DISCUSSION PAPER SERIES

Varying Youth Cohort Effects on Regional Labour Market Outcomes in Germany

Thomas de Graaff
Ceren Ozgen
Duncan Roth

![IZA Institute of Labor Economics — Initiated by Deutsche Post Foundation]

Varying Youth Cohort Effects on Regional Labour Market Outcomes in Germany

Thomas de Graaff
Vrije Universiteit Amsterdam and Tinbergen Institute

Ceren Ozgen
University of Birmingham, IRiS and IZA

Duncan Roth
IAB Institute for Employment Research

ABSTRACT

Varying Youth Cohort Effects on Regional Labour Market Outcomes in Germany*

We assess how changes in youth cohort sizes effect employment rates in German labour market regions. Replicating the conventional approach, we estimate that a percentage increase in the youth share reduces regional employment rates by −0.2%. We challenge the assumption that cohort size effects are homogenous across space and find robust evidence that the negative effect of youth cohort size is more pronounced in the labour markets of metropolitan regions. These results suggest an upward pressure on urban regional employment rates as a result of the projected decrease in the size of the German youth share.

Corresponding author:
Thomas de Graaff
Department of Spatial Economics
Vrije Universiteit Amsterdam
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
E-mail: t.de.graaff@vu.nl

1. Introduction

Employment and unemployment outcomes vary considerably and persistently across regions. Better understanding the reasons for these disparities has been the aim of much research. Overman and Puga (2002) and Niebuhr (2003) show that European regions with similar unemployment rates form clusters that can also extend across national borders. Zeilstra and Elhorst (2014) stress the relevance of national institutions as well as socioeconomic conditions at the regional level, while Rios (2017) argues that the relative importance of these factors has changed in the wake of the Great Recession. Likewise, regional variation also exists with respect to employment outcomes (Martin and Tyler, 2000). Saito and Wu (2016) find that sectoral diversity and human capital are the most relevant factors for regional employment growth in the US, whereas Di Cataldo and Rodríguez-Pose (2017) argue that the importance of different factors depends on the conditions prevailing in a region. Finally, Holl (2018) provides evidence that regional employment develops differentially depending on the degree of urbanisation.

This paper ties into the subject of disparities in regional labour market outcomes by empirically assessing how a region's employment rate is affected by changes in its age structure and, in particular, by addressing the question of spatial heterogeneity in this relationship. In doing so, we build on contributions to the cohort size literature which have shown that differences in the size of age cohorts have profound effects on labour market outcomes such as wages (Mosca, 2009; Brunello, 2010; Moffat and Roth, 2016), (un-)employment (Korenman and Neumark, 2000; Biagi and Lucifora, 2008) and educational choice (Fertig et al., 2009). Specifically, we adopt the approach of Shimer (2001) and employ the share of young age groups in the working-age population as a measure of the regional age structure. Using data from German labour market regions from the period 2001–2010, we estimate the effect that this variable has on a region's

employment rate. In contrast to previous studies, we are able to include East Germany in the analysis which allows us to account for the different demographic developments of both parts of the country.

While the extant literature (Shimer, 2001; Skans, 2005; Garloff et al., 2013) has implicitly assumed that the effect of the youth share is constant across regions, we adopt a novel estimation strategy in the form of finite mixture modelling (FMM) which allows for regional heterogeneity in the effect of the youth share on the employment rate. An attractive feature of this approach is that it does not impose an ex-ante grouping of regions based on observable characteristics but endogenously assigns regions to clusters within which the estimated effect of the youth share is similar. An analysis of this type is relevant because it provides a basis for understanding how projected changes in age structures will, ceteris paribus, affect regional employment levels and hence regional inequalities in this outcome.

We argue that a change in the youth share shifts the labour supply curve for two reasons. Since labour market participation is lower among youths than among older age groups, a change in the regional age composition has a direct effect on total labour supply. Moreover, there may also be an indirect effect if a change in the youth share affects participation among youths as well as among other age groups. The effect that a shift in the labour supply curve has on employment will depend on two factors. First, under perfect competition the labour demand elasticity will determine to what extent there will be an adjustment of employment or wages. The more elastic the labour demand curve, the larger will be the effect on employment vis-à-vis wages. Second, the change in employment will also depend on the size of the labour supply shift that is brought about by a change in the youth share.

These factors are likely to differ between regions and, in particular, between urban and

non-urban areas. Urban labour markets work differently, as evidenced by a greater probability of finding employment (Di Addario, 2011) and a better matching efficiency (Glaeser and Mare, 2001; Duranton and Puga, 2004; Wheeler, 2008). More specifically, empirical evidence suggests that, ceteris paribus, an increase in the youth share should have a larger negative impact on regional employment in urban areas. First, Maiti and Indra (2016) provide evidence that the elasticity of labour demand at the regional level varies with the sector structure and Herwartz and Niebuhr (2017) use data from European regions to show that the labour demand elasticity increases in magnitude with population density. Second, urban areas provide more opportunities for non-employment such as enrollment in education (Newbold and Brown, 2015) which could imply that an increase in the regional youth share will lead to a larger inward shift of the labour supply curve in urban areas.

To the best of our knowledge, the hypothesis that changes in the share of youths in a region have a different effect on the employment rate of urban and non-urban areas has so far not been explored. Our empirical strategy is well suited to assess this hypothesis without having to first separate the data according to the degree of urbanisation. Instead, we would expect the grouping produced by the FMM approach to reflect the distinction between urban and non-urban areas if the estimated relationship indeed differed in the described way.

Based on the empirical approach of Shimer (2001), we start by estimating a model in which the effect of youth cohort size is homogenous across space. Our results suggest that an increase in the relative size of the regional youth age group leads to a significant reduction in the overall employment rate. Unlike Shimer (2001), we therefore do not find any evidence for the size of the youth share having a beneficial effect on labour market outcomes. Furthermore, we show that this result does not merely represent a change in composition caused by the fact that young individuals typically have a lower probability

of being employed than older age groups. With respect to regional effect heterogeneity, the FMM analysis identifies two clusters that can be separated into large urban regions on the one hand and all other regions on the other. We find robust evidence that the negative effect of youth cohort size is larger in the cluster of urban regions. This finding is consistent with existing evidence that the labour demand elasticity is larger in urban regions and that a greater availability of educational options in urban areas affects leads to a differently sized shift in labour supply.

The remainder of the paper is structured as follows: Section 2 provides a description of the data, while Section 3 discusses the empirical model including the FMM framework. The results are presented in Section 4 and Section 5 concludes.

2. Data and variables

Our empirical analysis uses employment and population data from the Federal Employment Agency and the German Statistical Office, respectively. The data span the period 2001–2010 and contain information on the size of the population as well as the number of employed. Both variables can be further disaggregated along various dimensions (age, sex, nationality). Since we do not have comparably disaggregated unemployment data at the regional level, our empirical analysis focuses on employment outcomes.[1]

The cross-sectional units of the analysis are the 141 labour-market regions defined by Kosfeld and Werner (2012). These regions contain one or more administrative entities at the third level of the Nomenclature of Territorial Units in Statistics (NUTS) which are combined based on the commuting flows between these entities. Since they take account

---

[1]A census was conducted in the year 2011 which led to a structural break in the population data. We therefore limit the analysis to the period before the census.

of commuting, we treat these regions as separate labour markets within which the effects of changes in the size of the youth population on the labour market can be assessed.

The labour-market outcomes of interest in this paper are the regional employment rates. These are defined as the number of employed individuals divided by the population aged between 18 and 64 in a labour-market region. As a measure of the size of the youth cohort we use the number of individuals in the age group 18 to 24 relative to the population aged 18 to 64 which closely resembles the youth share used by Shimer (2001). Reflecting cyclical variation, the means of the employment rate revolve around values of about 50%, during the sample period. In addition, the rates display substantial cross-sectional variation within a given year, with the employment rate ranging from roughly 35% to 65%. The regional youth share has a mean value of about 13% and also varies considerably across labour-market regions.
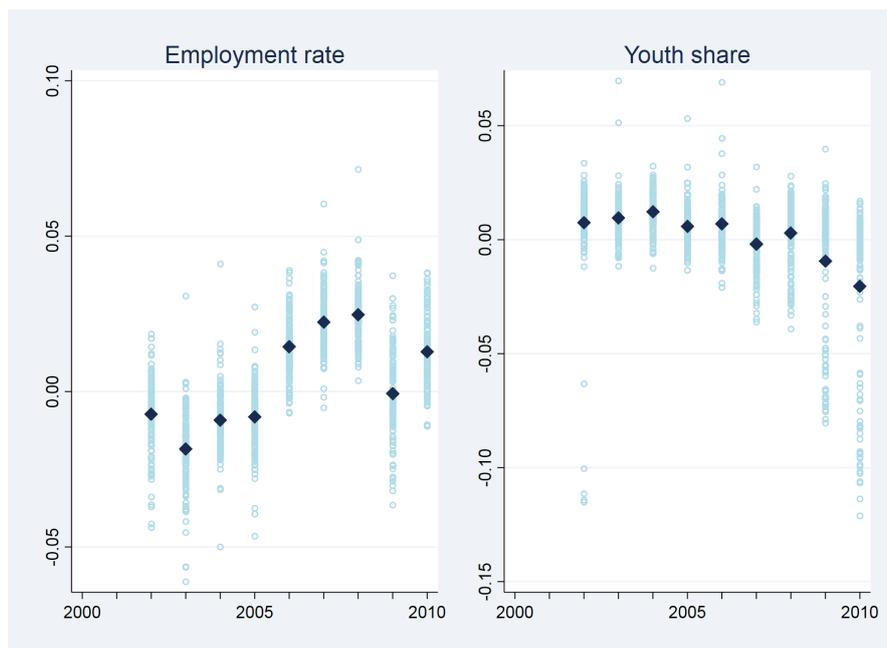
Figure 1 – Variation in the growth rates

To account for unobserved regional heterogeneity in the levels, we transfer the variables into growth rates. The cyclical development of the employment rate can be seen in Figure

7

1. The rates of growth were on average negative for the first half of the sample period before increasing substantially in 2006 and remaining high with the exception of a drop in 2009 as a result of the Great Recession. The regional youth share, in contrast, initially displayed average year-to-year growth, but started falling towards the end of the sample period with particularly large decreases to be found in East German regions. Finally, Figure 2 illustrates the residual variation that remains after regressing each variable's growth rate on a set of year dummies. Overall, the cyclical variation in the growth rates appears to be well accounted for through the use of annual fixed effects.
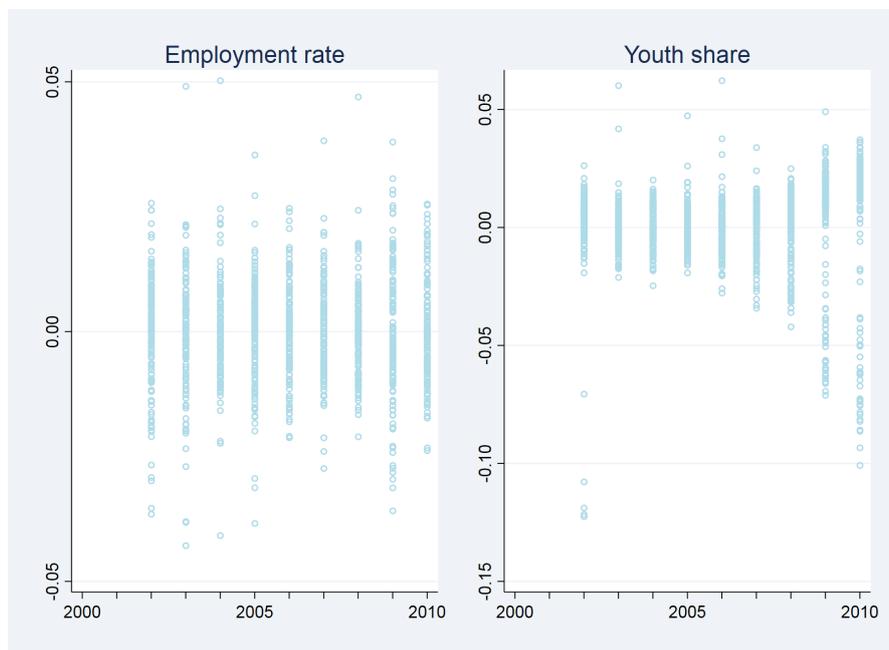
Figure 2 – Residual variation in the growth rates

Note: Observations represent the residuals from a regression of the growth rates of the respective variables on a set of year dummies.

In order to control for region-year-specific labour demand shocks, we add the variable l_{rt} to the model which is defined as the weighted sum of employment across all industries k in region r in year $t-1$. The weights are given by the ratio of industry-specific employment

at the national level in the years t and $t-1$:

$$l_{rt} = \ln\left[\sum_k \frac{E_{k,t}}{E_{k,t-1}} E_{rk,t-1}\right]. \tag{1}$$

This variable, which is often referred to as a Bartik index (see Bartik, 1991), is a measure of the predicted level of employment in region r that would have been realised if each regional industry had displayed the same rate of employment growth as its equivalent at the national level. Descriptive statistics of the main variables of interested used in the empirical analysis are shown in Table 1.

Table 1 – Descriptive statistics

Variable	Obs.	Mean	St. Dev.	Min	Max
Youth Share	1,410	0.133	0.009	0.108	0.173
Employment rate	1,410	0.496	0.061	0.318	0.657
log(Bartik's index	1,410	11.724	0.914	9.632	14.081

3. Model and identification strategy

3.1. Baseline model

To estimate the effect of the youth share on the overall employment rate, we specify the following model in which subscripts r and t denote labour-market regions and years, respectively:

$$p_{rt} = \beta y_{rt} + \gamma l_{rt} + \nu_r + \mu_t + \epsilon_{rt}, \tag{2}$$

In this specification the dependent variable p_{rt} is the natural logarithm of the employment rate, while the variable y_{rt} represents the logarithm of the youth share and l_{rt} the Bartik index. The terms ν_r and μ_t are region and year fixed effects which account for unobserved

time-invariant regional variation as well as common year-specific shocks and ϵ_{rt} is an *i.i.d.* error term.

To tackle possible autocorrelation, Equation 2 is estimated in first differences which removes unobserved time-invariant regional heterogeneity and implies that the parameters of interested are identified from the variation in the annual growth rates of the variables:

$$\Delta p_{rt} = \beta \Delta y_{rt} + \gamma \Delta l_{rt} + \Delta \mu_t + \Delta \epsilon_{rt}, \tag{3}$$

Estimation of Equation 3 is complicated by a potential endogeneity problem. Should individuals systematically migrate in response to a region's contemporaneous labour-market conditions (for example, by migrating into regions that offer better employment opportunities), ordinary least squares estimation (OLS) would produce biased and inconsistent coefficient estimates. To ensure consistent estimation we employ an identification strategy that is based on instrumental variable (IV) estimation. We follow the extant literature (Korenman and Neumark, 2000; Shimer, 2001; Garloff et al., 2013; Moffat and Roth, 2016) and use the size of the youth age group at an earlier point of time as an instrument for its current size. Specifically, we instrument the share of the age group 18 to 24 in the population aged 18 to 64 that is observed in region r at time t by the size of the age group 3 to 9 in the population aged 3 to 49 in region r 15 years previously.

In the absence of large-scale migration, a cohort that has been relatively large in the past can be expected to continue to be large in the present. The strength of the association between the lagged and the current size of the youth share (as evidenced by the first-stage F-statistics) makes this variable an attractive instrument. For the instrument to be invalid past shocks to a region's employment rate (which might have induced migration by the parental generation of the current youth age group) would have to be correlated with current shocks. We argue that this is unlikely in light of a 15-year lag and that the

regional fixed effects already accounts for the fact that certain regions realise permanently higher employment rates than others.

Equation 3 estimates an aggregate effect of youth share on employment rate. However, the effects of a change in the youth share can be different across the regions because local labor markets might function structurally different. A possible approach to incorporating these differences in the empirical analysis would be to further differentiate the youth-share variable along potentially relevant characteristics (e.g., education) and to include a full set of youth-share variables in the model. We abstain from taking this approach for two reasons. First, a model including differently defined youth shares would require a separate instrument for each of these variables in order to be identified. Given the number of valid instruments required for properly estimating these models, such an approach would not be feasible. Second, differentiating the youth share variable already requires making an assumption concerning the dimensions that are relevant for regional heterogeneity. Rather than imposing a priori assumptions, we employ an estimation strategy—i.e., latent class modelling—that is able to uncover groups of regions within which the effects of changes in the youth share are reasonably similar. In a final step, we assess whether these groups differ with respect to a variety of characteristics.

3.2. Regional heterogeneity: Hypotheses

We argue that an increase in the size of the youth population leads to a shift of the labour supply curve. This hypothesis is motivated by the resulting change in the age composition of the population. Labour force attachment is lower among youths than among older age groups in the working-age population. Turning to the employment rate, the former can be expressed as a weighted average of age-specific employment rates. An

increase in the youth share shifts weight to younger age groups whose employment rate is typically lower, thereby putting downward pressure on the overall rate. In addition to the compositional effect, empirical evidence suggests that changes in the youth share also affect the employment rates of the own and of other age groups (Shimer, 2001; Skans, 2005; Garloff et al., 2013). The negative impact on the employment rate would be reinforced if an increase in the youth cohort also reduced the employment rate of youths and of older age groups. We assess the existence of such an indirect effect in the empirical analysis.[2] Unless the indirect effects outweigh the compositional effect, we expect that an increase in the size of the youth share leads to an inward shift of the labour supply curve and to a lower overall employment rate.

Two factors are central to determining the magnitude of the employment effect. First, the labour demand elasticity will affect the extent to which a shift in the labour supply curve will be reflected in an adjustment of employment as opposed to wages. With a more elastic demand curve the effect on employment will be more pronounced relative to the change in wages. Second, the extent of the shift in the labour supply curve brought about by a change in the youth share will be influenced by the existence of non-employment opportunities such as enrolment in education. In our view both of these factors point towards a more negative employment effect in urban areas. Herwartz and Niebuhr (2017) show that labour demand becomes more elastic in denser areas. Moreover, labour demand tends to be more elastic when close substitutes are available which is arguably more likely in urban regions that offer a larger pool of workers. Finally, the existence of an urban-wage premium (Glaeser and Mare, 2001; Duranton and Puga, 2004) suggests that the share of labour costs will be higher in urban areas for a given firm compared to non-urban regions which should also make the labour-demand curve more elastic. Urban

---

[2]In principle, the size of the youth cohort can also have an effect on the demand for labour. For example, youths might demand different types of goods and services than older age groups. However, we believe that the inclusion of the Bartik variable in the empirical specification controls for shifts in the labour demand curve and we therefore focus on the effect that changes in the youth share have on the supply of labour.

regions also provide easier access to education implying a larger inward shift of the labour supply curve.

3.3. Regional heterogeneity: Latent Class Estimation Strategy

To empirically assess the existence of spatial effect heterogeneity, we employ a latent class estimation in the form of a finite mixture model (FMM) which produces an endogenous grouping of regions into clusters within which the effects of the youth share variable are similar. FMM constitutes a flexible and powerful probabilistic modelling tool which provides convenient heuristics to derive a number of homogenous subgroups from a heterogeneous population (Wedel and Kamakura, 2012).

While this segmentation can be achieved based on observable characteristics, most a priori or post hoc clustering methods remain relatively normative. The advantage of mixture models is that it does not impose a distribution or grouping of observations a priori while being able to achieve a segmentation along various dimensions. In our case the segments are defined on the basis of heterogeneity in the responses of the regional employment rate. This procedure then yields a finite number of homogenous subsets where each subset is accompanied by its own regression results and where each region is assigned a specific probability of belonging to each of these subsets.

More specifically, finite mixture model provides further insight in form of a typical statistic $y = f(y|\theta)$ which depends on unknown parameters θ. Following the notation of Wedel and Kamakura (2012), the population is assumed to be a mixture of S segments, in proportions $\pi_1, \ldots, \pi_S$. It is, however, not known ex-ante how the sample is distributed over the segments. Moreover, the probabilities π_s should meet the constraint, $\sum_{s=1}^{S} \pi_s = 1$, $\pi_s \geq 0$,

$s = 1, \ldots S$. Each segment s then has a uniquely estimated $y_s = f(y|\theta_s)$. Subsequently, if the segments were known a priori the conditional distribution of y_n would then be obtained by $f(y_n|\phi) = \sum_{s=1}^{S} f_s(y_n|\theta_s)$, where $\phi = (\pi, \theta)$. A conventional approach to estimate such a model is to apply the two stage Expectation Maximisation (EM) procedure by Dempster et al. (1977), where in the first stage the probabilities π_s are estimated and in the second stage the parameters θ_s. This is done in an iterative way until convergence is achieved.

3.4. Preview of results

A first glance at the results from estimating the model of Equation 3 within an instrumental-variables framework displays that an increase in the relative size of a region's youth population is associated with a decrease in the regional employment rate.

As the aim of this paper is to assess the existence of regional heterogeneity in the effects that changes in the size of the youth population have on regional labour-market outcomes, Figure 3 offers a first insight into how this effect differs in highly urbanised areas. This is done by showing the identifying variation and the associated linear fit separately for those ten labour-market regions containing the cities with the largest population in the year 2010 and for all remaining regions. The slightly steeper linear fit for the city regions provides first evidence that the negative effect of a change in the youth share on the employment rate is more pronounced in large cities. We provide a more detailed discussion of the baseline results and the existence of regional heterogeneity in the following section.
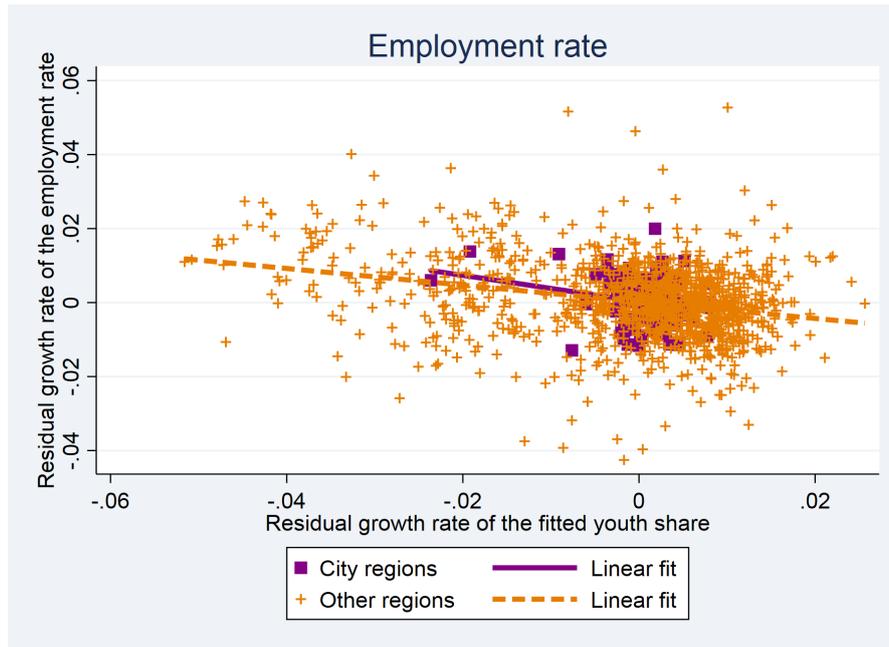
Figure 3 – Identifying variation by type of region

Notes: The residual growth rate of the employment rate corresponds to the residuals obtained from a regression of the growth in the employment rate on year fixed effects and the growth rate of the Bartik index weighted by each region's share of national population. The residual growth rate of the fitted youth share is obtained in two steps. First, the fitted values of a weighted regression of the youth share's growth rate on year fixed effects, the growth rate of the Bartik index and the growth rate of the instrument are computed. Second, a weighted regression of these fitted values on year fixed effects and the growth rate of the Bartik index is used to form the residuals. The dashed line represents the linear fit. City regions are the labour market regions of Berlin, Bremen, Cologne, Dortmund, Düsseldorf, Essen, Frankfurt, Hamburg, Munich and Stuttgart.

4. Results

This section presents the results from estimating the relationship between the regional employment rate and the youth share. We first start by discussing the generic impact of this variable before turning to the results of the FMM analysis.

4.1. Homogeneous youth-share effects

Table 2 contains the coefficients and standard errors from estimating the model in Equation (3) by OLS and two-stage least squares (2SLS). The OLS results show that an increase in the share of the youth population by 1 percent is predicted to reduce the employment rate by 0.2 percent, ceteris paribus. The corresponding 2SLS estimate is in the same order of magnitude, but about 40% lower. Both effects are significantly different from zero at the 1 percent significance level. So, accounting for possible selection due to inter-regional migration through the use of an IV identification strategy does not materially affect the estimated coefficients. This finding is in line with our expectation of lower youth share participation translating into lower employment rates.

Table 2 – Generic impact of youth share (18–24) on log(employment rate)

	OLS	2SLS
log(Youth share)	−0.199***	−0.274***
	(0.007)	(0.011)
Time dummies		Yes
Bartik index		Yes
First-stage F-stat.		1257.71***
N		1,269
R^2	0.938	0.932

Note 1 : All variables are first differenced ad observations are weighted by regional population shares.
Note 2 : significance levels are denoted by: *p $< .05$; **p $< .01$; ***p $< .001$.

The qualitative results of this analysis stand in contrast to the estimated coefficients reported by Shimer (2001) who finds an increase in the youth share to raise the employment ratio. This difference may be seen as an indication that the German labour market operates differently from its US counterpart. The 2SLS results are qualitatively comparable to those reported by Garloff et al. (2013), but the estimated coefficient is considerably smaller. Given that the time periods considered are almost identical, a

possible explanation for this difference is the inclusion of data from East Germany.[3]

Shimer ([2001](#)) stresses that regressing the overall employment rate on the youth share does not provide any insight on the underlying mechanisms. Since the overall employment rate can be expressed as a weighted average of the corresponding ag-specific rates, different mechanisms are possible. On the one hand, the estimated effect may represent a change in the composition of the population: younger age groups have a lower probability of being employed and an increase in the youth share shifts weight to younger age groups which would then automatically lead to a decrease in the overall employment rate. On the other hand, changes in the size of the youth share may also have an indirect effect by changing age-specific employment rates, since in a develop country like Germany age distributions typically display a diamond-like shape. Thus, an increase in the share of youth cohorts would mean an increase in participation of similarly aged groups, hence higher impact on employment. Table 3 shows the result from estimating the model of Equation (3) for separate age groups.

Table 3 – Generic impact of youth share (18–24) on employment rate of specific age cohorts

	20–24	25–34	30–44	45–54	55–64
log(Youth share)	-0.85^{***}	-0.29^{***}	-0.12^{***}	-0.03^{*}	-0.53^{***}
	(0.04)	(0.02)	(0.02)	(0.02)	(0.06)
Time dummies			Yes		
Bartik index			Yes		
First-stage F-stat.			866.85^{***}		
N			1,269		
R^2	0.78	0.87	0.83	0.81	0.82

Note 1 : All variables are first differenced ad observations are weighted by regional population shares.
Note 2 : significance levels are denoted by: *p $< .05$; **p $< .01$; ***p $< .001$.

We show that, as expected, changes in the youth share have the largest effect on the own age group: an increase by 1 percent is estimated to reduce the employment rate

---

[3]This hypothesis is supported by the results of the FMM analysis which shows that the effect of the youth share is smallest (in absolute values) in the East German regions. See Table 4.

of those aged 20–24 by 0.85 percent. The size of the effect generally decreases as the youth age group increases. The magnitude of the estimated coefficient falls considerably for age groups 25–34 and 35–45 but remains statistically significant, whereas it becomes insignificant for the employment rate of those aged 45–55. Indeed a possible explanation of this pattern is that the substitutability with younger workers decreases with the age difference and that therefore changes in the supply of younger workers are less relevant for the labour-market outcomes of older age groups. Contradicting this pattern, there is also a sizeable negative impact on the employment rate of those aged between 55 and 64. A similar effect at the upper end of the age range is reported by Garloff et al. (2013) who speculate that older workers might find it easier to retire early when a large youth cohort enters the labour market.

4.2. Regionally heterogeneous youth share effects

4.2.1. Model performance

After discussing the results of estimating a model in which the effect of the youth share on the employment rate is assumed to be constant, this section focuses on the possible existence of spatial heterogeneity in the relationship between the youth share and the employment rate by using an FMM approach.

One of the prevailing questions with this type of analysis is which number of clusters should be chosen. Figure 4 shows the values of different information criteria across the number of clusters. According to the Akaike information criterion the model improves as the number of clusters grows. In contrast, the Bayesian information criterion and the integrated completed likelihood criterion, which both penalise an increase in the number

of model parameters, suggest that the optimal number of clusters is rather small—two to be precise. As discussed subsequently, the main conclusions from the analysis are robust to the use of a larger number of clusters as shown in Table 4.
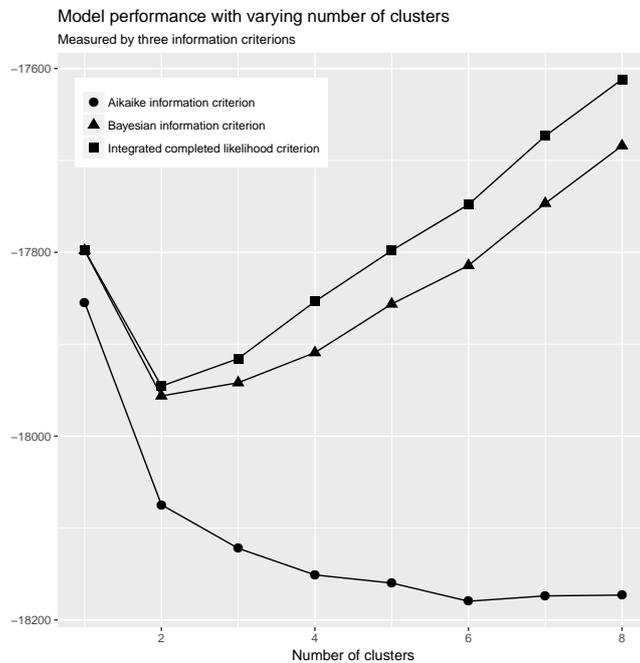
Figure 4 – Model performance of using various numbers of clusters

The results of the ex-post analysis reveal a well-behaved segmentation of regions. To illustrate this point, Figure 5 shows the distribution of the probabilities of being assigned to each group identified by the 2-cluster specification.[4] In each case the histograms are strongly polarised, indicating that for most regions the probability of belonging to a specific cluster is either close to zero or close to unity.

---

[4]Histograms for 3 and 4-cluster specifications can be found in Appendix A.1.
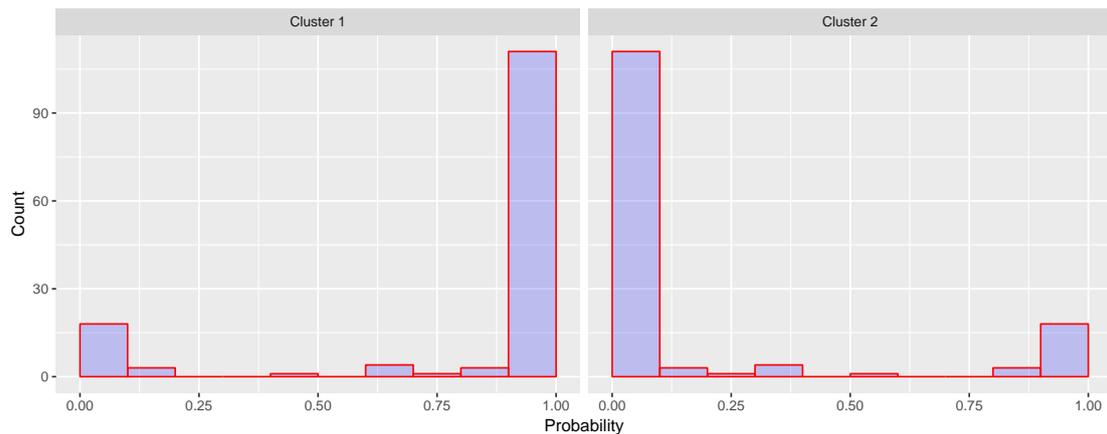
Figure 5 – Histogram of the probabilities of the 2 cluster specification

4.2.2. Regression results

The results of the FMM analysis suggest that there are two unique clusters of regions that differ in terms of how changes in the size of the youth share affect the employment rate, though similar conclusions can be drawn from using solutions with three or four clusters. Figure 6 shows the geographical distribution of models with two to four clusters if we assign a region to that cluster for which the corresponding probability is largest.

The specification with 2 clusters in panel (6a) illustrates clearly that the labour market regions containing large cities—such as Berlin, Hamburg, Munich, Frankfurt and Düsseldorf—form a separate group. In the case of Berlin and Frankfurt as well as the Eastern Germany cities of Dresden and Leipzig this relationship appears to also spill over into neighbouring regions. Once the number of clusters is increased to three, (panel (6b)), most regions in Eastern Germany are grouped into one cluster, while the cluster of cities remains largely unchanged. Finally, allowing for 4 clusters (panel (6c)) leads to a further differentiation of the Western part of Germany into two parts. Again, the city cluster is unaffected by this change.
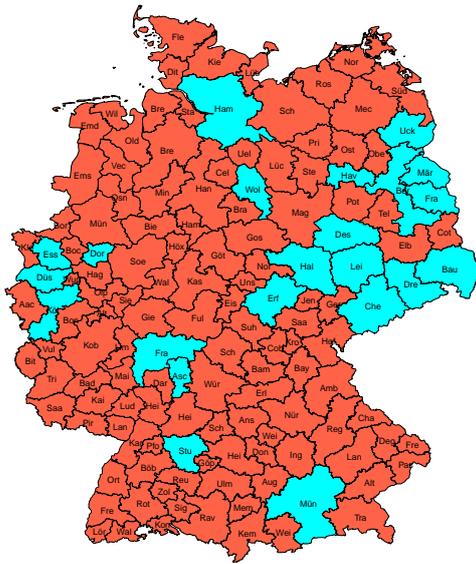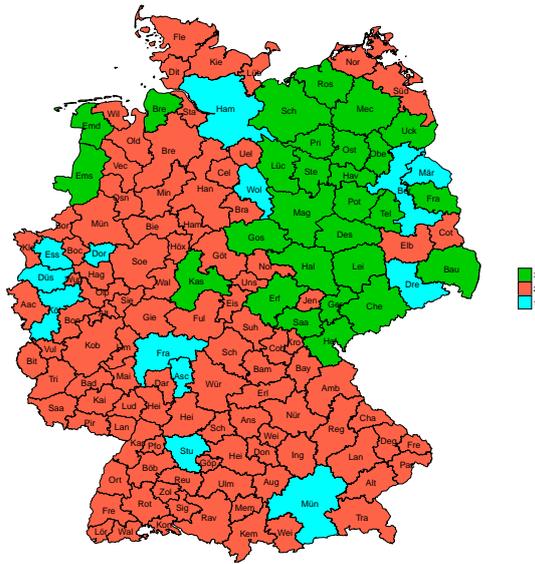
(a) 2 Clusters
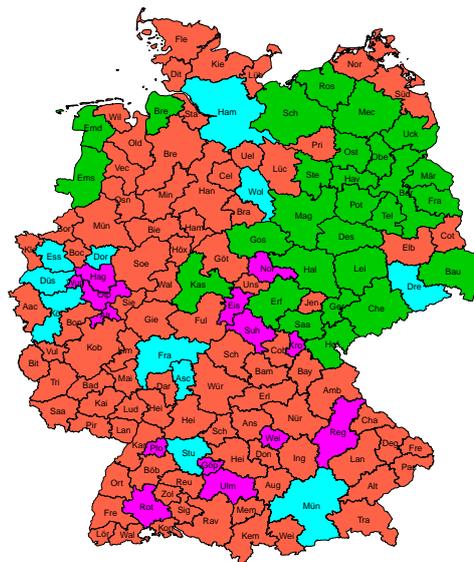
(b) 3 Clusters

(c) 4 Clusters

Figure 6 – Clusters of homogeneous impacts of youth share on employment across Germany

Table 4 shows the 2SLS coefficients of the youth share variable for specifications containing 2, 3 or 4 clusters. In each case the estimated coefficients are negative and highly

statistically significant. While there is no difference in terms of sign, the magnitude of the coefficient estimates varies considerably across groups of regions. For the 2-cluster specification the elasticity of the employment rate with respect to the youth share is estimated to be -0.33 percent in regions containing large cities, but stands at only -0.20 percent in the remainder of the country.[5] When the number of clusters is increased to 3 or 4, the negative effect of the youth share continues to be largest among city regions, while it is smallest within Eastern Germany.

Table 4 – Impact of youth share (18–24) on log(employment rate) by clusters for 2, 3 and 4 cluster specification

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
2 cluster specification				
log(Youth share)	-0.328^{***}	-0.201^{***}		
	(0.022)	(0.020)		
3 cluster specification				
log(Youth share)	-0.281^{***}	-0.220^{***}	-0.144^{***}	
	(0.029)	(0.025)	(0.021)	
4 cluster specification				
log(Youth share)	-0.380^{***}	-0.228^{***}	-0.129^{***}	-0.230^{***}
	(0.032)	(0.024)	(0.024)	(0.024)
First differenced	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes
Bartik index	Yes	Yes	Yes	Yes

Note 1 : observations are weighted by regional population shares.
Note 2 : significance levels are denoted by: *p $< .05$; **p $< .01$; ***p $< .001$.

4.2.3. Ex-post analysis of clusters

This subsection provides a comparison of the two clusters identified by the FMM analysis according to different demographic and labour market characteristics. We hypothesised above that the impact of youth cohort changes are different between urban and non-

---

[5]Due to the precision of the estimation the 95% confidence intervals of the estimates do not overlap.

urban labour markets and that these differences can be attributed to differences in labour demand elasicties. Indeed, we find that urban and non-urban labour market regions differ in their impact and that urban labour market regions have more pronounced effects of changes in youth cohort sizes on employment rates. However, to understand the underlying mechanisms, it is interesting to see which regional characteristics differ the most between the two clusters.

Table 5 gives the mean regional characteristics for the two clusters. Obivously, Cluster 1 regions—those containing predominantly large cities—are larger in terms of employment and also display a larger population density. The city cluster also has a higher share of foreign employees. In contrast, the age structure of the city cluster does not appear to differ substantially from that of the non-metropolitan Cluster 2, with perhaps the exception of the older cohorts (with age betwen 55–64), which are represented slightly more in the city cluster.

The difference in sector structure between the two clusters is clearly more pronounced. The city cluster displays relatively more workers in services while the non-city cluster has relatively more production workers.

In terms of the two groups' occupational structure, the city cluster displays clearly a larger share of employees working in engineering and managerial occupations, therefore occupations with higher degree of task complexity. Whereas simple manual and qualified manual occupations account for larger fractions in the non-metropolitan cluster. There is some evidence that job turnover is larger in the regions of the city cluster though the differences are not as pronounced as the occupation or sector structure.

Overall, Table 5 confirms the evidence from Figure 6 that the effect of youth cohort size on the employment rate is different in the large and densely populated labour

23

Table 5 – Mean regional characteristics per cluster

Region characteristic	Cluster 1	Cluster 2
Age cohort (18–19)	0.037	0.039
Age cohort (20–24)	0.094	0.095
Age cohort (25–29)	0.091	0.089
Age cohort (30–34)	0.096	0.096
Age cohort (35–39)	0.117	0.120
Age cohort (40–44)	0.132	0.135
Age cohort (45–49)	0.126	0.127
Age cohort (50–54)	0.113	0.113
Age cohort (55–59)	0.097	0.095
Age cohort (60–64)	0.096	0.092
Employment in sector agriculture	0.022	0.019
Employment in sector mining	0.001	0.001
Employment in sector other employment	0.015	0.010
Employment in sector production	0.259	0.320
Employment in sector services	0.636	0.590
Employment in sector technical occupations	0.067	0.060
Employment in task agriculture	0.020	0.018
Employment in task engineering	0.031	0.023
Employment in task managerial occ.	0.033	0.022
Employment in task professions	0.019	0.016
Employment in task qualified buss. and adm. occ.	0.219	0.192
Employment in task qualified manual occ.	0.132	0.161
Employment in task qualified service occ.	0.057	0.057
Employment in task semi-professions	0.087	0.084
Employment in task simple buss. and adm. occ.	0.088	0.085
Employment in task simple manual occ.	0.115	0.147
Employment in task simple service occ.	0.126	0.127
Employment in task technical occ.	0.046	0.046
Foreign working population	0.086	0.076
Job ends	0.024	0.022
Job starts	0.025	0.023
Population density	423.431	215.825
Total employment	473960.773	140511.468

24

markets regions of Germany. These findings confirm our intuition of a larger elasticity of demand in urban areas with respect to higher density and differentiated workforce, and a differentiated occupational structure.[6]

5. Robustness checks

Table 5 provides the regional characteristics that we have found most important in the impact of the size of the youth cohort on employment rates. However, some other drivers might be envisaged to have impact as well: namely, the potential impact of (internal) migration, regional differences in education enrollment and alternative measurements of employment rates. As a bottom-line, we have found none of these alternative drivers to be alternating our main results. We discuss these three further below sequentially.[7]

5.1. Impact of internal migration

Internal migration is an important factor in Germany. Especially, youngsters in eastern Germany migrate out after education. If our instrument does not properly account for this, then our results will potentially be biased.

Our main assumption is that our instrument—the size of the youth cohort 15 years ago—is valid. That is, that it is independent from current employment rates except through the size of the current cohort and conditional on time-invariant regional unobserved characteristics. So, theoretically, internal migration should not play a role: we identify

---

[6]Ideally, we would like to see differences in educational attainment as well between the two clusters. Unfortunately, our data does not allow us to investigate this.

[7]For sake of brevity we do not report all regression results, only the relevant ones—also because they do not change noteworthy. But they are available upon request.

the impact of the youth cohort by the size of that age group before *potential* regional migration has an effect. However, a valid concern might be that our instrument only picks up the *average* relation and does not take properly the outer tails into account. So, if our instrument fails to take extreme migration out of eastern Germany into the larger metropolitan regions into account, then our estimates (especially those of the clusters) are biased.

To check for this, we have used polynomial specifications (squared and cubic) for our instrument as well for both our generic and cluster analyses. These analyses do not change the results (with, e.g., an impact of -0.26 for both the squares and cubic specifications for the generic model).

5.2. Impact of education

A second understandable concern is the impact of education in labour markets where the age group 18–24 is overrepresented. For example, in 'university towns' where most of the economic activity is centred around/related to the universities, the economic potential may be limited in terms of job creation. Germany is a country where educational attainment is rather high and takes relatively long to acquire a diploma (once apprenticeship is considered). So, presence of large groups of youths in education may drive up the unemployment rates. In order to address the potential impact of youth clustering in some regions and artificially influencing (un-)employment rates we employ two different types of analysis. We firstly repeat model (3), by excluding university towns from our analysis. The university towns in our analysis are defined as the non-city regions (the non-city regions defined in this analysis does not coincide with cluster 2) that host relatively large universities. These are the regions of Gottingen, Giessen and Marburg, Tubingen, Passau,

Regensburg and Greifswald. The generic impact of the youth cohort does not changes significantly and becomes -0.274 with a standard error or 0.012.

Secondly, we re-define the youth share by excluding the age group 18–19 and base our definition on the share of the age group 20–24 in the population aged 20–64. This new variable potentially excludes those who are still registered in education or apprenticeship. This selection has a disadvantage of leaving out the youngest cohort who are actively in the labour market. However, only some 30 per cent of the population in this age group in Germany is in employment. Moreover, in Germany the students in their early ages are allocated to areas of specialisation depending on their skills and education has a strong impact on people's career path. Therefore, because of these distinct features we expect the workers in this group to influence the regional outcomes differently compared to other age groups. This new definition does impact our results slightly, as, e.g., the generic impact of youth cohort changes to -0.373 (with a standard error of 0.027).

5.3. Do definitions alter the story?

Another concern may be due to varying definitions used in the literature. In particular, Shimer (2001)'s definition of employment. The main difference in his definition of employment relates to the denominator defined as the sum of unemployed and employed workers. This approach excludes discouraged and inactive working age population which would otherwise be counted in the denominator.

Moreover, including eastern Germany might affect our results as well. We therefore run our analyses with a subset of regions excluding eastern Germany.

Again, altering the definition of employment rates and excluding eastern Germany does not change our generic results. The magnitude of the coefficients become even stronger when using the more conservative definition of Shimer (2001) with an impact of -0.343 (with standard error of 0.020). The coefficient for analysis when excluding eastern Germany does not change much and becomes -0.274 (with standard error of 0.011).

6. Conclusion

This paper estimates the effect that changes in the size of youth cohorts have on the employment rate using data from German labour market regions covering the period 2001–2010. Specifically, it challenges the assumption made by the extant cohort size literature that these effects are constant across space. We argue that the magnitude of the employment effects will depend on the regional elasticity of labour demand as well as on the magnitude of the shift in labour supply caused by a change in the population share of young age groups. Existing evidence suggests that the working of these channels differs between regions, thereby casting doubt on the assumption of spatial effect homogeneity.

We formulate the hypothesis that the effect that a change in the share of the youth population has on the employment rate is more pronounced in urban areas due to labour demand being more elastic and a greater availability of non-employment options. This hypothesis is tested using a novel empirical approach which does not rely on an ex-ante grouping of regional labour markets but instead endogenously assigns regions to clusters within which the model's estimated coefficients are similar.

Overall, we find that increases in the youth share lead to a decrease in regional employment rates. These results, therefore, do not provide any evidence that larger cohorts are

associated with favourbale labour market outcomes as suggested by other studies. We then proceed to show that the size of these effects vary across regions and that our empirical approach indeed identifies a grouping that distinguishes between highly urbanised regions and all other regions. Our results show that within the large and densely populated labour markets of metropolitan regions increases in youth cohort size have a considerably larger negative impact on the employment rate with an estimated elasticity of 0.33 compared to 0.20 in non-metropolitan regions. The fact that urban regions stand out is robust to the use of a larger number of clusters.

These results provide a basis for evaluating the impact that the projected decline in the relative size of young age groups will have on regional employment rates. According to our findings, the process of declining youth shares will put upward pressure on regional employment rates. This effect is expected to be distributed unequally across space. In particular, an equiproportional reduction in the regional youth share is estimated to have a more pronounced impact on employment rates in metropolitan areas, thereby contributing to a divergence between urban and non-urban employment rates. However, this divergence would be less pronounced if the reduction in the youth share turned out to be larger in non-metropolitan areas.

References

Bartik, T. J. (1991). "Boon or Boondoggle? The debate over state and local economic development policies". In: (cit. on p. 9).

Biagi, F. and C. Lucifora (2008). "Demographic and education effects on unemployment in Europe". In: *Labour Economics* 15.5, pp. 1076–1101 (cit. on p. 3).

Brunello, G. (2010). "The effects of cohort size on European earnings". In: *Journal of Population Economics* 23.1, pp. 273–290 (cit. on p. 3).

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). "Maximum likelihood from incomplete data via the EM algorithm". In: *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38 (cit. on p. 14).

Di Addario, S. (2011). "Job search in thick markets". In: *Journal of Urban Economics* 69.3, pp. 303–318 (cit. on p. 5).

Di Cataldo, M. and A. Rodríguez-Pose (2017). "What drives employment growth and social inclusion in the regions of the European Union?" In: *Regional Studies* 51.12, pp. 1840–1859 (cit. on p. 3).

Duranton, G. and D. Puga (2004). "Micro-foundations of urban agglomeration economies". In: *Handbook of regional and urban economics.* Vol. 4. Elsevier, pp. 2063–2117 (cit. on pp. 5, 12).

Fertig, M., C. M. Schmidt, and M. G. Sinning (2009). "The impact of demographic change on human capital accumulation". In: *Labour Economics* 16.6, pp. 659–668 (cit. on p. 3).

Garloff, A., C. Pohl, and N. Schanne (2013). "Do small labor market entry cohorts reduce unemployment?" In: *Demographic Research* 29, p. 379 (cit. on pp. 4, 10, 12, 16, 18).

Glaeser, E. L. and D. C. Mare (2001). "Cities and Skills". In: *Journal of Labor Economics* 19.2, pp. 316–42 (cit. on pp. 5, 12).

Herwartz, H. and A. Niebuhr (2017). "Regional labor market performance in Europe: error correction dynamics and the role of national institutions and local structure". In: *International Regional Science Review* 40.3, pp. 270–296 (cit. on pp. 5, 12).

Holl, A. (2018). "Local employment growth patterns and the Great Recession: The case of Spain". In: *Journal of Regional Science* 58.4, pp. 837–863 (cit. on p. 3).

Korenman, S. and D. Neumark (2000). *Cohort Crowding and Youth Labor Markets: A Cross-National Analysis.* Ed. by D. G. Blanchflower and R. B. Freeman. January. University of Chicago Pres, pp. 57–106 (cit. on pp. 3, 10).

Kosfeld, R. and D.-Ö. A. Werner (2012). "Deutsche Arbeitsmarktregionen–Neuabgrenzung nach den Kreisgebietsreformen 2007–2011". In: *Raumforschung und Raumordnung* 70.1, pp. 49–64 (cit. on p. 6).

Maiti, A. and D. Indra (2016). "Regional Variations in Labor Demand Elasticity: Evidence from US Counties". In: *Journal of Regional Science* 56.4, pp. 635–658 (cit. on p. 5).

Martin, R. and P. Tyler (2000). "Regional employment evolutions in the European Union: a preliminary analysis". In: *Regional Studies* 34.7, pp. 601–616 (cit. on p. 3).

Moffat, J. and D. Roth (2016). "The Cohort Size-Wage Relationship in Europe". In: *Labour* 30.4, pp. 415–432 (cit. on pp. 3, 10).

Mosca, I. (2009). "Population Ageing and the Labour Market: Evidence from Italy". In: *LABOUR* 23.2, pp. 371–395 (cit. on p. 3).

Newbold, K. B. and W. M. Brown (2015). "The Urban–Rural Gap in University Attendance: Determinants of University Participation Among Canadian Youth". In: *Journal of Regional Science* 55.4, pp. 585–608 (cit. on p. 5).

Niebuhr, A. (2003). "Spatial interaction and regional unemployment in Europe". In: *European Journal of Spatial Development* 5, pp. 1–26 (cit. on p. 3).

Overman, H. G. and D. Puga (2002). "Unemployment clusters across Europe's regions and countries". In: *Economic policy* 17.34, pp. 115–148 (cit. on p. 3).

Rios, V. (2017). "What drives unemployment disparities in European regions? A dynamic spatial panel approach". In: *Regional Studies* 51.11, pp. 1599–1611 (cit. on p. 3).

Saito, H. and J. Wu (2016). "Agglomeration, congestion, and US regional disparities in employment growth". In: *Journal of Regional Science* 56.1, pp. 53–71 (cit. on p. 3).

Shimer, R. (2001). "The Impact of Young Workers on the Aggregate Labor Market". In: *The Quarterly Journal of Economics* August, pp. 969–1007 (cit. on pp. 3–5, 7, 10, 12, 16, 17, 27, 28).

Skans, O. N. (2005). "Age effects in Swedish local labor markets". In: *Economics Letters* 86.3, pp. 419–426 (cit. on pp. 4, 12).

Wedel, M. and W. A. Kamakura (2012). *Market segmentation: Conceptual and methodological foundations.* Vol. 8. Springer Science & Business Media (cit. on p. 13).

Wheeler, C. H. (2008). "Local market scale and the pattern of job changes among young men". In: *Regional Science and Urban Economics* 38.2, pp. 101–118 (cit. on p. 5).

Zeilstra, A. S. and J. P. Elhorst (2014). "Integrated analysis of regional and national unemployment differentials in the European Union". In: *Regional Studies* 48.10, pp. 1739–1755 (cit. on p. 3).

A. Appendices

A.1. Histograms for 3 and 4 cluster specifications
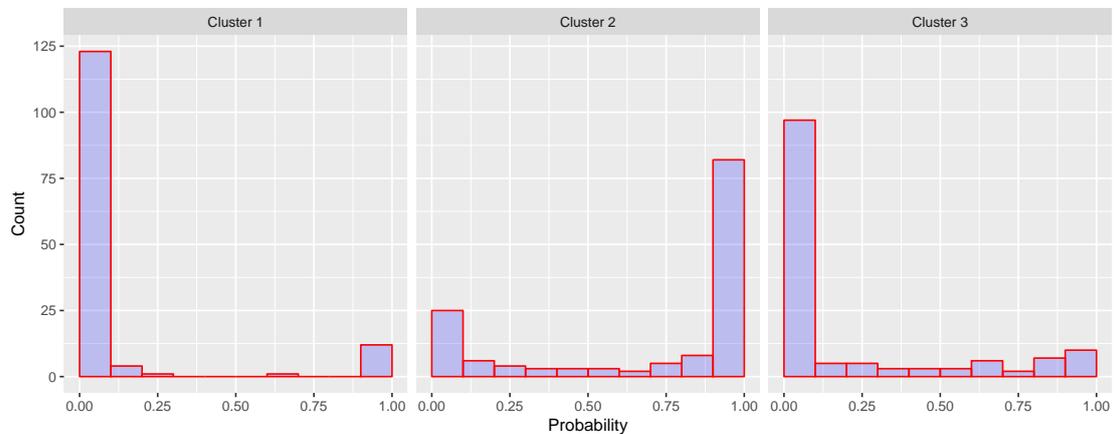
**Figure 7** – Histogram of the probabilities of 3 cluster specification
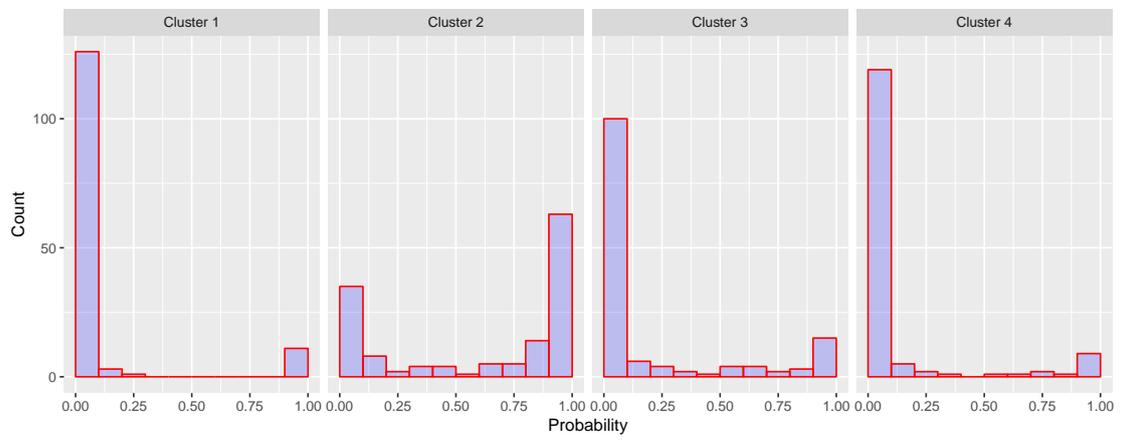
Figure 8 – Histogram of the probabilities of 4 cluster specification