly the same frequency as the non-readers (60%). The reading behaviour is, however, very different in the two sub-groups: the highly-educated group has much higher probabilities of reading both magazines (.60 and .80) than the poorly-educated group (.20 and .20). Hence, the association between A and B can be fully explained from the dependence of A and B on the education level.
- In latent class analysis, the indicators are assumed to be independent given the latent class membership; this means that the latent categories have the same role as the education level in the example above. The term local underlies that the independence between two indicators must hold within the latent groups and not in the overall population.
- The assumption of a categorical variable is what makes LCA suitable for profiling analysis, as it allows organizing individual observations into meaningful and homogeneous subgroups starting from the complex array of empirical indicators.

Step 2: Goodness-of-fit and classification analysis

The favourite goodness-of-fit statistics used in this paper to select the optimal number of latent classes are the *Bayesian Information Criterion* (BIC; Schwartz, 1978) and the *Akaike Information Criterion* (AIC; Akaike, 1987) The BIC and the AIC are measures that capture the *trade-off* between the model's ability to fit the data and the model's parametrization: a model with a higher number of latent classes always provide a better fitting of the underlying data but at the cost of complicating the model's structure. The BIC and the AIC summarise this trade-off into a single index, which provides guidelines for choosing between an adequate representation of the population into a finite number of sub-groups and an increasing complexity of the statistical model. A *smaller* value of these indices indicates a *more optimal balance* between model fit and parsimony; thus, the model with a number of latent classes that minimizes the AIC and BIC is typically the best choice.[22]

22 BIC=-2log(LL)+log(N)q, while AIC=-2log(LL)+2q. In these equations LL is the log-likelihood function of the LCA model, q the number of parameters and N the sample size.

Classification statistics provide further information for the choice of the optimal number of latent classes. These measures summarise how-well the model is able to classify individuals into clusters. The simplest classification statistics is based on the *share* of individuals that are estimated to be miss-classified according to the *modal* assignment.[23] In general, a lower value signals a better classification of individuals into specific latent classes (Vermunt and Magdison, 2005). Although a certain amount of classification error is natural in latent class analysis, values above 30% signal that the model is not able to discriminate between classes in the allocation of a significant amount of cases.

Figure 5 - Panel A shows for **Estonia** the percentage variations of the BIC and the AIC for an increasing number of latent classes. Panel B provides the same information for **Spain**. For models with few latent classes the percentage variations are relatively large because the model's ability to fit the data increases significantly compared to the model's parameterization.[24] For a higher number of classes the increment of the goodness-of-fit is progressively compensated by the higher parameterization, thus producing smaller, and eventually positive, changes in the two measures.

The BIC is minimized in **Estonia** for a model with 6 classes whereas the AIC points to a 12-class solution. Similarly, in **Spain** the BIC points to an 8-class model and the AIC to a 12-class solution. The difference in the two statistics depends on the different penalty that the two measures apply to the increasing goodness-of-fit: the AIC takes into account only the higher number of parameters whereas the BIC considers also the overall sample size. Thus, in general, the BIC points to a more parsimonious specification than the AIC.

The classification error in **Estonia** is below 30% for both the 6-class and the 12-class solution. Both models provide therefore a good representation of the underlying data, even though the 6-class solution points to a lower classification error than the 12-class model (22% versus 25%). The 12-class model for **Spain** has a classification error of about 30%, a value that signals a poor predictive power. The 8-class model has a lower classification error, corresponding to about 24% – a value in line with the model for **Estonia**.

When BIC and AIC point to different numbers of classes the classification error statistic provides further information for the selection of the optimal model. The 6 and the 8-class solution are therefore the optimal choices for Estonia and Spain, respectively.

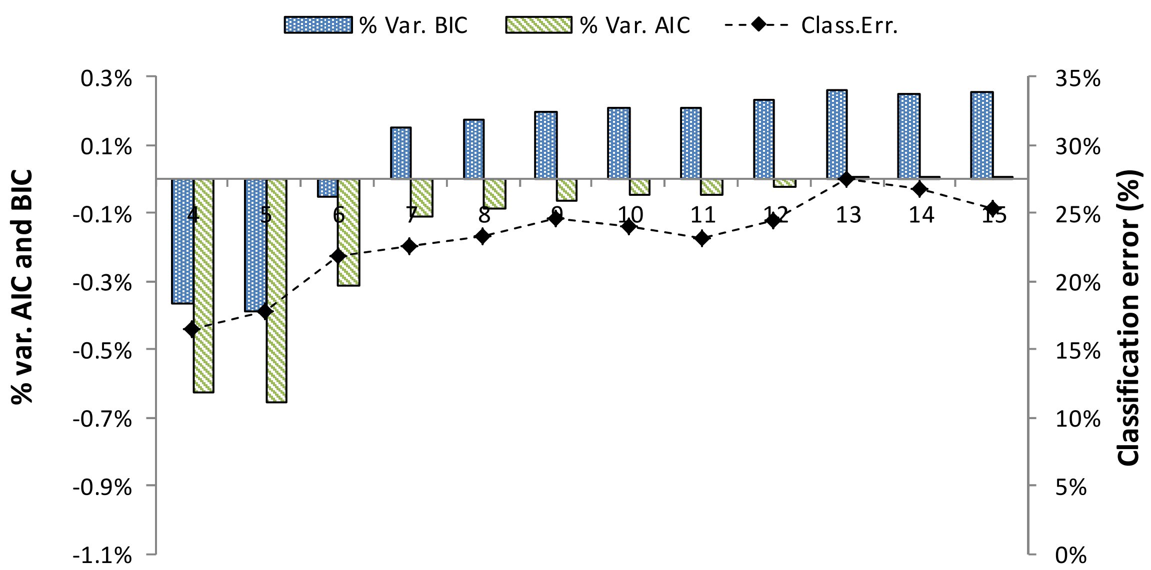
---

23 Latent class models do not assign individuals to specific classes but estimate probabilities of class membership. After estimation one has in general two options to analyses the results: assigning individuals to each cluster based on the highest probability of class-membership (the modal assignment) or weighting each person with the related class-membership probability in the analysis of each class (the proportional assignment). Let us consider an individual with class-membership probabilities of 0.6 and 0.4 for classes 1 and 2, respectively. In case of modal assignment such person would be assigned to class 1 whereas in the case of proportional assignment the same person would enter the class 1 with a weight of 0.6 and class 2 with a weight of 0.4. When class-membership probabilities are far from 0 or 1 the modal assignment can produce a significant classification error. The classification-error statistic discussed in the text is based on this intuition.

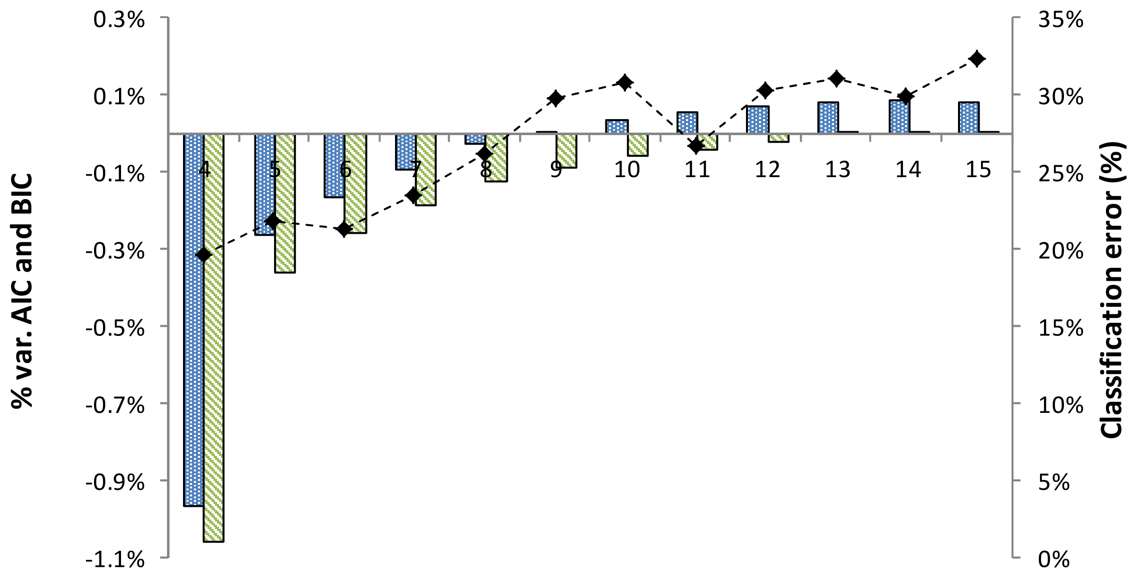
24 The variation of the two criteria for increasing number of classes is somehow counterintuitive: negative at the beginning and positive at the end. This depends on the mathematical specifications of the two measures: the number of parameters contributes positively and the goodness-of-fid negatively.

Figure 7. Selection of the optimal number of latent classes

Panel A - Estonia



Panel B - Spain



Source: Authors calculations based on EU-SILC 2013.

Step 3: Misspecification tests

The model selected through goodness-of-fit and classification statistics may not be optimal due to misspecification issues. The most common misspecification issue in latent class analysis is related to the violation of the Local Independence Assumption (LIA) described above (Box 5). This assumption shapes the mathematical specification of the statistical model and, in practice, requires the indicators to be *pairwise* independent *within* the latent groups. When this requirement is not met the model is not be able to reproduce the *observed* association between the indicators, at least for the indicators showing some residual within-class (*local*) dependency.

In this paper the presence of local dependencies between pairs of indicators is tested using a statistical test on the so-called *Bivariate Residuals* (Vermunt and Magdison, 2016). The idea is to compare the *observed* associations between pairs of indicators with the *expected* association under the assumption of local *independence*. Large differences between estimated and observed associations signal violations of the LIA. The formal statistical test on the bivariate residuals takes the form of a standard Pearson chi-squared test, which can be easily computed once the model has been estimated. In general, *bivariate residuals* with values higher than 1 imply the presence of some residual correlation within latent groups.

Violations of LIA can be addressed by simply increasing the number of latent classes or modelling explicitly the local dependencies between pairs of indicators. Local dependencies between pairs of indicators are typically included in the model with ad-hoc parameters, the so-called *direct effects* (Vermunt and Magdison, 2016). The inclusion of direct effects in the model specification eliminates any residual correlation between the indicators (by construction) but it also requires repeating the entire model selection process, as the new baseline model with local dependencies may lead to a different optimal number of classes.[25]

In **Estonia**, the 6-class model selected with the BIC shows clear signs of misspecification, with bivariate residuals significantly higher than 1 for several pairs of indicators.[26] Increasing the number of latent classes always reduces the local dependencies between indicators. For instance, the 12-class model show signs of local dependencies only for one pair of indicators but, as discussed above, has the drawback of a poorer predictive power (i.e. a higher classification error).

When the higher number of latent classes does not eliminate the residual associations between indicators or when the higher number of latent classes further increases the classification-error statistic, the only alternative is to change the baseline model specification, modelling explicitly the possibility of local dependencies between pairs of indicators. Seven direct effects are necessary in the 6-class model for **Estonia** to eliminate the local dependencies. For this new baseline model the BIC now points to a 4-class solution whereas the AIC to a 10-class solution. The 4-class solution shows however new residual associations between other pairs of indicators, while the 6-class model with direct effects, by construction, does not show misspecification issues. Considering the higher classification error characterising the 10-class model with direct effects, the **6-class** model therefore remains the favourite option for **Estonia**.

In **Spain** the 8-class model chosen according to the BIC shows signs of misspecification for nine pairs of indicators. The 12 class model reduces the misspecifications to only three pairs but the model has a too-high classification-error statistic (Figure 5). The inclusion of direct effects in the 8-class model eliminates any residual associations between pairs of indicators. When running again the goodness-of-fit analysis for the model with direct effects the BIC now points to a 7-class solution, which has also all the bivariate residuals well below 1. The inclusion of the direct effects for Spain has also the beneficial effect of reducing the classification error for the 7-class model from 24% to 20%. The **7-class** specification with direct effects represents therefore the favourite option for **Spain**.

Step 4: Model refinements – inclusion of active covariates

In most empirical applications the aim of latent class analysis is not just to build a classification model based on a set of indicators but also to relate the class membership to other individual and household characteristics identifying specific population sub-groups of interest, such as *youth* and *women*. For instance,

---

[25] Direct effects change the likelihood function of the underlying statistical model. They correspond to the parameters of the interaction terms between the pair of indicators with local dependencies. These interaction terms enter the linear specification of the logit models characterizing the probability structure of the indicators with local dependencies; the related parameters are then estimated together with the other model's values.

[26] Results are available upon request.

one might want to know if there are gender differences between the latent groups and test whether these differences are significant. There are two ways for dealing with additional individual and household characteristics in latent class analysis: a **one-step** and a **three-step** approach. The former involves estimating the latent class model with the additional characteristics (covariates) *actively* contributing to the definition of the group-membership probabilities. The stepwise approach, instead, keeps the covariate out of the model and focuses on them only during the post-estimation analysis. This implies the following three steps: 1) estimating the model without covariates; 2) allocating the individuals to the latent groups using the estimated group-membership probabilities; 3) computing two-way tables summarizing the relation between group-membership and covariates.

In general, the one-step approach is more efficient than the stepwise approach. Under the assumption that the model is well specified one can directly *test* if there are significant group differences (e.g. by gender) between latent classes by simply using the (correct) standard errors estimated with the other model's parameters. An extensive empirical literature has also shown that active covariates with large and significant effects on the group-membership probabilities can reduce the classification-error statistics and, more generally, have a positive effect on the overall model predictive power (Wurpts and Geiser, 2014 and Vermunt and Magidson, 2005).

An important assumption of the latent-class model with covariates is that they affect *directly* the probabilities of group-membership, but not the marginal distributions of the indicators in each latent group. For instance, the share of *women* can be different between groups but the way man and women react to a given indicator *within* the same group must be the same. In other words, latent class models with covariates implicitly assume that covariates and indicators are *independent* within latent groups. This assumption closely resembles the LIA described above and can be tested with a very similar strategy, i.e. with tests on the residual associations between indicators *and* covariates.

When there is evidence that some categories of interest (*e.g.* women) are structurally different from the rest of the individuals *in the same latent group* – i.e. they differ not only in terms of shares between groups but also in terms of responses to the indicators – the inclusion of active covariates can produce misspecification problems similar to those described in Step 3 above. Depending on the extent of such misspecifications (i.e. the number of bivariate residuals between combinations of indicators and covariates) one has in general three options: 1) address the misspecifications by explicitly modelling the associations between indicators and covariates with *direct effects* (as discussed in Step 3 above); 2) depending on the interest in a specific covariate, run a separate LCA between the categories of interest (e.g. for men and women separately); 3) include the covariate(s) causing the major misclassifications issues directly in the classification model.

In the present paper the inclusion of active covariates is primarily driven by the interest in specific population sub-groups that are typically considered in the breakdown of common labour market statistics. The selected active covariates are **age** (3 categories), **gender** and the presence of **young children**. The choice of these three variables relies on practical considerations, i.e. the relevance of these categories in the policy debate on AESPs and also on the possibility for the public employment services to actually collect such information, and on their impact on the statistical model in terms of misspecification issues and classification-error statistics.

The inclusion of active covariates in **Estonia** and **Spain** generates several associations between indicators and covariates within the latent groups. This signals that the employment barrier indicators capture only part of the heterogeneity in the data. The inclusion of such variables in the model is therefore beneficial in terms of the overall model goodness-of-fit and predictive power, with the classification error statistics dropping from 22% to 14% in **Estonia** and from 20% to 10% in **Spain**.[27]

A significant reduction of the classification-error statistics in models with active covariates is the sign that, for some individuals, the employment-barrier indicators alone do not produce a clear-cut latent-class assignment and that, therefore, the covariates are playing an important role not only in improving the latent-class membership but also in shaping the main barrier profile characterizing some of the latent groups. While this does not typically affect the barrier profiles of the biggest groups (i.e. those with the biggest shares in the target population) the barrier profiles of the smallest groups could be partially shaped around the interaction between the information provided with the active covariates and the indicators.[28]

27 The residual within-class associations between indicators and covariates are modelled explicitly with the inclusion of *direct effects* in the final model specification. The other strategy, i.e. running a separate LCA for the covariates showing the main associations with the indicators, are outside the scope of the present paper but can be considered in more-applied country-specific studies. The inclusion of the covariates directly in the classification model led to results that are broadly comparable to the model with active covariates and direct effects. The latter represent therefore the final specification used in this paper for both Estonia and Spain.

28 This should be considered as an improvement with respect to a model without covariates whose indicators do not always produce clear-cut latent-class assignment for some individuals. In fact, without additional information, the allocation of these individuals into a specific latent group would be done almost at random, whereas in models with covariates the allocations of this individuals depends on the active covariates and how they interact with the indicators.