Developing a cross-national comparative framework for studying labour market segmentation: Measurement equivalence with latent class analysis

Martin Lukac^{a,b}, Nadja Doerflinger^{a,c} & Valeria Pulignano^{a,d}

^a Centre for Sociological Research, KU Leuven, Parkstraat 45, 3000 Leuven, Belgium

^b Corresponding author, martin.lukac@kuleuven.be, +32 16 37 79 08, ORCID: 0000-0002-1747-3168

^c nadja.doerflinger@kuleuven.be, +32 16 32 31 84, ORCID: 0000-0003-2894-6132

^d valeria.pulignano@kuleuven.be, +32 16 32 31 62, ORCID: 0000-0002-6466-491X

This is a post-peer-review, pre-copyedit version of an article published in Social Indicators Research. The final authenticated version is available online at: <http://dx.doi.org/10.1007/s11205-019-02101-3>

Abstract

This article proposes a novel measurement model of labour market segmentation in Europe for cross-national comparisons, tackling three drawbacks of current approaches: First, as segmentation is a multi-dimensional concept, it necessitates a complex measurement approach combining several indicators. Second, to date, we lack methodological evidence that earlier used measures are comparable across countries. Third, as any measure of social phenomena contains measurement error, segmentation research may be confounded by misclassification error. To overcome these drawbacks, we argue for modelling segmentation as a latent categorical concept by means of characteristics of the employment relationship. Our analysis shows that accounting for measurement non-equivalence in cross-national labour market segmentation research is crucial to arrive at reliable and unbiased comparative conclusions. The results demonstrate the importance of increased complexity in measuring labour market segmentation. Overall, this article serves as a methodological cross-national comparative framework for future quantitative analysis of labour market segmentation.

Keywords: labour market segmentation; latent class analysis; measurement equivalence; measurement error; cross-national comparative research

Acknowledgements

We are very grateful to the editor and the reviewers for their insightful comments and suggestions. We also thank Jeroen Vermunt and Steven P. Vallas for their feedback during the initial phases of this research as well as the participants of the Industrial Relations in Europe Conference 2018 (10-12 September 2018 in Leuven, Belgium) for the stimulating discussion. .

Funding

This research is part of the research project “*How are social divides produced in contemporary labour markets?*”, which is financed by the Federale Wetenschappelijk Onderzoek (FWO) with a project number G071716N and KU Leuven C1 Grant number C14/16/015.

1. Introduction

Contemporary scholarly work has illustrated the relevance of using an occupational perspective on patterns of non-standard and low pay employment to study labour market segmentation (Eichhorst and Marx 2015). In particular, findings show that divisions between standard and non-standard forms of employment and levels of pay retain analytical and empirical importance when examining social cleavages in labour markets within and across different occupations. On the other hand, research has also emphasized the risk of reliance on pre-defined patterns and characteristics of labour market segments, such as standard and non-standard forms of work (e.g. Yoon and Chung 2015). Based on this, researchers have put forward the need to empirically measure actual patterns of labour market segmentation across countries (Davidsson and Naczyk 2009). In this paper, we propose a novel methodology to measure labour market segmentation, which deals with the aforementioned shortcomings.

European labour markets have witnessed increasing inequality and dualisation. New social divides have been created by conditioning social rights on labour market status (Emmenegger et al. 2012; Clasen and Clegg 2011). These divides seem to favour social security for labour market insiders (i.e. people with secure employment), while disfavouring labour market outsiders (i.e. those trapped in unstable and insecure employment). Ongoing institutional changes reinforce these social divides (Hassel 2004; Palier and Thelen 2000), leading to a gradual increase of the proportion of outsiders. Labour market segmentation can be linked to the dual labour market hypothesis (Doeringer and Piore 1971; Goldthorpe 1984; Kalleberg 2003). Accordingly, the primary labour market is insulated from market fluctuations, offering “relatively high wages, good working conditions, chances of advancement, [...] and above all, employment stability” (Piore 1975: 126). In contrast, the secondary labour market is governed by unrestrained market mechanisms of supply and demand and characterized by low-paid (service) jobs, fixed-term employment, part-time work, poorer conditions, and low employment stability.

As analyses of labour market segmentation and insider-outsider divides illustrate, segmentation might be sensitive to the institutional context—especially to employment and welfare regimes (e.g. Häusermann and Schwander 2012). As a result, the universal nature of the pre-assumed categories according to the traditional understanding of labour market segmentation should be questioned empirically. Although a wide range of research on work,

employment and social policy acknowledges the complex and dynamic nature of labour market segmentation (see Davidsson and Naczyk 2009 for a review), few studies have (empirically) tackled this issue or reflected upon measurement approaches towards segmentation. This article aims to fill this gap and investigate to what extent labour market segments are comparable across a selection of European countries.

Investigating labour market segmentation in a cross-national comparative way requires dealing with the challenge of applying one definition of labour market segments across different institutional contexts of countries. According to Locke and Thelen (1995) this implies developing “contextualized comparisons”, which includes identifying not only similar, but also apparently different context-specific phenomena, whose manifestations are revealing of similar meanings within distinctive institutional arrangements. Accordingly, this paper tackles the three main issues emerging from the effort to measure labour market segmentation across countries in Europe:

(1.) *Complexity of measurement* — In line with scholars calling for a more complex measurement (Davidsson and Naczyk 2009; Yoon and Chung 2015), we argue for using several indicators concurrently for measuring labour market segmentation as a latent variable and propose a theory based and empirically driven model of segmentation.

(2.) *Comparability across countries* — By applying the theoretical framework of measurement equivalence (Billiet 2003; Davidov et al. 2014), we investigate the extent to which labour market segments can be compared across countries. We present a model that corrects for inequivalence of measurement and explains which comparisons of labour market segments are attainable with the proposed measurement model.

(3.) *Misclassification error* — As all measurement instruments suffer from misclassification error (i.e. incorrectly labelling insiders as outsiders and *vice versa*), the proposed measurement model of labour market segmentation quantifies this error and enables correcting for misclassification in its relation to other variables (Bolck, Croon and Hagenaars 2004), hence giving a more precise estimation of the various segments’ populations.

This paper develops a quantitative methodological approach addressing these three issues by latent class (LC) analysis (Goodman 2002; McCutcheon 1987; Hagenaars and McCutcheon 2002). LC analysis is used to extract a number of scientifically meaningful latent

classes that are capable of explaining complex relationships between several observed variables—addressing the first issue of complexity of measurement. Labour market segment membership is operationalized as a multi-dimensional concept through insecurity (Rodgers 1989). In this paper, we assume that the degree of insecurity linked to the terms and conditions of the employment relationship is the main concept distinguishing labour market segments in Europe (Doerflinger, Pulignano and Lukac 2017). Hence, measures of temporal, organizational, and economic insecurity are jointly used to estimate labour market segment membership. Moreover, the LC analysis framework, allows for investigation of comparability in cross-national settings and for accommodation of various levels of measurement inequivalence (Kankaraš, Moors and Vermunt 2010)—responding to the second issue mentioned above. Finally, LC analysis is a stochastic model that assumes that respondents answer survey items only with a specific probability, therefore allowing quantification of the misclassification error (Eid, Langeheine and Diener 2003). All of this together brings substantial improvement to previous efforts to measure labour market segmentation. This is essential as the conclusions drawn about labour market segmentation are contingent upon the quality of measurement that is used.

To illustrate this framework, we present an empirical analysis of labour market segmentation in Europe, using the European Union Labour Force Survey (EU-LFS) data from 2015 and a sub-sample of 11 countries (Belgium, Germany, Denmark, Finland, France, Ireland, Italy, Poland, Portugal, Slovenia, and the United Kingdom) representing liberal, coordinated, and mixed market economies (Hall and Soskice 2001) and various welfare regimes (Esping-Andersen 1990). For data manipulation and cleaning, we used R (R Core Team 2018), RStudio (RStudio Team 2016), and dplyr package (Wickham et al., 2017). Models were estimated in Latent GOLD (Vermunt and Magidson 2005), and all R code and Latent Gold syntax used for this paper can be found in the article’s online appendix.

The remainder of the article will proceed as follows. First, we review the main literature for labour market segmentation and elaborate about its measurement. Second, we introduce the concept of measurement equivalence, how it is possible to measure it with LC models, and how LC models tackle the problem of misclassification. Third, we present analytical results of labour market segmentation in Europe. We show the resulting patterns of segmentation, how they are distributed across the countries, and investigate their social

configurations. Finally, we summarize the findings and discuss their implications for the measurement of segmentation and for labour market segmentation theory in general.

2. Measuring Labour Market Segmentation

Recent research has emphasized the need to shed light on the actual patterns of labour market segmentation across different countries (Doerflinger et al. 2017). Scholars using existing dichotomies—e.g. industry vs. service sector work, or skilled vs. unskilled labour—to investigate labour market segmentation often rely on pre-defined patterns and characteristics of segmentation (for example, temporary employment contract is often equated with disadvantaged labour market position in the secondary segment). Based on such preconceptions, labour markets are subsequently probed about the extent to which segmentation exists; however, the underlying assumption that the drivers of segmentation are functionally equivalent across countries is hardly empirically investigated. In other words, the pre-assumed distinction between the primary and secondary labour markets is expected to be equivalent across countries. As a result, labour market segments are treated as homogeneous social categories across countries. In so doing, there is a risk of not uncovering country-specific major dividing lines, which are likely to vary across countries due to differences in social, political, economic and cultural institutions (Doerflinger et al. 2017).

The study of Yoon and Chung (2015), using LC analysis, addresses the issue of complexity of measurement of labour market segmentation. Combining various indicators to measure segmentation, they identify a third segment (“future insecure”) in between the primary and secondary one. Yet, the study is limited to one country and therefore does not tackle the issue of the cross-national comparability of the uncovered divides. Accordingly, it remains a challenge to address “the limitations of applying one single definition of labour market divisions across different countries” (Yoon and Chung 2015: 3).

Contemporary analyses of labour market segmentation and the insider-outsider divide demonstrate that segmentation is contingent upon a country’s institutional context. This suggests that segmentation is institutionally sensitive, with variation across countries in accordance with different employment and welfare regimes in Europe. For example, Häusermann and Schwander (2012) revealed four different types of segmentation in their

cross-national study, which they combine to diverse typologies of employment and welfare regimes. Those are claimed to be resulting from the evolution of labour market and welfare policy reforms in Europe. Furthermore, other studies examine the extent to which this evolution affects the stability of the primary (insider) and secondary (outsider) segments across and within countries over time (Biegert 2014; Blossfeld et al. 2008). Such studies underline that segmentation configurations are not necessarily stable in time and may be subject to change. Variation in segmentation and over time makes Huws (2014) hypothesize that labour market segments are not homogenous; she therefore advocates considering the primary and secondary segments as two extremes of a continuum, with other segments possibly existing in between.

Changes in segmentation patterns could for instance be related to the decline of industrial employment and growth of service work from 1970s/80s onwards. Accordingly, occupations became a device to investigate segmentation lines and non-standard employment across and within countries. Non-standard employment is widespread in personal and elementary service occupations that often also feature low pay (Eichhorst and Marx 2015). However, non-standard work also occurs in some high-skill service occupations; therefore, solely using this indicator to study segmentation may generate inaccurate results. Although much comparative research in the tradition of the sociology of work, employment and policy acknowledges the complex and dynamic nature of labour markets, little research has attempted to empirically measure segmentation cross-nationally to date. Methodological approaches incorporating the complex nature of segmentation are scarce, particularly in terms of identifying segments that are in between what is commonly referred to as primary and the secondary segments.

In the 2000s, various scholars have used the concept of 'job quality' as a way to study segmentation between different groups in the labour market (e.g. Munoz-Bustillo et al. 2009, Drobnič et al. 2010). A common characteristic of such studies is the agreement that the concept is multidimensional and should therefore be measured by multiple indicators (for a more detailed discussion on job quality, see for example Van Aerden, Moors, Levecque and Vanroelen 2014). As a result, a variety of measurement approaches and indices were created, however, the issue of cross-national comparability of job quality was hardly questioned (e.g. Vandenbrande et al. 2012; Davoine, Erhel and Guergoat-Larivière 2008).

This paper proposes a quantitative methodological framework to tackle both challenges—complexity of measurement and cross-national comparability. But how can segmentation in contemporary labour markets be measured cross-nationally? In line with Doerflinger et al. (2017), we argue that insecurity is a useful concept to distinguish labour market segments nowadays. Ongoing debates highlight the relevance of insecurity as a distinctive feature of labour market segmentation (e.g. Eichhorst, Marx and Wehner 2016; Vosko 2005). While insecurity has often been measured by contract type as a proxy indicator, we argue that a more complex operationalization of insecurity is needed to distinguish between secure and insecure jobs. Solely using contract type is problematic as within-group variation is likely to remain hidden (Doerflinger et al., 2017). Relatedly, a wide range of research reveals that permanent contracts do not always come along with good and temporary contracts with bad working conditions (e.g. Lautsch 2002; Weil, 2014). Therefore, we argue that insecurity does not only relate to the form of the employment relationship (i.e. contract type), but to its terms and conditions (Doerflinger et al., 2017).

To operationalize this approach, we use Rodgers' (1989) distinction between four dimensions of insecurity—namely temporal, organizational, economic and social. These dimensions are inherently linked to the employment relationship's terms and conditions. Specifically, temporal insecurity refers to the continuity of employment; i.e. stable or continuous jobs generally grant security. This concept can be linked to the type of employment contract. Permanent jobs are continuous and stable, whereas temporary contracts are by definition more unstable and, therefore, insecure due to ending at a particular date. Furthermore, aspects of working time also play a role. On the one hand, we make a distinction between full- and part-time work. Although some part-time work can come along with stability, many part-time jobs, particularly in private services, tend to come along with high levels of instability. Therefore, they may not be continuously available and come along with insecurity. On the other hand, the incidence of non-standard working hours is included in the analysis. Non-standard hours (night and weekend work) can not only be found in stable manufacturing jobs involving shift work, but also in many (unstable) service jobs. Hence, we assume higher incidence of non-standard hours to come along with temporal insecurity.

Organizational insecurity refers to the degree of control a worker has over work processes. Workers with low levels of control over their working conditions tend to have more unstable jobs than those who are able to exercise control. Therefore, organisational insecurity is inherently linked to hierarchies in companies, as more control comes along with higher positions (which are also less standardised) in organisational hierarchies. Based on this consideration, we use the variable ‘supervisory responsibilities’ as a proxy to measure organizational insecurity.

Economic insecurity is linked to wages, with lower wages corresponding to higher levels of insecurity due to possible problems to make ends meet and the increased risk of poverty. Thus, wage deciles are used to cover the economic insecurity dimension.

Social insecurity relates to the extent to which workers are protected against issues like dismissals or discrimination and to aspects of social security provision. Although this is relevant, social protection is related to the macro-institutional and legal framework; therefore, it is not the terms and conditions of the employment relationship which entitle a worker to protection, but rather the fact that the worker has an employment relationship. Therefore, social insecurity is not considered in our analyses.

Overall, we argue that this operationalization reflects the latent and multi-dimensional nature of labour market segmentation. However, it is important to acknowledge that insecurity is not only generated by objective characteristics as highlighted in our approach and operationalization. There is also a subjective side to insecurity focused on workers’ perceptions, experiences, and expectations (see for instance Green 2009; Chung and Mau 2014; Chung and van Oorschot, 2011). However, integrating subjective insecurity in our approach is out of the scope of this paper.

3. Measurement Equivalence in Comparative Research

Comparative research questions often span across various social groups or nations. To ensure that we draw valid comparative conclusions, we need to scrutinize whether the investigated concept is equivalent (i.e. comparable) across the analysed groups. This crucial issue in the design and analysis of comparative studies is known as measurement equivalence (Billiet 2003; Davidov et al. 2014). A concept is equivalent if it works in the same way in each

investigated group. Conceptually, this condition can be fulfilled by theoretical reasoning that aims to show that there are no substantial reasons why the investigated concept should not be understood likewise between the groups and thus compared. Measurement equivalence is, however, also a statistical property that indicates whether we are measuring the same concept empirically. On the one hand, the conceptual requirement is to select a sample for the comparison that makes theoretical sense to compare (see Locke and Thelen 1995). On the other hand, the statistical test of measurement equivalence provides evidence that survey items used to measure the theoretical construct are interpreted similarly and measure the underlying concept equivalently in a cross-national setting (Singh, 1995). If this property remains unsatisfied, comparisons between the groups will not only reflect the actual true difference but also non-equivalence of the measurement instrument and hence likely provide biased results (Schuman and Presser 1981; Tourengau, Rips, and Rasinski 2000; Kuha and Moustaki 2015). Testing for measurement equivalence therefore enables researchers to judge if the investigated concept is comparable across the selected groups or whether they should refrain from drawing comparative conclusions without further precautions.

3.1 Testing for Measurement Equivalence with LC Models

To model labour market segmentation and to test for measurement equivalence, we use multi-group LC models (McCutcheon 1987; Clogg and Goodman 1984). Multi-group LC models are characterized by three types of variables—observed (indicator) variables, unobserved variables (latent classes), and a grouping variable (McCutcheon 2002; Kankaraš, Moors and Vermunt 2010). A multi-group LC model with a single latent variable (X), four observed indicators (A, B, C, D), and one grouping variable (G) can be expressed as:

$$\pi_{ijklts}^{ABCDX|G} = \pi_{ts}^{X|G} \pi_{its}^{A|XG} \pi_{jts}^{B|XG} \pi_{kts}^{C|XG} \pi_{lts}^{D|XG}$$

where the latent class probability ($\pi_{ts}^{X|G}$) is the probability that a randomly chosen observation in group s will be located in class t . The conditional probabilities (e.g. $\pi_{its}^{A|XG}$, ...) show the probability of observing a specified answer i to the observed indicator variable A given their membership in t and s . In the context of labour market segmentation, $\pi_{ts}^{X|G}$ denotes the size of a segment in country s , while $\pi_{its}^{A|XG}$ could, for example, indicate the probability of answering that the person has a full-time contract to the question asking ‘Do you have full-

time or part-time contract?', given that the individual is part of labour market segment t and country s . Conditional probabilities can be expressed as:

$$\pi_{its}^{A|XG} = \frac{\exp(\lambda_i^A + \lambda_{it}^{AX} + \lambda_{is}^{AG} + \lambda_{its}^{AXG})}{\sum_i \exp(\lambda_i^A + \lambda_{it}^{AX} + \lambda_{is}^{AG} + \lambda_{its}^{AXG})}$$

where lambdas (λ) are log-linear parameters, with λ_{it}^{AX} being of the highest interest, because it shows the relationship between the observed indicator and the latent variable that is modelled. The term λ_{its}^{AXG} shows how the relationship between the latent variable and its indicators differ across the groups. In our case, combination of λ_{it}^{AX} and λ_{its}^{AXG} shows the association between labour market segments and its observed indicators (working hours, contract type, supervisory responsibilities, ...) across all countries in the sample. Terms λ_i^A and λ_{is}^{AG} are used as intercept probabilities for each indicator, where λ_{is}^{AG} is a direct effect between the group and the indicator which differentiates the intercept probability by country. This term is used when the group differences in item responses cannot be fully explained by the latent variable—when the group variable influences the indicators independently of the latent variable. Figure 1 visualizes the probability of having a part-time contract for a hypothetical example of two individuals from the same segment t but from two different countries across all levels of measurement equivalence.

Fig. 1. Levels of Measurement Equivalence

In the framework of measurement equivalence, multi-group LC models are called *heterogeneity models*, because all parameters are allowed to differ across groups. In fact, it is equal to estimating a LC model in each country separately (Clogg and Goodman 1985). Concluding that a measurement model is heterogeneous means that the connection to the latent concept and the baselines differ across the countries—shown by different lengths of

lines in Figure 1. The different effect of the latent class is shown by the grey arrow ($\lambda_{it}^{AX} + \lambda_{its}^{AXG}$), while the different baseline for the indicator variable across countries is shown by the black solid line ($\lambda_i^A + \lambda_{is}^{AG}$).

In order to move to higher levels of measurement equivalence, and thus to higher comparability, some parameters of the model are constrained to be equal across the groups. The *partial homogeneity model* is estimated by constraining some or all interaction terms to zero:

$$\lambda_{its}^{AXG} = \lambda_{jts}^{BXG} = \lambda_{kts}^{CXG} = \lambda_{lts}^{DXG} = 0$$

If the model with the constraint indicates a better (or equivalent) model fit than the heterogeneity model, we consider this as evidence for partial homogeneity between the groups. This means that the relationships between the latent variable and the indicators do not differ by group, so there is no need to include the interaction term in the model. Following the example in Figure 1, this means that two individuals from two different countries who are assigned to the same labour market segment experience the same increase/decrease of the probability of having a full-time contract (shown as λ_{it}^{AX} ; the same length of the grey arrow).

Achieving the level of *partial homogeneity* improves the possibilities of drawing comparative conclusions about the sizes of labour market segments across the countries (McCutcheon and Hageaars 1997). Moreover, having established that the concept of labour market segmentation functions equivalently across the countries, the extracted segments can be associated with socio-demographic and job-related covariates of interest. This allows us to identify, for example, which social groups and demographics have a higher probability to experience insecure employment.

A further step towards equivalence is the estimation of the *structural homogeneity* model. This is done by taking the partial homogeneity model and adding the following constraint.

$$\lambda_{is}^{AG} = \lambda_{js}^{BG} = \lambda_{ks}^{CG} = \lambda_{ls}^{DG} = 0$$

In the structural homogeneity model, the interactions (λ_{its}^{AXG}) and the direct effects (λ_{is}^{AG}) between the grouping variable and the indicators are constrained to zero and hence excluded from the model. As a result, conditional probabilities of items are restricted to be equal across the groups and independent from the grouping variable, when controlled for the latent

variable. In this model, we assert that the conditional probability of having a full-time contract is not different across countries once we account for the labour market segment. In Figure 1, this can be seen by the same probability of having a part-time contract in both countries, given the observations come from the same segment (equal lengths of the grey arrow and the black line).

In the structural homogeneity model, the only aspect left to vary by group is the probability of latent class membership $\pi_{ts}^{X|G}$. If we constrain $\pi_{ts}^{X|G} = \pi_t^X$, the model becomes a *complete homogeneity model*, which represents the highest level of equivalence. The estimated model essentially boils down to a single-group LC model that can be expressed as:

$$\pi_{ijklt}^{ABCDX} = \pi_t^X \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X}$$

and asserts that latent classes are identical across all groups and the data is virtually pooled together. If this model is reached, it means that there is no statistical evidence for a difference between the countries in labour market segmentation. Not only are the segments indistinguishable across the countries with regard to their content, but also the proportions of people in particular segments are equal.

Testing for measurement equivalence with LC models is performed in a stepwise manner—at each step constraining an additional set of parameters, and evaluating the change in model fit to the observed data. If a model with higher level of measurement equivalence is estimated without significantly deteriorating the model fit, we consider this as evidence for that particular level of measurement equivalence (Clogg and Goodman 1984, 1985). LC models are conventionally evaluated by several model fit criteria: the likelihood-ratio (L^2), Pearson's chi-squared (χ^2), and information criteria (IC) indices. As the former two options have a number of important limitations (see McCutcheon 2002), we rely on the latter. Although Akaike information criterion (AIC) and Bayesian information criterion (BIC) are often used in the literature for comparing LC models, there are also more recent measures available, such as CAIC (Bozdogan 1987) and SABIC (Sclove 1987). According to a simulation study by Nylund, Asparouhov and Muthén (2007), BIC and SABIC were the most successful at identifying the correct number of classes with larger sample sizes and thus will also guide model selection in our analysis. Model selection will also be evaluated upon satisfying the assumption of local independence by inspecting bivariate residuals. In case of large residuals,

models with additional allowed residual correlations will be tested to model unaccounted local dependence (Oberski, van Kollenburg and Vermunt 2013).

Results of statistical tests of measurement equivalence can have important implications for what kind of comparisons researchers use and which conclusions can be drawn from their analyses. Yet, the approaches to deal with results of tests of measurement equivalence can vary according to the investigated research question.

3.2 What to do about measurement non-equivalence?

Most comparative studies are likely to be situated somewhere between the extremes—between the heterogeneity model and the complete homogeneity model. The former asserts complete lack of comparability, thus prevents the researcher from drawing any valid contrasts between the groups. The latter, on the contrary, asserts full correspondence between the two groups, hence nothing remains to be compared. Unless the aim is to investigate whether the groups are equal on the latent concept, complete homogeneity is of little interest for comparative purposes.

Measurement non-equivalence commonly occurs in applications; see for example Kankaraš and Moors (2009) for LC model and Meuleman and Billiet (2012) for factor analysis. Generally, there are two approaches that researchers can take to tackle questions of measurement equivalence in their models: either *allow* for differences in the measurement of the latent concept between the groups, or *ignore* the differences between the groups and proceed with a model that assumes equivalence. In labour market segmentation, allowing for measurement inequivalence would mean letting the model vary at occasions where the model fit would indicate a difference between the compared groups. Ignoring the inequivalence would mean to model the data as if the equivalence of the concept of labour market segmentation was achieved across the groups and do not differentiate the analysis by the grouping variable whatsoever.

It could be argued on the grounds of preference for model parsimoniousness that ignoring inequivalence is acceptable if it yielded equal findings as a model that accounts for the inequivalence. Nevertheless, as Kuha and Moustaki (2015) show, latent variable models can be sensitive to assumptions about measurement equivalence and ignoring possible

violations might lead to substantial bias in cross-group comparisons. Yet, accounting for measurement non-equivalence can also be problematic. Although statistically the model that allows some level of heterogeneity will fit the data better, there are conceptual issues emerging at the same time. Accounting for measurement non-equivalence means allowing one or more indicators to perform differently across the groups. In labour market segmentation, this would mean, for example, that membership in a segment would increase the probability of having a part-time contract in one country, while decreasing the probability in another. Although this association can easily be modelled, it is questionable if we would still deal with an equivalent latent concept across countries.

In confirmatory factor analytical literature, some authors argue that at least two indicators per construct should be equivalent across the groups for drawing valid comparisons (Byrne, Shavelson and Muthén 1989; Steenkamp and Baumgartner 1998). There is, however, no such guidance for LC models to apply. The most profound conceptual problem is that a well-fitting partially measurement equivalent model might simply not exist (Kuha and Moustaki 2015)—basically, there might not be two or more indicators, as advised above, to construct the conceptual link between the groups. Furthermore, even if there are two equivalent indicators but several that are inequivalent, the basis for interpretability of the measured differences is weak.

4. Correcting for Misclassification Error

Closely related to measurement of latent constructs is the issue of misclassification. Measurement error occurs as a difference between the true value of the measurement and the value obtained during the measurement process (Lessler & Kalsbeek, 1992). Misclassification error is a type of measurement error for categorical data, where the true value is in one response category of a variable, while the measurement process yields an answer in another response category. LC models are one of the methods that were developed to address the impact of measurement and classification error on survey estimates (Biemer 2010; Berezofsky, Biemer and Kalsbeek 2008).

Accordingly, LC models come along with key advantages. First, LC models do not assume that measurement is free of error and hence directly separate the systematic part of

the measurement from the error part. Second, LC models provide model-based estimates of the truth as well as estimates of false-negative and false-positive probabilities (Biemer 2004). Finally, in addition to assessing the quality of the estimates, LC analysis can assess the quality of the indicators that are being used to estimate the latent class (Biemer and Wiesen 2002; Kreuter, Yan and Tourangeau 2008).

Misclassification error becomes a salient issue after defining the LC measurement model, when the subsequent aim is to explain class membership by means of covariates—explanatory variables of interest like socio-demographic characteristics, that could provide more information about the extracted classes. The commonly taken approach is to estimate the LC model, obtain posterior class membership probabilities, classify observations into respective segments (Goodman 1974; McLachlan and Peel 2000), and use this classification as a dependent or independent variable in a statistical model. However, this approach is problematic. Regardless of the classification method, common statistical procedures underestimate the relationship between the latent classes and investigated covariates (Bolck, Croon and Hagenaars 2004). This problem generally emerges due to the misclassification between the aforementioned true and the predicted latent membership (Croon 2002). Bolck, Croon and Hagenaars (2004) and later also Vermunt (2010) developed methods to correct for this bias. Simulation studies show that proposed correction methods work well under conditions generally encountered in social sciences and the corrected estimates and their standard errors can therefore be trusted in applied research (Bakk, Tekle and Vermunt 2013).

Omitting any of these considerations when investigating labour market segmentation might lead to biased estimates of segment populations and subsequently incorrect conclusions. Although the proposed measurement of labour market segmentation is well grounded in theory, it is unlikely that it would produce a perfect measurement. Therefore, it is crucial to consider cases where a person's true labour market segment differs from the predicted one by the model. LC analysis inherently accommodates this problem by predicting each observation's segment membership as a probability. This probability is then, in return, used to correct for the uncertainty produced by the measurement model.

As a result, LC analysis not only enables developing more valid measurement models of labour market segmentation by accounting for the possible error attached to survey responses, but also more precise estimates of associations between segments and other

variables of interest. We therefore argue that conclusions about labour market segmentation obtained by the presented method are more precise and less biased in comparison with traditional approaches to measuring labour market segmentation.

5. Empirical Results

5.1 Data

In order to illustrate the proposed approach, we analyse data from the European Union Labour Force Survey (EU-LFS). EU-LFS is a large annual survey focused on the working population, conducted in 28 member states of the EU, two candidate countries, and three countries of the European Free Trade Association (EFTA). EU-LFS contains a relatively stable set of indicators in time and across countries along the three dimensions of insecurity that are used in the measurement. Extensive efforts of data harmonization at the European level come along with high quality social survey data, which is a valuable asset for the paper.

We selected 11 countries (Belgium, Germany, France, Denmark, Finland, Italy, Portugal, Poland, Slovenia, the UK and Ireland) in accordance with existing typologies of coordinated, liberal and mixed market economies (Hall and Soskice 2001) and welfare state regimes (Esping-Andersen 1990) to model labour market segmentation in 2015. First, the subgroup of Belgium, Germany, and France were selected to represent coordinated market economies and conservative welfare regimes. Second, Denmark and Finland represent the Nordic type of coordinated economies and social-democratic welfare regimes. Third, Italy and Portugal were included as cases of mixed market economies and representing the Southern type of welfare regimes. Finally, the United Kingdom and Ireland serve as examples of the liberal market economy and liberal welfare regimes. We included at least two cases per cluster to warrant not only between-, but also within-cluster comparisons. Although inclusion of more countries would seem desirable, it comes at the cost of tractability of the model, with only marginal improvements in generalizability of the findings. Moreover, some countries lack the data that we required for our models (e.g. Spain missing income deciles), thus could not be used for the analysis. As a result, this paper focuses on a sub-sample of countries reflecting the major types of political economies and welfare regimes in Europe.

The operationalization of labour market segmentation is shown in Table 1. The temporal aspect is operationalized by standard weekly working hours (full-time/part-time), the contract type (permanent/temporary), and occurrence of atypical working schedules (constructed as a sum of incidence of shift work, evening work, night work, and weekend work—each indicator having three values: 0 for “Person never does shift work”, 1 for “Person sometimes does shift work”, and 2 for “Person does shift work”. We attached “low” to no atypical work, “middle” to values of the sum between 1-3, and “high” to sum values above 3). Organizational insecurity is measured by a proxy measure of supervisory responsibility (yes/no). Finally, economic insecurity is measured by wage (low/middle/high, constructed from country-specific income deciles, lowest three deciles as “low”, four middle deciles as “middle”, and highest three deciles as “high”). Besides employment characteristics pertaining to segmentation, various job-related and personal characteristics are used as covariates in the later stages of analysis to generate a deeper understanding of the extracted segments and processes eliciting segmentation: age (“low” between 17-32/“middle” between 33-50/“high” between 51-70), gender (male/female), education level (one-digit ISCED-11: “low” as lower secondary/“middle” as upper secondary/“high” as third level education), having a second job (yes/no), sector (NACE categories), nationality (country/EU/non-EU), firm size (small/medium/large), and occupation (one-digit ISCO).

Table 1. *Operationalization of insecurity (Rodgers, 1989)*

Concept	Measurement	Levels
Temporal insecurity	Working hours	<i>Full-time</i> <i>Part-time</i>
	Contract type	<i>Permanent</i> <i>Temporary</i>
	Non-standard working hours	<i>Low</i> <i>Middle</i> <i>High</i>
Organizational insecurity	Supervisory responsibility	<i>Yes</i> <i>No</i>
Economic insecurity	Income	<i>Low</i> <i>Middle</i> <i>High</i>

5.2 Model Fit

The model selection proceeds in three steps: (1) we run a sequence of models to determine the number of latent classes by gradually increasing the number of modelled classes; (2) at each increase of the number of latent classes, we test four levels of measurement equivalence by placing respective constraints on parameters of the LC model, and the best model is selected based on BIC and SABIC; (3) we evaluate the local independence assumption by inspecting bivariate residuals and assess changes in the model by adding relevant residual correlations.

When running the models, we explicitly consider the possibility that some segments might be present in some countries and not in others. This result would be in line with the models estimated in each country separately (not presented here due to space limitations). A class would be considered as not occurring in a country, if its modelled probability in the country is close to zero. Possibly, we could constrain latent class probability parameters to zero and test the model performance by BIC and SABIC. However, it is not the aim of this paper to test specific hypotheses about the number of labour market segments.

The results reveal that regardless of the number of latent classes, the best fitting model is a partially homogenous model (see Table 2). As described above, this model is equivalent with respect to the conceptual meaning of the extracted latent classes, because indicator loadings (λ_{it}^{AX} , etc.) are equivalent across countries. However, the model differs in distributional characteristics between countries due to different intercepts on indicators and allocation to latent classes (direct effects between the group and indicators λ_{is}^{AG} , etc., and direct effect between the group and latent class probability λ_{st}^{GX}). A partially homogenous model provides evidence that the *link* (i.e. the loadings) between the indicators and the latent concept of labour market segmentation is equivalent across the countries in the sample and hence provides ground for cross-national comparisons.

A possible explanation for the necessity to include the direct effects between the grouping variable and indicators is that the countries in the sample have different baseline levels of indicators that cannot be fully explained only by the model of labour market segmentation. In other words, the model captures the relationships between the indicators, but there are idiosyncratic country differences beyond the model, so the model fit indicators suggest the inclusion of the country-to-indicator direct effect. The different distribution of

latent class membership is a conclusion following from analyses of countries separately, where we found different numbers of latent classes between countries. Inclusion of a direct effect between the group variable and the latent variable of segment membership is therefore necessary to capture the different number of segments across countries.

Table 2. Model Fit

		L²	df	Class. Err.	BIC(L²)	SABIC(L²)
5-class	Heterogeneity	793.9	352	0.19	-3168	-2049
	Partial Homogeneity	2427.5	632	0.18	-4685	-2677
	Structural Homogeneity	5431.6	702	0.13	-2469	-238
	Complete Homogeneity	12145.5	742	0.16	3795	6153
6-class	Heterogeneity	553.4	264	0.22	-2418	-1579
	Partial Homogeneity	2100.3	614	0.24	-4810	-2859
	Structural Homogeneity	4668.6	684	0.28	-3029	-856
	Complete Homogeneity	12011.8	734	0.25	3751	6084
7-class	Heterogeneity	402.8	176	0.27	-1578	-1019
	Partial Homogeneity	1817.3	596	0.26	-4890	-2996
	Structural Homogeneity	3924.1	666	0.32	-3571	-1455
	Complete Homogeneity	11950.7	726	0.36	3780	6087
8-class	Heterogeneity	321.3	88	0.28	-669	-389
	Partial Homogeneity	1515.6	578	0.29	-4989	-3152
	Structural Homogeneity	3295.8	648	0.34	-3997	-1938
	Complete Homogeneity	11914.0	718	0.36	3833	6115
9-class	Heterogeneity	206.3	0			
	Partial Homogeneity	1402.6	560	0.28	-4900	-3120
	Structural Homogeneity	2768.8	630	0.32	-4321	-2319
	Complete Homogeneity	11908.3	710	0.33	3918	6174

With regard to the number of latent classes, the 8-class solution has the best model fit, as the 9-class solution would increase both BIC and SABIC. The relative difference between a 7-class and 8-class model is small so we consider both models. We proceed by inspecting bivariate residual tables for local dependence which can reveal more information about the best fitting model (see Table 3). Bivariate residuals show the need to account for the residual correlation between non-standard working time and contract type—9.15 for 7-class model, 20.36 for 8-class model.

Table 3. Standardized bivariate residuals

		7-class PH	8-class PH	7-class PH + resid.
Working hours	Contract type	0.05	0.39	1.34
	Non-standard w.h.	0.12	0.18	0.98
	Supervisory resp.	0.03	0.14	0.09
	Income	0.52	0.08	0.08
Contract type	Non-standard w.h.	9.15	20.36	0.00
	Supervisory resp.	2.28	0.01	1.30
	Income	1.13	0.61	0.70
Non-standard w.h.	Supervisory resp.	1.22	0.38	0.18
	Income	1.64	0.11	0.79
Supervisory resp.	Income	0.46	0.48	0.22

Adding the residual correlation (model fit shown in Table 4 below), the 8-class model shows worse model fit in BIC and SABIC, indicating that the eighth class might be redundant if we account for the local dependence between indicators. As a result, the 7-class model shows the best trade-off between model fit and parsimony. Both information criteria, BIC and SABIC, highlight this model as the best fit to the data and its model fit is better than any of the previously estimated models. Moreover, inspection of bivariate residuals (back in Table 3) shows that we have satisfied the assumption of local independence by no other residuals showing extensively high values.

Table 4. Extended models with residual correlation

		L^2	df	Class. Err.	BIC(L^2)	SABIC(L^2)
7-class	Partial Homogeneity + residual correlation	1600.1	594	0.28	-5085	-3197
8-class	Partial Homogeneity + residual correlation	1525.2	576	0.31	-4957	-3127

5.3 Segmentation Profiles

The profiles of employees in each segment are shown in Table 5. Conceptually, we identify three larger encompassing groups that we labelled as standard, managerial, and disadvantaged, with varying proportions between the countries.

Table 5. Latent class profiles

	Standard		Managerial		Disadvantaged		
Size	0.26	0.15	0.15	0.14	0.13	0.15	0.03
Working hours							
Full-time	0.87	1.00	0.97	0.97	0.27	0.45	1.00
Part-time	0.13	0.00	0.03	0.03	0.73	0.55	0.00
Contract type							
Permanent	0.91	0.92	0.97	0.98	0.76	0.82	0.20
Temporary	0.09	0.08	0.03	0.02	0.24	0.18	0.80
Non-standard working hours							
Low	0.53	0.50	0.32	0.57	0.69	0.33	0.64
Middle	0.22	0.24	0.43	0.19	0.24	0.28	0.18
High	0.25	0.26	0.25	0.24	0.07	0.39	0.17
Supervisory responsibility							
Non-sup.	0.79	0.99	0.23	0.59	0.93	0.92	0.98
Supervisory	0.21	0.01	0.77	0.41	0.07	0.08	0.02
Income							
Low	0.01	0.00	0.04	0.00	0.85	1.00	0.87
Middle	0.97	0.64	0.23	0.01	0.10	0.00	0.13
High	0.01	0.36	0.73	0.99	0.04	0.00	0.00
Cluster size per country							
	Standard		Managerial		Disadvantaged		
Belgium	0.46	0.02	0.09	0.16	0.15	0.12	0.00
Germany	0.26	0.14	0.20	0.07	0.14	0.12	0.07
Denmark	0.27	0.16	0.24	0.05	0.03	0.24	0.01
Finland	0.09	0.25	0.30	0.01	0.00	0.30	0.05
France	0.41	0.00	0.09	0.23	0.16	0.08	0.03
Ireland	0.23	0.19	0.13	0.20	0.12	0.13	0.00
Italy	0.22	0.22	0.03	0.23	0.12	0.17	0.01
Poland	0.40	0.00	0.07	0.20	0.14	0.20	0.00
Portugal	0.27	0.13	0.09	0.20	0.21	0.11	0.00
Slovenia	0.13	0.26	0.06	0.24	0.08	0.24	0.00
the UK	0.09	0.29	0.29	0.03	0.10	0.20	0.01

The group of standard employees is the largest group, on average covering 40.7% of the workforce in the investigated countries. They tend to be employed on full-time, permanent contracts, with relatively low incidence of non-standard working schedules. They enjoy medium and medium/high salaries, with occasional supervisory responsibility. We identify two sub-groups of standard employees that differ mainly in two aspects—supervisory

responsibilities (higher probability in the first (21%) than the second group (1%)), and salary (very high probability for a medium income (97%) in the first group, while the second group may also fall into the higher income category (36%)). The first group is predominant in Belgium, Germany, Denmark, France, Poland, and Portugal, whereas the second group is present in Finland, Slovenia, and the UK. Ireland and Italy show an equal proportion of both subgroups.

The group of 'managerial employees' constitutes 28.9% of employees. Similar to the standard group, the managerial group has a very high probability of having a full-time and permanent contract. Unlike the standard group, 'managerial' employees are likely to have supervisory responsibilities combined with high salaries. The identified two subgroups differ in the incidence of non-standard working schedules: the first group is likely to work non-standard hours (higher occurrence of medium non-standard working hours: 43%) than the second (19%). Wage levels differ as well; the first group has a lower probability of having high income (73%) in comparison with the second group (99%). The first 'managerial' group is concentrated in Germany, Denmark, Finland, and the UK, while the second 'managerial' group can be found in Belgium, France, Ireland, Italy, Poland, Portugal, and Slovenia.

Finally, the group of 'disadvantaged' employees covers 30.4% of the investigated workforce. With varying degrees of probability, disadvantaged workers have part-time or temporary contracts, are unlikely to have any supervisory responsibilities, and are predominantly concentrated in the lowest income strata. The model identifies three distinct subgroups: The first group has the highest probability to have a part-time contract (73%), which is similar to the second group (55%), while the third group seems to always enjoy full-time contracts. The third group, however, has the highest probability to be employed on a temporary contract (80%), followed by the first group (24%) and second group (18%). The second group has the highest incidence of non-standard working hours. All groups align on the lack of supervisory responsibilities and low income. The first and second groups seem to co-occur with very similar probabilities in Belgium, Germany, Ireland, Italy, and Poland. The first group is solely present in Portugal and France, while the second occurs mostly in Denmark, Finland, Slovenia, and the UK.

Table 6: R-squared values

	R-squared
Wages	0,72
Working hours	0,42
Supervisory resp.	0,34
Contract type	0,17
Non-std. working hrs.	0,05

The results show strong divides in labour markets in wages, working hours, and supervisory responsibilities—these three categories seem to be conceptually defining the extracted segments. Contract type and incidence of non-standard working hours seem to be of lower importance. In order to quantify how much of the variance of the original indicators are segments able to explain once extracted, segments (latent classes) are used as an independent variable to predict the indicators and R-squared values are calculated. Table 6 shows R-squared values for each indicator—these range between 0-1 and are similar to the conventional interpretation of the variance explained in ANOVA¹ (Vermunt and Magidson 2005). These results confirm our observation of wages, working hours, and supervisory responsibilities are playing a major role in defining labour market segments presented in this study. Nevertheless, these are average R-squared values over all countries, which conceals differences between countries. For a more comprehensive discussion, see Doerflinger, et al. (2017).

All in all, the measurement model with seven classes, partial homogeneity, and with included residual correlation fits the model selection criteria best: It has the best model fit in terms of BIC and SABIC, is more parsimonious than the 8-class model, and also better accommodates the assumption of local independence.

¹ We use Goodman-Kruskal tau-b coefficients as R^2 statistic (see Southwood 1974). It indicates how well an indicator is explained by the latent variable, which is similar to conventional interpretations of the variance explained in analysis of variance and to item communalities in factor analysis (Vermunt and Magidson 2005).

5.5 The social configurations of labour market segments

To gain more information about the identified labour market segments, the extracted latent classes are associated with a set of covariates that were selected based on existing research: age, gender, education, nationality, existence of a second job, firm size, sector, and occupation. Although we model the association between these covariates and the segments, it is important to underline that the covariates cannot influence the measurement model or assignment of workers into respective segments. Any of the socio-demographic and economic differences that we report below emerged only due the segmentation by the five indicator variables used in the measurement model (i.e. working hours, contract type, non-standard working hours, supervisory responsibilities, and income).

Table 7: Social configurations of labour market segments (% add up to 1 in columns)

		Standard		Managerial		Disadvantaged		
	Size	0.26	0.15	0.15	0.14	0.13	0.15	0.03
Age category								
	17-32 years	0.33	0.34	0.25	0.12	0.20	0.47	0.95
	33-50 years	0.40	0.35	0.45	0.51	0.42	0.30	0.05
	51-70 years	0.27	0.31	0.31	0.37	0.38	0.22	0.00
Gender								
	Male	0.48	0.68	0.71	0.64	0.13	0.36	0.71
	Female	0.52	0.32	0.29	0.36	0.87	0.64	0.29
Education (recoded ISCED)								
	Low	0.15	0.26	0.02	0.08	0.28	0.28	0.49
	Medium	0.61	0.50	0.31	0.39	0.44	0.64	0.43
	High	0.24	0.24	0.67	0.53	0.28	0.08	0.08
Nationality								
	Country	0.96	0.84	0.93	0.99	0.91	0.85	0.88
	EU-national	0.02	0.09	0.05	0.00	0.03	0.07	0.05
	Non-EU	0.02	0.07	0.03	0.00	0.05	0.09	0.07
Existence of a second job								
	No	0.96	0.98	0.95	0.97	0.91	0.96	1.00
	Yes	0.04	0.02	0.05	0.03	0.09	0.04	0.00
Firm size								
	50+	0.47	0.52	0.67	0.60	0.29	0.27	0.48
	11-49	0.33	0.32	0.22	0.30	0.29	0.37	0.35
	1-10	0.21	0.16	0.11	0.10	0.42	0.35	0.17
Sector (recoded NACE)								
	Agricult. & Mining	0.01	0.01	0.02	0.01	0.02	0.02	0.05
	Care services	0.16	0.08	0.10	0.08	0.18	0.14	0.14
	Industry & Manuf.	0.27	0.34	0.25	0.29	0.16	0.11	0.41
	Other services	0.03	0.03	0.04	0.01	0.11	0.10	0.01
	Private services	0.34	0.39	0.43	0.33	0.26	0.63	0.26
	Public admin.	0.18	0.16	0.17	0.29	0.27	0.00	0.12

The social configurations of the identified labour market segments are shown in Table 7. Standard employees are relatively balanced by age categories and gender, and prevalently feature medium education levels. It is unlikely that employees in this group have more than one job and approximately half of the group is employed in companies with more than 50 employees. Managerial employees are concentrated in medium-to-higher age strata—above 33 years old, with a greater number of men, and predominantly high education levels. Similar

to standard employees, they are unlikely to have a second job and work mainly in large companies. The disadvantaged employees are from mixed age categories: while the first group covers people above 33 years, the second and third groups are mainly composed of the young. The gender composition varies as well. The first and second disadvantaged groups are dominated by women (87% and 64%), while the third is dominated by men (only 29% of women). Overall, disadvantaged workers have attained low or middle education levels. More than other groups, the disadvantaged workers are more likely to work in medium or small companies with less than 49 employees. Migrant workers—either EU or non-EU nationals—are concentrated mainly in two groups: second standard and second disadvantaged groups.

Finally, the distribution of segments in occupation groups (ISCO) is shown in Table 8: Managers and professionals are predominantly in the low insecurity managerial segment, while moving towards occupations that are stressed by the literature (e.g. Eichhorst and Marx 2015) as having a high risk of insecurity, we observe a rising proportion of disadvantaged workers. Specifically, avoiding the “extremes”—e.g. managers vs. elementary occupations—we observe much more diversity across the results than the occupations literature would conventionally suggest by using a single indicator approach (such as using temporary employment contract to analyse insecurity in the workforce).

Table 8: Occupations (ISCO) distribution of labour market segments (% add up to 1 in rows)

	Standard employees		Managerial employees		Disadvantaged employees		
Size	0.26	0.15	0.15	0.14	0.13	0.15	0.03
Armed Forces Occup.	0.34	0.00	0.00	0.62	0.00	0.03	0.01
Managers	0.05	0.00	0.73	0.16	0.01	0.04	0.00
Professionals	0.19	0.18	0.19	0.35	0.06	0.02	0.02
Tech. & Assoc. Prof.	0.37	0.09	0.12	0.25	0.07	0.07	0.03
Clerical Support Workers	0.36	0.18	0.03	0.11	0.15	0.13	0.04
Serv. & Sales Workers	0.27	0.12	0.04	0.03	0.06	0.47	0.01
Skilled Agricult., Forestry and Fishery Workers	0.36	0.11	0.06	0.00	0.07	0.24	0.16
Craft & Rel. Trades Work.	0.33	0.27	0.05	0.11	0.03	0.13	0.08
Plant & Mach. Oper. and Assemblers	0.32	0.31	0.01	0.12	0.04	0.16	0.04
Elementary Occupations	0.20	0.14	0.01	0.00	0.28	0.34	0.03

The presented partially equivalent measurement model of labour market segmentation allows us to draw cross-national comparative conclusions about the identified segments. We would be unable to reach these conclusions without using LC models. Importantly, the analysis shows that comparisons across countries regarding labour market segmentation without accounting for the detected measurement inequivalence will necessarily yield biased results, as it would not be possible to distinguish the indicator level effect originating either from a country or from specific segment. Moreover, association of the extracted segments without correcting for the misclassification would be inevitably underestimated and potentially not show important labour market dividing lines. On the whole, accounting for measurement inequivalence by means of the proposed model results in more precise and unbiased comparisons about labour market segmentation in Europe.

6. Discussion and conclusion

Although labour market segmentation has been a topic of high importance and interest in labour market research, measurement approaches and models that are used are seldom scrutinized (see Jackman 2008). This article proposed a methodological framework for measuring labour market segmentation in contemporary European labour markets from a comparative perspective. The approach offered here is applicable to measuring a wide range of concepts appearing in labour market research—especially, directly unobservable concepts like precarity or job quality. However, without conducting the proposed statistical tests, it is hardly possible to judge if concepts are cross-nationally comparable or not; therefore, approaches to measure segmentation would benefit from the aforementioned corrections that LC models offer. To ensure that cross-national comparisons are valid and unconfounded, comparative research requires equivalent instruments that work in all selected countries identically (Mills, van de Bunt and de Bruijn 2006).

The benefits of the latent variable approach to measure labour market segmentation are illustrated based on an empirical example using EU-LFS data. We argue that the proposed framework has three major benefits:

First, we tackled the need for increased complexity of the measurement of labour market segmentation. As demonstrated in the literature review, instead of viewing

segmentation as a binary phenomenon, we assume more fragmented configurations to exist throughout contemporary European labour markets. To measure these configurations, we operationalized the dividing lines based on insecurity, which is conceived as a major concept generating segmentation (e.g. Eichhorst et al. 2016). Our results show the expected fragmented configurations with seven classes, which can be conceptually assigned to three broader segments. Analyses of employees' socio-demographic and job-related characteristics in each segment also shed light on the social aspect of segmentation.

Second, a major benefit of our approach is that it allows for examining if the concept of labour market segmentation is comparable across countries. We conclude that by using the proposed LC model and accounting for distinctive country differences, the concept of labour market segmentation is comparable across the investigated countries. For example, we observe that some of the uncovered segments occur in a number of countries yet not in others. Such nuances in connection with the flexibility of LC models offer possibilities to test specific hypotheses about institutional multilevel effects on formation of labour market segments in Europe (see Biegert 2014).

Third, we tackled the issue of misclassification that commonly occurs in social survey research. Labour market segmentation has often been measured by means of proxy variables like employment contract. This approach implicitly assumes no misclassifications. In other words, traditional multivariate statistical tools are not built to tackle instances where the worker is classified to one segment, while in reality being a member of another segment. LC models relax this assumption and directly quantify the misclassification. Moreover, available correction procedures (Vermunt 2010) confront the problem of underestimation of associations with other variables. By definition (Bolck, Croon and Hagenaars 2004), we argue that the LC model offers stronger conclusions than traditional tools for analysis of labour market segmentation.

Although the main focus of this paper is on measurement, it also contributes to current debates on labour market segmentation. LC analysis is a useful tool for overcoming the problem of using pre-defined categories when studying segmentation. Instead of measuring assumed segmentation lines, LC analysis allows for empirical identification of segmentation lines and shows which indicators contribute the most. Furthermore, it allows for investigating the cross-national comparability of segments—an issue which is hardly scrutinised in previous

research. Our approach uses a multidimensional concept of (objective) insecurity to measure labour market segmentation. It allows for an in-depth analysis of the phenomenon, identifying segments which differ on their levels of insecurity. Specifically, insecurity grounded in the terms and conditions of employment relationships is used as the starting point for measuring and uncovering segmentation lines. Based on this, external variables such as occupations or education levels can be used to analyse the characteristics of the members of particular labour market segments. Therefore, our methodology allows us to distinguish insecurity variations within occupational or educational groups, which is not possible with traditional single-indicator approaches.

Finally, the proposed approach allows for a quantitative investigation of comparability across different countries and thus helps avoiding comparisons between apples and oranges (Locke and Thelen 1995). With regard to labour market segmentation, this is of particular importance due to the fact that it is supposedly contingent upon national institutional frameworks. Therefore, following the logic of contextualised comparisons (Locke and Thelen 1995), the proposed framework allows for investigating varieties of segmentation while judging to what extent particular segmentation lines are equivalent across countries. This provides useful guidance in interpreting research results as the influence of institutions across countries can be measured, which may further facilitate contextualisation.

Bibliography

- Bakk, Z., Tekle, F. B. and Vermunt, J. K. (2013). Estimating the Association between Latent Class Membership and External Variables Using Bias-adjusted Three-step Approach. *Sociological Methodology* 43(1), 272-311.
- Berzofsky, M., Biemer, P. P. and Kalsbeek, W. (2008). A Brief History of Classification Error Models. American Statistical Association, Joint Statistical Meetings, Denver, CO, August 3-7, 2008. Available at <https://www2.amstat.org/sections/srms/Proceedings/y2008/Files/302416.pdf>
- Biegert, T. (2014). On the outside looking in? Transitions out of non-employment in the United Kingdom and Germany. *Journal of European Social Policy* 24(1), 3-18.
- Biemer, P. (2004). Simple response variance: then and now. *Journal of Official Statistics* 20, 417-439.
- Biemer, P. and Wiesen, C. (2002). Latent class analysis of embedded repeated measurements: An application to the National Household Survey on Drug Abuse. *Journal of the Royal Statistical Society: Series A* 165(1), 97-119.
- Biemer, P. P. (2011). *Latent Class Analysis of Survey Error*. Hoboken, NJ: John Wiley & Sons.
- Billiet J. (2003). Cross-cultural equivalence with structural equation modeling. In J. A. Harkness, F. J. R. van de Vijver, & P. P. Mohler, *Cross-Cultural Survey Methods* (pp. 247-264), Hoboken, NY: John Wiley.
- Blossfeld, H.-P., Buchholz, S., Bukodi, E. and Kurz, K. (Eds.) (2008). *Young Workers, Globalization and the Labour Market: Comparing early working life in eleven countries*. Cheltenham, UK: Edward Elgar.
- Bolck, A., Croon, M. and Hagenaars, J. (2004). Estimating Latent Structure Models with Categorical Variables: One-Step Versus Three-Step Estimators. *Political Analysis* 12(1), 3-27.
- Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika* 52, 345-370.
- Byrne B. M., Shavelson R.J. and Muthén B. (1989). Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance. *Psychological Bulletin* 105, 456-66.
- Clasen, J. and Clegg, D. (2011). *Regulating the risk of unemployment: National adaptations to post-industrial labour markets in Europe*. Oxford: Oxford University Press.
- Clogg, C. C. and Goodman, L. A. (1984). Latent structure analysis of a set of multidimensional contingency tables. *Journal of the American Statistical Association* 79, 762-771.
- Clogg, C. C., and Goodman, L. A. (1985). Simultaneous latent structure analysis in several groups. In N. B. Tuma (Ed.), *Sociological Methodology* (pp. 81-110), San Francisco: Jossey-Bass.
- Croon, M. (2002). Using Predicted Latent Scores in General Latent Structure Models. In G. A. Marcoulides, & I. Moustaki (Eds.), *Latent Variable and Latent Structure Models* (pp. 195-223), Mahwah, NJ: Erlbaum Associates.

- Chung, H., & Mau, S. (2014). Subjective insecurity and the role of institutions. *Journal of European Social Policy* 24(4), 303–318.
- Chung, H., & van Oorschot, W. (2011). Institutions versus market forces: Explaining the employment insecurity of European individuals during (the beginning of) the financial crisis. *Journal of European Social Policy* 21(4), 287-301.
- Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., and Billiet, J. (2014). Measurement Equivalence in Cross-National Research. *Annual Review of Sociology* 40, 55-75.
- Davidsson, J. and Naczyk, M. (2009). The Ins and Outs of Dualisation: A Literature Review. *REC-WP Working Paper No. 02/2009*. Available at <https://ssrn.com/abstract=1489906>
- Davoine L., Erhel C. and Guergoat-Larivière M. (2008). *A Taxonomy of European Labour Markets Using Quality Indicators*. Université Paris1 Panthéon-Sorbonne, HAL. Available at <http://ideas.repec.org/p/hal/cesptp/halshs-00317280.html>
- Doerflinger, N., Pulignano, V. and Lukac, M. (2017). Varieties of labour market insecurity: social configurations of divides in Europe during the crisis. Annual Conference of the International Working Party on Labour Market Segmentation, 13-15 September 2017, Manchester.
- Doeringer, P.B. and Piore, M. J. (1971). *Internal Labor Markets and Manpower Analysis*. New York and London: M. E. Sharpe.
- Drobnič, S., Beham, B., and Präg, P. (2010). Good job, good life? Working conditions and quality of life in Europe. *Social Indicators Research* 99(2), 205-225.
- Eichhorst, W., and Marx, P. (Eds.) (2015). *Non-Standard Employment in Post-Industrial Labour Markets: An Occupational Perspective*. Cheltenham, UK: Edward Elgar Publishing.
- Eichhorst, W., Marx, P., and Wehner, C. (2016). Labor Market Reforms in Europe: Toward more flexicure labor markets? IZA Discussion Paper No. 9863, Institute for the Study of Labor, Bonn, Germany.
- Eid, M., Langeheine, R. and Diener, E. (2003). Comparing typological structures across cultures by multigroup latent class analysis: A primer. *Journal of Cross-Cultural Psychology* 34(2), 195-210.
- Emmenegger, P., Häusermann, S., Palier, B., and Seeleib-Keiser, M. (Eds.) (2012). *The Age of Dualization: The Changing Face of Inequality in Deindustrializing Societies*. Oxford: Oxford University Press.
- Esping-Andersen, G. (1990). *The three worlds of welfare capitalism*. Cambridge, UK: Polity Press.
- Goldthorpe, J. H. (1984). The End of Convergence: Corporatist and Dualist Tendencies in Modern Western Societies. In J. H. Goldthorpe (Ed.), *Order and Conflict in Contemporary Capitalism* (pp. 315-344), Oxford: Clarendon Press.
- Goodman, L. A. (1974). The Analysis of Systems of Qualitative Variables When Some of the Variables Are Unobservable. Part I: A Modified Latent Structure Approach. *American Journal of Sociology* 79, 1179-1259.

- Goodman, L. A. (2002). Estimating Latent Structure Models with Categorical Variables: One-Step Versus Three-Step Estimators. In J. A. Hagenaars, & A. L. McCutcheon (Eds.), *Applied Latent Class Analysis* (pp. 3-55), Cambridge, UK: Cambridge University Press.
- Green, F. (2009). Subjective employment insecurity around the world. *Cambridge Journal of Regions, Economy and Society* 2(3), 343–363.
- Hagenaars, J. A. and McCutcheon, A. L. (Eds.) (2002). *Applied Latent Class Analysis*. Cambridge, UK: Cambridge University Press.
- Hall, P. A. and Soskice, D. (2001). *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford: Oxford University Press.
- Hassel, A. (2004). The governance of the employment–welfare relationship in Britain and Germany. In B. Ebbinghaus, & P. Manow (Eds.), *Comparing Welfare Capitalism: Social policy and political economy in Europe, Japan and the USA* (pp. 146-168), London and New York: Routledge.
- Häusermann, S., Schwander, H. (2012). Varieties of Dualization? Labor Market Segmentation and Insider-Outsider Divides Across Regimes. In P. Emmenegger, S. Häusermann, B. Palier, & M. Seeleib-Kaiser (Eds.), *The Age of Dualization* (pp. 27-51), Oxford: Oxford University Press.
- Huws, U. (2014). *Labor in the global digital economy: The Cybertariat Comes of Age*. New York: Monthly Review Press.
- Jackman, S. (2008). Measurement. In J. M. Box-Steffensmeier, H. E. Brady, & D. Collier, Oxford *Handbook of Political Methodology* (pp. 119–151), Oxford: Oxford University Press.
- Kalleberg, A. (2003). Flexible firms and labor market segmentation: Effects of workplace restructuring on jobs and workers. *Work & Occupations* 30(2), 154-175.
- Kankaraš, M. and Moors, G. (2009). Measurement Equivalence in Solidarity Attitudes in Europe: Insights from a Multiple-Group Latent-Class Factor Approach. *International Sociology*, 24(4), 557-579.
- Kankaraš, M., Moors, G. and Vermunt, J. (2010). Testing for measurement invariance with latent class analysis. In E. Davidov, P. Schmidt, & J. Billiet (Eds.), *Cross-cultural analysis: Methods and applications* (pp. 359-384), New York, NY: Routledge.
- Kreuter, F., Yan, T. and Tourangeau, R. (2008). Good item or bad—Can latent class analysis tell? The utility of latent class analysis for the evaluation of survey questions. *Journal of the Royal Statistical Society: Series A* 171(3), 723-738.
- Kuha, J. and Moustaki, I. (2015). Nonequivalence of Measurement in Latent Variable Modeling of Multigroup Data: A Sensitivity Analysis. *Psychological Methods* 20(4), 523-36.
- Lessler, J. T. and Kalsbeek, W. D. (1992). *Nonsampling error in surveys*. New York: Wiley.
- Locke, R. M. and Thelen, K. (1995). Apples and Oranges Revisited: Contextualized Comparisons and the Study of Comparative Labor Politics. *Politics & Society* 23(3), 337-367.
- McCutcheon, A. L. (1987). *Latent Class Analysis*. London: Sage Publications.
- McCutcheon, A. L. (2002). Basic Concepts and Procedures in Single- and Multiple-Group Latent Class Analysis. In J. A. Hagenaars, & A. L. McCutcheon (Eds.), *Applied Latent Class Analysis* (pp. 3-55), Cambridge, UK: Cambridge University Press.

- McCutcheon, A. L. and Hagenars, J. A. (1997). Simultaneous latent class models for comparative social research. In R. Langeheine, & J. Rost (Eds.), *Applications of Latent Trait and Latent Class Models* (pp. 266–77). New York: Waxman.
- McLachlan, G. and Peel, D. (2000). *Finite Mixture Models*. NY: John Wiley & Sons.
- Meuleman, B. and Billiet, J. (2012). Measuring Attitudes toward Immigration in Europe: The Cross-Cultural Validity of the ESS Immigration Scales. *Ask: Research and Methods* 21(1), 5-29.
- Mills, M., van de Bunt, G. G. and de Bruijn, J. (2006). Comparative Research: Persistent Problems and Promising Solutions. *International Sociology* 21(5), 619-631.
- Munoz-Bustillo, R., Fernández-Macías, E., Antón, J. I., Esteve, F. (2009). *Indicators of job quality in the European Union*, European Parliament, Brussels. Available from <http://www.europarl.europa.eu/document/activities/cont/201107/20110718ATT24284/20110718ATT24284EN.pdf>
- Nylund, K.L, Asparouhov, T. and Muthén, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling* 14(4), 535-569.
- Oberski, D., van Kollenburg, G., and Vermunt, J. (2013). A Monte Carlo evaluation of three methods to detect local dependence in binary data latent class models. *Advances in Data Analysis and Classification* 7(3), 267-279.
- Palier, B. and Thelen, K. (2000). Institutionalizing Dualism: Complementarities and Change in France and Germany. *Politics & Society* 38(1), 119-148.
- Piore, M. (1975). Notes for a theory of Labor Market Stratification. In R. Edwards, M. Reich, & D. Gordon (Eds.), *Labor Market Segmentation* (pp. 215-150), Lexington, Mass.: D. C. Heath.
- R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at <http://www.R-project.org/>.
- Rodgers, G. (1989). Precarious work in Western Europe: The state of the debate. In G. Rodgers, & J. Rodgers, *Precarious Jobs in Labour Market Regulation: The Growth of Atypical Employment in Western Europe*, International Labour Organization, International Institute for Labour Studies.
- RStudio Team (2015). *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA Available at <http://www.rstudio.com/>.
- Schuman, H. and Presser, S. (1981). *Questions & answers in attitude surveys*. New York, NY: Academic Press.
- Sclove, L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika* 52, 333-343.
- Singh, J. (1995). Measurement Issues in Cross-National Research. *Journal of International Business Studies* 26(3), 597-619.
- Southwood, K. E. (1974). Goodman and Kruskal's Tau-b as Correlation Ratio: Some Implications. *Sociological Methods & Research* 3(1), 82-110.
- Steenkamp J., and Baumgartner H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research* 25, 78–90.

- Tourangeau, R., Rips, L. J. and Rasinski, K. (2000). *The psychology of survey response*. New York, NY: Cambridge University Press.
- Van Aerden, K., Moors, G., Levecque, K., & Vanroelen, C. (2014). Measuring employment arrangements in the European labour force: a typological approach. *Social Indicators Research* 116(3), 771-791.
- Vandenbrande, T., Vandekerckhove, S., Vendramin, P., Valenduc, G., Huys, R., Hansez, I., et al (2012). *Quality of work and employment in Belgium*. Dublin: EUROFOUND.
- Vermunt, J. and Magidson, J. (2005). *Latent GOLD 4.0 User's Guide*. Belmont, Massachusetts: Statistical Innovations Inc.
- Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis* 18, 450-469.
- Vosko, L. F. (2006). *Precarious Employment: Understanding Labour Market Insecurity in Canada*. Montreal: McGill-Queen's University Press.
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Available at <http://ggplot2.org>
- Wickham, H. and Francois, R. (2016). *dplyr: A Grammar of Data Manipulation*. R package version 0.5.0. Available at <https://CRAN.R-project.org/package=dplyr>
- Yoon, Y. and Chung, H. (2015). New Forms of Dualization? Labour Market Segmentation Patterns in the UK from the Late 90s Until the Post-crisis in the Late 2000s. *Social Indicators Research* 128(2), 609-631.