ERASMUS SCHOOL OF ECONOMICS

Labor market developments in Bulgaria. Analyses of regional differentials

Master Thesis

MSc. Economics and Business

Student name: Bilyana Vencheva

Student number: 330824

Supervisor: Dr. B. Crutzen

Co-reader: Prof. Dr. B. Jacobs

Executive summary

The question of the main determinants of changing labor market landscape has been among the center of the political discourse. In addition to output losses due to underutilization of the productive capacity of the economy, unemployment might have serious impact on individual`s physical and mental health. This necessitates active government engagement in understanding the reasons behind observed unemployment levels. Recently the focus has been shifted towards a more elaborate exploration of the determinants of labor market outcomes at a country level as a useful tool to assess individual outcomes.

Despite the fact that unemployment can be divided into three types – frictional, structural and cyclical, the majority of empirical studies that aim at establishing the reasons behind observed trends tend to group together factors that explain either type. The neoclassical theory suggests that a flexible labor market without significant adjustment costs should in general lead to equalization of unemployment levels across different regions as excess labor is able to move freely to locations with low unemployment. Nevertheless, evidence on spatial differentials within a country tends to favor the notion of long terms persistence in rates of regional unemployment that is inconsistent with those assumptions. As a result, the empirical literature has proposed several main determinants of unemployment differentials.

While data availability explains the relative abundance of empirical studies on labor market conditions within developed countries, studies on economies in transition have been scarce. The present paper aims at addressing this gap. The main research question focuses on defining the main factors that explain regional variation in unemployment rates. In order to answer the research question the paper utilizes available data on 28 administrative districts in Bulgaria during the period 2000 – 2013. Following Perugini (2008) and Lottmann (2012), panel data analyses are performed on regional level data so as to take advantage of the cross sectional and time series properties of the data. The main results indicate that cyclical fluctuations as measured through the output gap, age structure and industry composition are the main determinants of local labor market outcomes. The data also shows the presence of clustering of districts with persistently similar unemployment rates and regional characteristics which necessitates the inclusion of an econometric specification that allows for spatial spillovers. Some preliminary evidence of statistically significant spatial interactions also point at these conclusions.

Table of contents

1. Introduction

The major political changes following the collapse of the communist regime in 1990 have had major repercussions on the Bulgarian economy. The fall of trade barriers and liberalization of prices alongside the sharp reductions in state subsidies on major economic sectors were considered necessary for a quick transition to a market economy. The immediate results of the sweeping reforms were related to sudden output contractions due to decrease in domestic consumption, exports and investment in the real economy. The initial shock and unstable political climate led to relatively slow market reorientation and persistent stagnation. As a result the labor market was adversely affected with rising levels of unemployment which produced nation-wide increase in poverty, social and personal insecurity (CSD, 1996). Regional disparities in labor market conditions that seemed non-existent throughput the communist period started to increase in magnitude with the contraction of economic activity (Boeri & Scarpetta, 1996). The first decade of the transition period was characterize by persistent emigration flows, slow to no employment growth in certain regions as well as contraction of the labor force and increase in unemployment rates. Despite the initial shock, the economy started to exhibit signs of recovery at the beginning of the new century, nevertheless, several key policy issues with respect to the general labor developments were outlined. Specifically, the labor marked was characterized by ageing population, high levels of low skilled labor force, high youth unemployment and rise in regional divergence (Bratoeva, 2009).

In an attempt to resolve the above mentioned problems, a seven year employment strategy was adopted by the Bulgarian government in 2008. It largely followed the framework of the broader EU Lisbon Agenda that outlined the need for specific employment, unemployment rates targets, higher cross regional cohesion and improvement in the output growth in an attempt to boost high quality job creation across the EU member states. The plan is a direct consequence of the acknowledgement that unemployment leads to output losses due to underutilization of the productive capacity of the economy. Furthermore, it might have serious impact on individual`s physical and mental health. Long term unemployment has persistent impact on personal development while youth unemployment might lead to desperation, drug addiction and socially unacceptable behavior (Kieselbach, 2003). For instance, a growing number of empirical studies

have stressed the relationship between poor economic conditions (unemployment, income inequality) and prevailing crime rates[1]. This necessitates active government engagement in understanding the reasons behind observed unemployment levels. Targeted policy is needed to address any existing labor market inefficiencies and promote employment - intensive growth. The question of the main determinants of changing labor market landscape has therefore been at the center of the political discourse (Maivali & Stierle, 2013).

There is a growing amount of empirical literature on labor market outcomes. Based on the unit of examination, most of the empirical literature examines between-countries variation in unemployment and emphasizes on the importance of country specific labor policies and the general institutional framework, for instance social security, tax and retirement legislation. Crivelli et al. (2012) for instance utilize a panel of 167 countries and provide evidence that an effective policy mix should focus on reducing labor and product market rigidities for improved labor market outcomes. In a similar panel data study within Europe, Dopke (2001) establishes that the industry mix, labor market flexibility and wage determination process have a disproportionate influence over labor market dynamics within a country. Furthermore, Orlandi (2012) performs a study on the determinants of structural unemployment[2] within EU countries. He finds that the main variables that capture labor market outcomes are the labor market structural indicators, demand shocks and crisis events. In a cross country comparison between Spain and Portugal, Blanchard and Jimeno (1995) find that insufficient adjustments of the wages with respect to unemployment is the main factor contributing to high unemployment levels and their persistence in Spain. In contrast, institutional factors such as lower unemployment benefits have led to lower adverse effects on employment provision within Portugal enabling wage rate decrease that can absorb excessive unemployment. Nickell (1997) reaches a similar conclusion on the impact of generous unemployment benefits and long entitlement periods on unemployment rates. However, they also note the importance of high unionization and collective bargaining and high minimum wages.

[1] See for instance: Saridakis and Spengler (2009); Tang et al. (2009); Maddah (2013) on the relationship between unemployment and different types of crimes committed.

[2] Structural unemployment is one of the major types of unemployment that is characterized with a supply of labor in excess of the available demand for labor for a given wage rate and is mostly as a result of skills mismatch (Ehrenberg & Smith, 2009). See *Section 2* for further details.

Recently the focus has been shifted towards a more elaborate exploration of the determinants of labor market outcomes at a country level as a useful tool to assess individual outcomes. While some studies utilize time series data and analyze factors affecting aggregate (un)employment[3] trends, a growing amount of the literature has turned its attention to regional labor market differentials within a country as disparities in unemployment between regions in one country is even higher than between different countries (Vega & Elhorst, 2012). A pioneering paper by Blanchard and Katz (1992) introduces a model on regional evolutions of employment, unemployment and labor participation rates which sets the stage for more detailed inquiries into regional labor outcomes determinants. Taylor and Bradley (1997) find that the main factors that account for the large divergence in regional unemployment rates are the difference in unit labor costs, the industry mix as well as employment density and those determine regional competitiveness and respectively the stability of labor market. Brunello et al. (2001) emphasize on the importance of social policy (such as transfers) and the failure of local wages to respond to regional conditions as an important determinant of unemployment outcomes in the provinces of Italy, while Lopez-Bazo et al. (2000) finds that region specific demographic factors explain the largest proportion of unemployment rates differentials in Spain.

Problem statement and research question

While data availability explains the relative abundance of empirical studies on labor market conditions within developed countries, studies on economies in transition have been scarce. In Bulgaria more than 1.2 million jobs were destroyed in the process of extensive socio economic reforms after the fall of the Soviet Union, the majority of which in the manufacturing and service sector (Garibaldi et al., 2001). Despite the general belief that sharp increases in unemployment in previously centrally planned economies would be temporary and reversed within a couple of years, data has evinced otherwise with approximately 40% of the unemployment in the 90s being of long term nature (Nesporova, 2002)[4]. Apart from overall unemployment persistence, limited workers` mobility and skills deterioration contributed to sharp increases in regional disparities in the beginning of the transition period. These trends remained relatively unchanged as based on

[3] See for instance: Kabaklarli et al. (2011); on macroeconomic determinants of youth unemployment; Türkyılmaz and Özer (2008)

[4] Recent figures are even more disturbing with 57% of the registered persons in 2014 being unemployed for more than a year (National Statistical Institute data and own calculations Available at: http://www.nsi.bg/en/content/6503/unemployed-and-unemployment-rates-national-level-statistical-regions-districts)

2014 figures, regional unemployment rates in Bulgaria varied from as high as 26% (Shumen region) to as low as 7% (Kardzhali region). This naturally raises the question of the main reasons behind the observed labor market developments in Bulgaria.

Based on the discussion above the following research question is formulated:

What are the main determinants of regional labor market dynamics in Bulgaria in the period 2000 – 2013?

A. What is the regional evolution of labor market variables?
B. What factors explain regional unemployment differentials?

In order to answer the research question and the pertinent sub-questions, the paper utilizes available data on 28 administrative districts in Bulgaria during the period 2000 – 2013. Following Perugini (2008) and Lottmann (2012), among others, panel data analyses are performed on regional level data so as to take advantage of the cross sectional and time series properties of the data. The main results indicate that cyclical fluctuations as measured through the output gap, age structure and industry composition are the main determinants of local labor market outcomes. The data also shows the presence of clustering of districts with persistently similar unemployment rates and regional characteristics which necessitates the inclusion of an econometric specification that allows for spatial spillovers. Some preliminary evidence of significant spatial interactions also point at these conclusions.

The contributions of the paper are several. First of all, it extends the relatively scarce literature on transition economies and labor market conditions by providing an in-depth analysis of regional evolutions within Bulgaria. Most of the studies on transition economies utilize a cross section of several countries, while regional disparities are largely neglected. Bornhorst and Commander (2006), for instance, employ regional data on six economies in transition (including Bulgaria), however, their paper models the general dynamics of employment, unemployment and non-participation rates without focusing on region specific factors. Furthermore, the time period under examination extends from 1990 to 2001 – right after major disturbances in the economy – while the present paper employs the most recent data during a relatively stable period. Overall, the research seeks to outline the major labor market dynamics in Bulgaria on a regional level

during the last 14 years in an attempt to provide useful insights that would enable informed decision making on a local level.

The paper is organized as follows. *Section 2* introduces the main concepts that will be analyzed and discusses both theoretical and empirical literature on determinants of regional labor disparities. It also includes the conceptual framework and the hypotheses for the analyses. *Section 3* focuses on explanation of data sources while *Section 4* elaborates on the econometric specifications employed in the paper. The main results are presented in *Section 5* alongside some additional robustness checks and specifications. *Section 6* provides a discussion on the results obtained while *Section 7* concludes.

2. Literature Review

The section provides the conceptual framework of the analyses on regional labor dynamics. The main concepts in relation to unemployment are defined and discussed in the first part. The second part briefly presents the major methods employed in modeling the difference in regional unemployment rates simultaneously discussing their theoretical merits. In the following part the variables most frequently found in the empirical literature as the main determinants of differential regional labor market outcomes are presented.

2.1 Unemployment

In general the population can be divided into two groups of people – those who are in the labor force and those who are outside the labor force. Furthermore, the people who are in the labor force can be either employed or unemployed. Respectively, academics and policy makers are interested in two main variables in relation to labor markets – unemployment rate, which is defined as the ratio of the numbers of unemployed to the total labor force, and employment rate, usually calculated as the ratio of employed over total adult population[5]. While it might seem reasonable to expect that unemployment and employment rates move in opposite directions, their dynamics are also determined by the change in the labor force. If the latter increases rapidly, high unemployment rates might as well be associated with high employment rates (Ehrenberg & Smith, 2009). Although the evolutions of both variables are important to consider in assessing

[5] Alternative specifications are also possible, for instance: employed over number of people within the age group 15 -64.

aggregate macroeconomic trends, policy makers tend to focus on unemployment rates as they provide valuable insights into the labor market structure. If unemployment is too high, then a substantial amount of the population is unable to support itself which might signal deeper social problems, furthermore, it means that a significant part of the economic resource of a country – labor – is being wasted thus the economy is not operating at full capacity. On the other hand, very low unemployment levels might have inflationary consequences due to upward wage pressure. Despite the apparent ambiguity, economists often stress on the concept of a **natural rate of unemployment** (also referred to as a state of full-employment). Albeit not precisely defined, it is generally believed to be a measure of a long term unemployment rate prevalent in the economy when the labor market is in equilibrium.

Theoretically, unemployment can be divided into three types – frictional, structural and cyclical (also known as demand deficient). **Frictional unemployment** arises as a result of normal labor market dynamics, as employees change jobs in order to find a better match. Thus even if aggregate labor demand equals labor supply, frictional unemployment might arise as it takes time for unemployed to find information on and apply for suitable vacancies or respectively for employers to find the best candidates. Therefore the main factors that affect the amount of frictional unemployment are the efficiency of the job searching process and the availability of information on vacancies. Furthermore, high unemployment benefits might discourage job search and thus increase frictional unemployment (Poterba & Summers, 1995). **Structural unemployment** is a relatively permanent form of unemployment that is the result of labor market imbalances of two types: skills mismatch and high mobility costs. The former indicates that within a given area the skills demanded by employers differ from the skills supplied by the labor, while the latter signals the existence of excess labor in one area that cannot move freely across different regions due to high costs involved. The inflexibility of wages (for instance due to minimum wage laws or high union bargaining power[6]) might exacerbate the incidence of structural unemployment in case of skills mismatch[7]. The third type of unemployment, **demand deficient**, is the result of normal business cycle fluctuations that urge firms to change employment patterns. Due to real wage rigidities, a negative shock on the demand side of the labor market would lead to a decrease in employment and a rise in demand deficient

[6] See for instance: Blanchard and Jimeno (1995)
[7] For a more elaborate study on structural unemployment see Phelps (1998)

unemployment. However, if wages are flexible enough or able to adjust within a relatively short period, employment levels would be lower compared to the prior state, nevertheless, there will be no cyclical unemployment.

Although theoretically several categories of unemployment are defined, most of the actual measures provided by official sources, such as public administration or labor market surveys, do not distinguish between the different types. Therefore the majority of empirical studies that aim at establishing the reasons behind observed unemployment levels, tend to group together factors that explain either type.

2.2 Modeling unemployment empirically

Traditional macroeconomic models do not provide a general framework on within-country variation in unemployment levels despite the fact that large disparities in regional labor market outcomes have been documented (Taylor & Bradley, 1997). Large differences in regional unemployment may be highly inefficient (Elhorst, 2003). Consistently depressed regions might find themselves in a vicious circle that would lead to persistent inability to improve socio-economic conditions for their population. Unfavorable labor market environment leads to migration of the most skilled as well as a net loss of population that has a negative impact on the demand for local production, which further exacerbates regional labor market problems. Thus understanding the underlying factors behind regional labor market developments is essential. This section outlines the four major types of models that are employed so as to explain local unemployment differentials. The discussion is largely based on Elhorts (2003) and additional sources are cited upon introduction.

The four major approaches that are used to model unemployment in the empirical literature are as follows:

- Single equation models;
- Implicit models;
- Accounting identity models and
- Simultaneous models with interactions.

Within a **single equation approach** the unemployment rate is the independent variable that is explained through various factors that have an effect on labor demand and supply. Absent any

theoretical explanation, it might be difficult to establish whether the factors included in relation to unemployment rates have the expected impact on unemployment rates. There are three main models that are estimated within the single equation approach: the *Beveridge curve*[8], the *cyclical sensitivity model* and the *amenity model*[9]. The *Beveridge curve models* the relationship between unemployment rate, on the one hand, and vacancy rates, on the other. Although direct causal inferences on the reasons for observed unemployment cannot be established, shifts in the Beveridge curve can provide useful insights on the efficiency in the labor market that might be tied to observed labor market outcomes. As normally one would expect a negative relationship between unemployment and vacancy rates, an outward shift that corresponds to increased unemployment at the same vacancy rates might indicate important structural problems such as immobile labor force or regional skills mismatch. The *cyclical sensitivity models* the relationship between regional unemployment rate and the observed national counterpart. Overall, it tries to establish the extent to which local unemployment varies with aggregate levels. Nevertheless, within this model the theoretical basis for the relationship is not well established. Furthermore, it does not provide any insights on the reasons behind regional disparities in labor outcomes. Lastly, the *amenity model* represents a framework within which certain regional characteristics have a direct impact on labor markets through their effect on firm`s production functions or workers utility functions.

The **implicit approach** is characterized by implicitly modeling unemployment rates within a framework that explains labor market dynamics. The most prevalent example is the *Blanchard and Kartz model*. They establish a theoretical framework based on four equations that describe local labor market dynamics, i.e. determinants of short run labor demand as well as supply, a wage setting equation that determines the relation between unemployment and prevailing wages, and another equation that models tong run labor demand. By construction the model allows for long run unemployment rates that differ between regions.

The **accounting identity approach** is a fairly straightforward method based on which regional unemployment level is derived as a natural function of regional labor supply (working age

[8] See for instance: Cheshire (1973); Gordon (1987); Jones and Manning (1992); Holzer (1993)
[9] See for instance: Roback (1982); Marston (1985); Montgomery (1993)

population, labor force participation and net commuting) and labor demand (employment)[10]. Alternatively the **simultaneous model with interactions** assumes that there exists a constant feedback effect between various regional labor market factors and unemployment rates.

Overall, independent of the model used, unemployment is essentially determined by the development and interactions between factors that affect three underlying aspects of the labor market – labor demand, labor supply and the wage-setting process. The differences between the models are related to the extent to which any of the individual models incorporates those variables in the determination process of unemployment. Based on that, Elhorts (2003) concludes that most of the empirical models are reduced form equations on the relationship between local unemployment and different explanatory variables that proxy for regional labor demand, supply and wage-setting. The following section presents the most commonly identified explanatory variables.

2.3 Determinants of regional unemployment differentials

The neoclassical theory suggests that a flexible labor market without significant adjustment costs should in general lead to equalization of unemployment levels across different regions as excess labor is able to move freely to locations with low unemployment (Lopez-Bazo et al., 2000). In addition to outmigration of workers, equilibrium on the labor market can be achieved through adjustment in wages that could also lead to increase in the number of firms that are attracted by a large pool of available workers and lower wages. Nevertheless, evidence on spatial differentials within a country tends to favor the notion of long terms persistence in rates of regional unemployment that is inconsistent with the assumption of rapid adjustments and equalization of unemployment rates across locations. In general low elasticity of wages with respect to unemployment levels as well as large migration costs can explain a substantial part of the divergence in distribution. However, unemployment can also be a function of region-specific endowments and characteristics and this relationship can rationalize long term persistence of heterogenous equilibrium rates of unemployment across regions (Martin, 1997). The empirical literature on spatial disparities provides guidance on the most important determinants of regional unemployment rates. The following part elaborates on the findings.

[10]Unemployment = (Working age population *Labor force participation) + Net Inward Commuting - Employment

2.3.1 Demographics

The age structure of the population within a region might have an impact on labor market dynamics. There are two alternative theories that relate proportions of young people to observed unemployment rates. On the one hand, it is expected that young people are highly flexible and mobile which would lead to the prediction that they would migrate in case of persistent unemployment. On the other hand, young workers tend to switch jobs more often relative to elders in an attempt to find a better match (Brown & Sessions, 1997). Furthermore, young people who live with their parents would have higher opportunity costs of migration and might prefer longer spells of unemployment than moving out to another region. Therefore, it is also likely to expect that areas that are characterized with relatively higher proportions of young people may also exhibit disproportionately higher rates of unemployment. Indeed, several empirical studies find support for this hypothesis[11].

Similar controversy exists in relation to the proportion of older workers in labor force. On the one hand, several studies suggest a negative association between high proportion of older employees and regional levels of unemployment (Elhorst, 1995; Molho, 1995; Partridge & Rickman, 1995). It is suggested that they are likely to be more efficient at finding new occupations due to experience at the job search process, furthermore, they would normally have acquired more skills relative to youngsters during their professional experience which is likely to increase chances of retaining a job during downturns. On the other hand, McPherson and Flores (2012) suggest that structural unemployment is more likely to be observed in cases of aging population as older workers might have long unemployment duration relative to younger colleagues. This would entail e positive relationship between unemployment rates and share of older people in the labor force. Therefore, it is not possible to indicate a priori which effect is expected to dominate.

2.3.2 Employment

The majority of empirical studies on spatial distribution of unemployment include local employment growth as one of the explanatory variables. Explicitly based on the accounting identity equation, employment growth leads to a reduction in the unemployment rate. Evans and McCormick (1994) find that a substantial part of the regional unemployment can be attributed to changing patterns in employment creation and destruction. Despite the fact that most of the

[11] See for instance Aragon et al. (2003), Elhorst, 1995, Lottman (2012)

literature finds significant negative correlation between the variables, theoretically, the relationship between unemployment and employment rates can be more complex. Demographic and institutional factors can have important implications on the interaction between employment and unemployment (Kosfeld & Dreger, 2006). For instance, in the case of high population growth that is followed by an increase in labor force and less than proportional magnitude of job creation would tend to move both variables in the same direction. Alternatively, Elhorst (2003) points out that rural-urban migration that is a result of efforts to increase job supply in areas with high unemployment, might in fact lead to higher levels of unemployment that would materialize in positive coefficients between increased employment possibilities and unemployment rates. A major criticism to the inclusion of employment growth as a right hand side variable pertains to the idea that this masks the underlying mechanisms that lead to a rise in employment itself.

2.3.3 Human Capital

The education level of the population is frequently included within the set of explanatory variables of spatial differentials. Highly educated individuals are more likely to possess skills that would render them demanded by the employers. It is more likely that they are better and more efficient during the job search process, while simultaneously being less likely to be laid off work (Elhorst, 2003). Therefore both their probability of falling into unemployment is lower and once unemployed the duration of unemployment period is expected to be relatively short lived compared to the less educated. Furthermore, relatively high skilled workers are more likely to migrate out of areas with persistent high levels of unemployment. The same might be less valid for low educated workers who might face disproportionately higher migration costs. Therefore the relatively higher proportion of low educated labor force in some locations might be able to explain some of the unequal distribution of unemployment rates (Lopez-Bazo et al, 2000). Elhorst (2003) notes that in case of persistent poor economic performance, regions might suffer from the so called "*low skill poverty trap*" which exacerbate local labor productivity problems as it induces constant outflow of highly skilled workers, therefore reducing overall regional attractiveness.

2.3.4 Industry mix

Prevailing spatial unemployment disparities are quite often associated with regional differences in industrial specializations. Locations that are characterized with relative abundance of employment opportunities within growing industries, such as the service sector, are expected to

have low unemployment rates in comparison to areas with production that is concentrated in agriculture or manufacturing, which are considered declining (Elhorst, 2003). This proposition assumes that unemployment is industry- rather than region-specific, however, some studies tend to find that unemployment rates in the same industry differ across regions (Martin, 1997).

One way in which academics model the impact of the industrial mix is through respective shares in regional employment. Based on conventional logic one would expect that regions specializing in growing industries should exhibit lower unemployment. Nevertheless, if the job growth in one industry is not enough to offset contraction of jobs in another, structural unemployment might still increase. Therefore, the empirical literature does not provide a definite answer on the expected impact of the industrial mix. Bradley (1997), for instance, finds that the share of agriculture and regional unemployment rates are positively associated in Italy and negatively associated in Germany and UK. Elhorst (2003) argues that industrial shares might be a noisy indicator of the impact of sectoral shifts on unemployment as due to the methodology of measurement, shares in one industry would automatically increase in case employment in others falls and this is not likely to contribute to lower unemployment rates.

Despite the methodological issues, the industrial mix within a region does play a role as confirmed by several authors that employ alternative measures and look at sectoral shifts at regional level[12]. Recently, the empirical literature has focused on industrial diversity as opposed to industry employment shares. Mizutani et al (2003) and Izraeli and Murphy (2003), among others, suggest that industrial diversity is negatively associated with unemployment rates. The most widely used index to measure the extent of industrial concentration within a region is the Herfindahl`s specialization index. It indicates the extent to which employment in a region is concentrated within one industry. If overall employment is more evenly distributed across different sectors then more opportunities are available for redeployment particularly in case of cyclical economic fluctuations, therefore unemployment in highly diversified regions is expected to be lower.

[12] See for instance Neumann and Topel (1991), Holzer (1991), Samsom (1994),Hyclak (1996), Lopez-Bazo et al.(2000)

2.3.5 Regional Output

Fluctuations in regional employment levels are also as a result of standard business cycle variations. The demand for regions output will tend to vary over time in accordance with the goods market in which a region specializes. Therefore economic fluctuations will have an impact on regional unemployment levels. Studies that model regional labor market outcomes also take into account measures that proxy for demand fluctuations. Taylor and Bradley (1997), for instance, measure region`s cyclical activity through deviations of local GDP from its long term trend. They argue that this is appropriate in case the fluctuations can indeed be attributed to real changes in regional demand. Despite the fact that they employ a relatively short time period to estimate the corresponding regional GDP gaps[13] and reach the conclusion that there is no statistically significant regional cyclical demand effect. Nevertheless, the present thesis also incorporates GDP deviations in the model as this effect may vary between countries. The deviations are expected to have a negative impact on unemployment levels.

2.3.6 Additional explanatory variables

Due to data availability, the above mentioned explanatory variables are most frequently used in the empirical literature that models regional unemployment. Nevertheless, I also discuss the relative importance of three other factors that might have implications for regional labor market outcomes, namely: participation rates, housing and public benefits.

The impact of **participation rates** on unemployment is ambiguous as various empirical studies have found conflicting results. Based on the accounting identity alone, an increase in the participation rate would cause an increase in unemployment rate, therefore a positive relationship should be more likely *ceteris paribus*. Nevertheless, high participation rates might act as an encouragement for job growth which might directly imply lower unemployment levels. Layard, (1997) argues that the possible negative impact following an increase in the labor force is mostly compensated by higher employment creation. Additionally, a rise in employment opportunities is more likely to induce people to enter the labor force, which in turn would reduce unemployment if the extra jobs are filled in by the new entrants. On the other hand, differences in regional participation rates might also signal variation in some fundamental characteristics of the working age population. For instance, a region characterized by relatively low participation rates might be indicative of a larger fraction of less committed workers or a relatively larger proportion of

[13] A regional panel data within Italy, Germany and UK ranging from 1983 to 1994 is employed.

people with low investment in human capital (Elhorst, 2003). This would essentially suggest a negative association between labor force participation and unemployment. Indeed, most of the empirical literature tends to support this view[14] .

The **housing** situation can have an important effect on unemployment levels that is largely evinced through its impact on search behavior. According to Eurostat Housing Statistics, approximately 87% of the Bulgarian population resides within owner occupied buildings. This would essentially reduce incentives to migrate due to the high transaction and opportunity costs that are associated with a decision to locate to a region with better employment opportunities. The resulting limited mobility would tend to increase unemployment in regions characterized by lower share of renters (Muellbauer & Murphy, 1991). In addition to the proportion of owner occupiers, house prices can serve as an important indicator of the relative attractiveness of certain working areas, therefore it is expected that regions with improved employment conditions and lower unemployment rates, would tend to exhibit higher house prices. Alternatively, house prices directly reflect costs of living. Therefore, even in the presence of high unemployment, workers might choose to remain within a certain area as living expenses are disproportionately lower.

Brunello et al. (2001) argue that an explanatory variable that might have a direct impact on labor supply is the amount of **social transfers** per head. Inter-household transfers from the employees and retirees to the unemployed that are usually younger family members might have a certain negative impact on labor supply though unemployment duration and the reservation wage. A higher amount of transfers might further reduce regional migration flows alongside participation decisions. Indeed the authors find evidence that this partially explains the persistent trends of higher unemployment in the Southern regions of Italy. Despite the fact that this variable is not included in the majority of studies, it might be particularly relevant for Bulgaria due to cultural peculiarities. Approximately 53 % of people between 25-34 years of age tend to live with their relatives[15] which might in essence lead to higher inter-household transfers. Although government transfers per head is a noisy proxy for the impact of household transfers on incentives to work,

[14] See for instance: Hofler and Murphy (1989), Blanchard and Katz (1992); Decressin and Fatás (1995)
[15] Euranet Plus BNR Discussion. See reference.

the variable is included in the analysis due to a lack of available data on real intergenerational transfers.

Despite the preceding discussion, it is important to note that it is relatively difficult to place a specific expectation sign on the impact of an individual explanatory variable on the regional unemployment rate as changes in each might have an impact on both labor supply and demand. Furthermore, a change in the independent variables can cause similar changes in both employment and unemployment if, for instance, increase in employment induces inflow into the labor force which might lead to increasing the pool of unemployed (Bronars & Jansen, 1987). Thus, the impact on unemployment is determined by the relative effect of the regional variable on workers and firms and is, therefore, a matter of empirical investigation.

3. Data

The main purpose of the thesis is to provide a more in depth analysis on labor market developments within Bulgaria with a specific focus on regional outcomes. The relevant units of examination are the 28 districts[16] in the country for the period 2000-2013 and sample represents the most recent regional data available after the changes in the Territorial Organization Act in 1999. The regions are administrative territorial units that correspond to level NUTS3 from the European NUTS classification. The scarcity of information on variables that are included as explanatory variables of regional outcomes precluded a more extensive decomposition (municipal level for instance) of differential labor markets. Due to the relatively small sample, there is inevitably a trade-off between high statistical efficiency and a better specified model, nevertheless, I believe that the fourteen years of data could still provide useful insights into the most recent trends in labor dynamics.

The major part of the data is readily accessible on the National Statistics Institute (NSI) website platform. The main regional level variables are obtained from the Statistical Reference Books published for the period 2000-2012. They include data on labor force, employment and unemployment rates, as well as the annual average number of employed by 15 major economic activity groupings. The auxiliary sources are presented upon their introduction below.

[16] NUTS 3 level according to European regional classification.

3.1 Dependent variables

The first part of the analyses focuses on modeling regional unemployment rates. A person is considered unemployed if within the relevant time period, he/she is not employed and is actively looking for a job for a period of four weeks at least. An active method of job search is considered contact through labor offices, through advertisements and other job boards, direct contact with employees and others. This measurement of unemployment is based on Labor Force Surveys that are conducted on quarterly basis and is as such is a subject of statistical bias. Therefore, for additional robustness of analyses, I also consider an alternative measure that is based on administrative data alone that provides records on the average number of persons registered at the labor offices within a given district. The latter also permits the examination of long term unemployment as information on the unemployed registered for more than a year is available under the Regional Statistics Publications/Labor Market at the NSI database.

3.2 Independent variables

In order to control for the impact of demographics on regional unemployment disparities, I employ a measure related to the age structure of the population equal to the proportions of working age population in a district that fall within the age categories 15-24 (Young) and 50-64 (Old) respectively. The information is available under the Demographics and Social Statistics Database. The impact of human capital is measured through the proportion of population aged 25-64 that have completed higher education[17] (HighEduc) as well the share with primary or lower educational attainment(LowEduc). The data was obtained through an official enquiry and is made available by the National Employment Agency. The regional output variables were obtained at the regional office of NSI at the city of Plovdiv. As regional price indexes are not available, the national harmonized CPI index was used to convert the latter into constant 2005 BGN currency.

As summarized in the literature review, the industry mix variable is likely to have significant explanatory power in intra- regional unemployment variation. I have utilized two approaches to analyze its impact. Similar to Taylor and Bradley (1997), Lopez-Bazo et al. (2002) and Korobilis and Gilmartin (2010) industrial specialization effects are incorporated through the regional shares of employment within agriculture, manufacturing and services. Nevertheless, as pointed out in *Section 2* industrial diversity might be a better indicator for the ability of the regional

[17] Colleges/Universities

economy to absorb unemployment and therefore a better explanation for labor outcome disparities. Alongside Izraeli and Myrphy (2003) and Mizutani et al. (2003) I calculate the Herfindahl Index of industrial diversification for each region based on the following formula: [18]

$$H_{it} = \sum_{j=1}^{n} \left(\frac{EMPL_{ijt}}{EMPL_{it}} \right)^2$$

Where H_{it} is the industrial diversification index for region *i* at year *t*, $EMPL_{ijt}$ is the number of employed in industry *j* for region *i* at year *t* and $EMPL_{it}$ is the total regional employment during the year. The index essentially presents industrial employment as a share of total regional employment. The higher the value of the index, the more concentrated regional employment is within a given sector, therefore lower index values correspond to higher diversity.

One important weakness of the Herfindahl index pertains to the fact that it is not capable of directly accounting for the impact of the industrial structure. Mizutani et al.(2003) show that regions with the same index can have completely different industrial structures which in turn would influence the structural component of labor market dynamics. An example can facilitate the reasoning behind it. Imagine there are two districts that have similar employment composition, except for manufacturing and services and for one of the districts employment in manufacturing equals 10% of total labor force and for services – 5% and vice versa for the other. The calculated index would be the same. In reality however, the industrial structure is very different. One way to capture this impact, as suggested by Mizutani et al (2003), is through the calculation of the *location quotient* which "*measures the extent to which a metropolitan area is providing employment in an industry compared with the national average*." (page 6). The location quotient is computed as follows:

$$LC_{ijt} = \frac{EMPL_{ijt} / EMPL_{it}}{EMPL_{jt} / EMPL_{t}}$$

[18] Alternatively, there are other indices on industrial concentration that are frequently utilized, including the Ogive and the National Average Index. Nevertheless, they essentially summarize the same information and variations tend to be small. For an extensive discussion on industrial concentration indices, see Kessler et al. (2007)

Where LC_{ijt} is the location quotient of industry *j* in region *i* at time *t*, $EMPL_{ijt}$ is the number of employed in industry *j* for region *i* at year *t* and $EMPL_{it}$ is the total regional employment during the year. $EMPL_{jt}$ and $EMPL_{t}$ are the national level counterparts and stand for aggregate employment in industry *j* relative to overall employment in the country at time *t*. Within the empirical analyses that follow, the regional location quotients for five main sectors are calculated: Agriculture (LC AGRI), Manufacturing (LC MAN), Construction (LC CONS), Wholesale and retail trade (TRADE), Public Administration (LC ADM). A LQ that is equal to 1 indicates that the regional specialization in that industry is equal to the national average, a LQ that is higher than one indicates a region that has a relatively higher concentration in the respective industry. Industrial diversity and concentration measures also aim at providing a robust inference w.r.t. relative contribution of sectoral composition in explaining regional unemployment disparities.

3.3 Descriptive statistics

In order to gain a better understanding of the dataset employed, descriptive statistics of the main variables of interest are provided in *Table 1*. It can be seen that unemployment rates based on the two available measures[19] exhibit quite similar patterns in their statistical properties therefore I do not expect the results of the subsequent econometric analyses to be contingent on the dependent variable employed. The average unemployment rate over the observed sample period is approximately 13.5% with substantial variation both across and within different regions. The lowest unemployment rate was observed in region Sofia capital with 1.3% in 2008[20] and the highest value was recorded in the North East region of Targovishte in 2000. The large within regional variability renders less likely the possibility that heterogeneous regional unemployment rates might be of equilibrium nature, on the contrary, substantial variations might be expected in case of high sensitivity to cyclical factors.

In addition to unemployment, several other variables deserve attention. First of all the educational attainment of the population exhibits considerable variation between regions signaling that varying human capital levels might help explain consistent discrepancies of labor patterns between different regions. Second of all, high within regional volatility in demand

[19] LO: registered at Regional Labor Offices; LF: based on Labor Force Surveys

[20] This is the region with lowest unemployment rates over the whole sample period. Even in the aftermath of the crisis, the average unemployment was around 3.5%.

driven explanatory variables such as employment and GRP growth[21] alongside LF participation rates provide some evidence that idiosyncratic shocks might exert substantial impact on regional growth and labor dynamics. On the other hand, large standard errors might signal the presence of outliers which might bias the results. In alternative specification I aim to address this issue.

In addition to aggregate sample statistics, *Table 2* in Appendix B provides some information on mean values contingent on varying unemployment rates based on four quartiles. On average the total number of observations within a quartile is 100. The table serves as a preliminary evidence of the direction of correlation between unemployment rate and the explanatory variables discussed in the literature review. It can be seen that demographic variables (gender and age structure) show relatively little variation w.r.t different quartiles of unemployment, while educational attainment and sectoral variables exhibit clear variation as registered unemployment rate grows. More specifically, time-region periods with high share of educated workforce, low share of relative employment in agriculture (as well as higher share of urban population) and high values for demand driven factors are associated with lower unemployment rates.

Analyses on the stationary properties of data were performed. The results for Levin-Lin-Chu (2002)[22] unit roots test are available in Table 3 in Appendix B. Some evidence of non-stationary patterns was found in the explanatory variables *Urban* and *Young*, nevertheless the null hypothesis of unit roots for unemployment rate and all other right hand side regressors were rejected and no variable transformations are implemented.

[21] The largest reduction of GRP was recorded in Pernik region in 2009. Nominal GRP decreased from 1.4bln BGN in 2008 to as low as 0.75blnBGN in 2009. The major part of the reduction was due to substantial reductions of industry GVA (from 0.7bln to a little more than 0.2bln BGN)

[22] Im-Pesaran-Shin (2003) and DF tests were additionally performed for cross validation of estimation methodologies without difference in the conclusions. Results not reported.

Table 1. **Descriptive Statistics Main Variables, 2000-2013**

Variables	Obs	Mean	Std.Dev.			Min	Median	Max
			overall	between	within			
Dependent								
Unemp LO	392	13.85	6.02	4.36	4.24	1.30	13.43	34.90
Unemp LF	392	13.41	6.81	4.71	4.99	1.20	12.50	34.40
Independent								
Demographics								
Female	392	47.77	11.18	3.05	10.77	45.51	47.20	268.04
Young	392	20.87	5.38	1.41	5.20	13.02	20.78	113.46
Old	392	34.77	8.99	3.87	8.14	26.44	34.25	200.92
Educational Attainment								
High Educ	392	18.35	5.89	5.69	1.85	8.90	17.40	46.00
Low Educ	392	27.98	10.41	9.19	5.18	3.30	27.20	55.50
Sectoral								
AGRI	392	4.75	2.79	2.69	0.88	0.21	4.11	13.08
IND & CONS	392	37.93	8.15	7.85	2.62	14.93	38.79	56.88
Services	392	31.48	8.25	7.39	3.90	16.29	29.91	65.01
ADM	392	5.76	1.60	5.42	6.05	2.58	5.66	11.91
Private	392	66.75	8.06	5.42	6.05	42.34	67.71	79.95
DIV index	392	0.15	0.04	0.04	0.01	0.10	0.15	0.30
Demand Driven								
Empl gr	364	0.01	4.62	1.31	4.44	-14.36	0.12	14.61
Output gap	392	0.00	0.51	0.00	0.51	-1.23	-0.06	4.34
GDP gr	364	1.47	8.11	1.91	7.89	-67.97	1.50	42.85
GDP cap gr	364	2.88	8.19	1.37	8.07	-67.09	2.84	43.81
Additional								
Act Rate	392	49.35	4.85	3.83	3.07	37.90	49.30	63.10
Urban	392	64.76	11.54	11.66	1.31	33.08	65.52	95.64
Density	392	88.92	161.57	164.22	5.82	32.80	53.90	961.20
Minorities 1	392	2.28	0.86	0.83	0.26	0.50	2.23	5.99

Note : The table provides descriptive statistics of variables of interest. All varialbes, except for DIV index, are in percentage points. Demographic variables are expressed as a ratio to working age population; Sectoral varialbes are in percenage of total employment. Detailed definitions are available in ApendixA.

Source: National Statistical Institute, National Employment Agency and own calculations

4. Methodology

The econometric specification employed in the analyses of the determinants of regional disparities in labor market outcomes is a single equation model of cross section time series data, i.e. regional unemployment level is a right-hand side variable that is explained through a series of factors. More specifically, I employ panel data analyses and utilize both FE and RE estimator. Analogous to Lottman et al. (2012), the finalization of the model to be estimated and analyzed subsequently is determined through several steps:

1. *Choice of explanatory variables*

As the total list of potential factors exceeds 20, I have decided to analyze their relative contribution before specifying the final model. Based on the Akaike`s (AICc) and Bayesian Information criterion (BIC) goodness-of-fit is estimated and alternative specifications are compared[23]. Nevertheless, based on the theoretical literature a set of explanatory variables (MAIN) is included in all models based on their significant contribution to unemployment rates. Those variables are as follows: demand side (EMPL growth & Output gap); demographic structure (Young & Old); sectoral structure (IND & CONS) and human capital (High/Low Educ). The selection of the additional factors is based on the fixed effects model as discussed in Baltagi(2008):

$$Un_{it} = \sum_{r=1}^{8} \alpha_r \text{MAIN} + \sum_{r=9}^{R} \alpha_r x_{rit} + \mu_i + \delta_t + \varepsilon_{it} \quad (1)$$

Where MAIN is the set of variables defined above, x_{rit} stands for the additional variables selected based on AIC and BIC, μ_i and δ_t stand for district and time specific fixed effects respectively[24]; α_r are the coefficients to be estimated and ε_{it} is the error term, assumed to be iid. The subscripts stand for region *i* at time *t*, *r* defines the independent variable. A detailed description of methodology is provided in Appendix C.

[23] Despite the fact that AICc and BIC cannot be informative about the absolute quality of the model, they can give an indication of relative performance, which I ultimately would like to assess.

[24] District specific FE control for time invariant individual specific unobservable characteristics that might influence the results and as such to some extent control for endogeneity bias; while time dummies are included to account for aggregate factors that affect district in a similar manner.

2. Final model analyzed

Based on the discussion above and AIC and BIC, the final set of explanatory variables that was chosen included MAIN as well as two additional indicators of sectoral composition (TRADE and ADM), the labor force participation rate (Act Rate), share of urban population (Urban) and a proxy for minorities (Min). The final model based on which the subsequent analyses are performed takes the following specification:

$$Un_{it} = \alpha_1 \text{EMP gr} + \alpha_2 \text{Output gap} + \alpha_3 \text{HightEduc} + \alpha_4 \text{LowEduc} + \alpha_5 \text{Young} + \alpha_6 \text{OLD} + \alpha_7 \text{MAN} + \alpha_8 \text{CONS} + \alpha_9 \text{TRADE} + \alpha_{10} \text{ADM} + \alpha_{11} \text{Act Rate} + \alpha_{12} \text{Urban} + \alpha_{13} \text{Min} + \delta_t + \varepsilon_{it} \quad (2)$$

The adequacy of the usage of least squares estimators depends crucially on the properties of the error term (Moon et al., 2004). Failure to meet the assumptions of the LS estimator[25] would result in biased and/ or inefficient estimates. Those assumptions are likely violated in a pooled cross sectional model. Unobserved factors and/ or endogenous explanatory variables would bias the estimates rendering simple OLS estimates inconsistent. Even if OLS is not inconsistent, it might not be optimal in the absence of i.i.d. error terms as the latter leads to incorrect estimation of the coefficient variance thereby biased standard errors and incorrect inferences (Beck & Katz, 1995). Even though time dummies might capture some of the error variance, Beck (2001) argues that this is unlikely to resolve problems related to other region specific unobservables, therefore panel data estimators should be always implemented. On the other hand, FE is less efficient than OLS as the former only explores within group variation thereby ignoring relationships between regions. Low variability in the explanatory variables within regions thus might hamper inferences on some useful interactions among unemployment and explanatory variables. One way to deal with the issue is the utilization of RE estimator, nevertheless potential endogeneity of regressors could also result in inconsistent and bias inferences. Therefore I report results for the three types of estimators: pooled OLS estimator as a baseline comparison, as well as the panel data estimators. The appropriateness of FE versus RE estimator is evaluated through the Hausman test.

[25] $E(e_t) = 0$, var $(e_t) = \sigma_e^2$ cov (e_t, e_s) and/or cov $(e_i, e_j) = 0$ for $t \neq s$ and $i \neq j$; and cov(x, e) = 0 (exogeneity of parameters).

Both FE and RE estimators include time fixed effects to control for common shocks that might affect the structure of the error term nevertheless the presence of serial correlation that is not accounted for can still lead to incorrect inferences. Panel clustered standard errors (PCSE) are deemed one way to deal with potential bias. Beck and Katz (1995) argue that they are perfectly suited to deal with those problems by adjusting standard errors correspondingly in the presence of serial correlation, nevertheless, they are equally accurate and similar to default standard error estimate in case of non-existence of the above mentioned issues.

5. Results

5.1 Sub question 1: Evolution of regional labor market variables

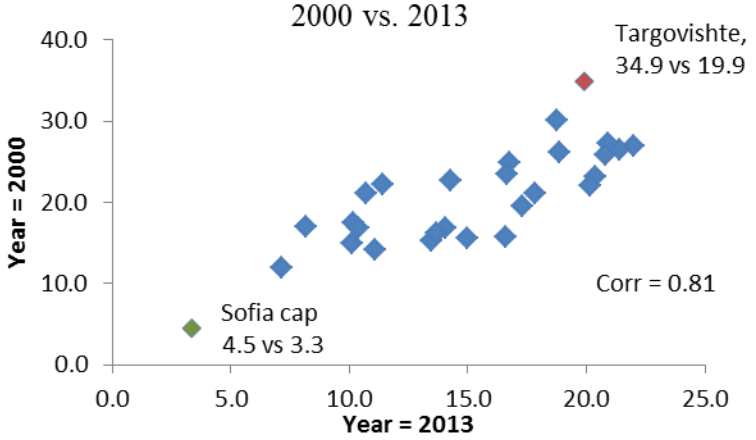
The following section aims at providing an overview of labor dynamics in Bulgaria for the past 13 years. Table 6 in Appendix C provides a more elaborate idea on the developments in five main labor market indicators: number of employed/unemployed, working age population (Pop WA) as well as LF participation (Activity Rate) and unemployment rate for the 28 districts and two distinct time periods (pre and post crisis). Two main things should be noted. First of all, the average unemployment rates in the aftermath of the financial crisis have been consistently lower or stable for all but one[26] regions compared to the pre-crisis period. This was mainly due to substantially high 2000-2001 unemployment rates driven by the collapse of the Bulgarian currency and banking crisis of 1998/1999. Nevertheless, despite the global economic recession of 2007/08, the highest rate observed over the second period was much lower (22.2%) compared to the observed 34.9% during the pre-crisis period, thereby showing signs of aggregate improvement in labor conditions.

Second of all, there is an indication of clustering of labor market outcomes on a regional level. Based on the accounting identity alone, the decrease in unemployment rates seem to have been driven by changes in labor force composition. The majority of districts witnessed an absolute decline in the number of registered unemployed. Nevertheless, neighboring districts seem to follow common trends and two major groups with contrasting developments are observed. Those within the Northern part of the country exhibit both a decrease in the number of persons employed as well as a substantial decline in working age population signaling that migration

[26] Blagoevgrad

outflows could explain part of observed labor dynamics and declining rates. An increase in migration flows is likely to contribute positively to a decline in structural unemployment on a regional level as people in depressed regions move towards regions with better opportunities. In the case of Bulgaria, those seem to be the Eastern and Southern parts of the country that have witnessed a decline in persons unemployed as well as more than proportional increase in employment.

Figure 1. Regional Unemployment rate 2000 vs. 2013



Despite aggregate trends of declining unemployment levels within regions, there is little evidence on convergence towards a common equilibrium. The majority of districts with the highest rates in the beginning of the sample period remained consistently among the worst labor market performers throughout. The high correlation coefficient of 0.81 (See Figure 1) between regional outcomes in 2000 and 2013 respectively signals that regional unemployment disparities are more likely to be of persistent nature rather than as a result of temporary market adjustments. Examination of demand side factors on a local level tends to support similar conclusions (See Figure 2-3). None of the demand side variables - employment growth and GDP per capita growth – provides evidence of strong association with unemployment levels. Although employment growth is weakly negatively associated with changes in unemployment rates (pp see Figure 2B), more robust econometric analyses are needed to assess the relative contribution of factors explaining regional disparities. The following sections explore this issue in detail.

Figure 2A. Employment growth & Unemployment

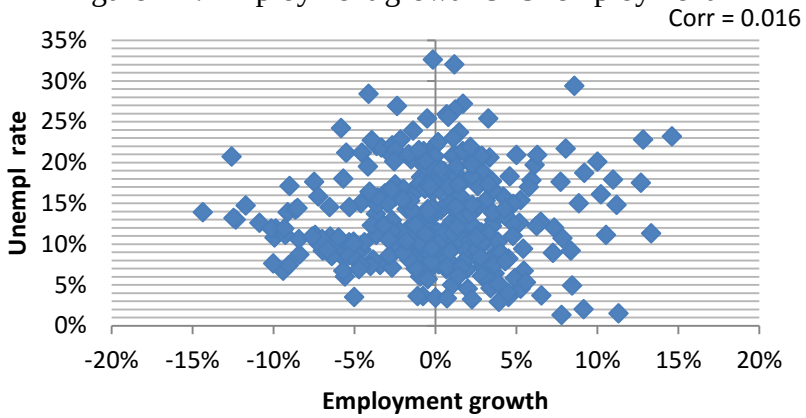


Figure 2B.

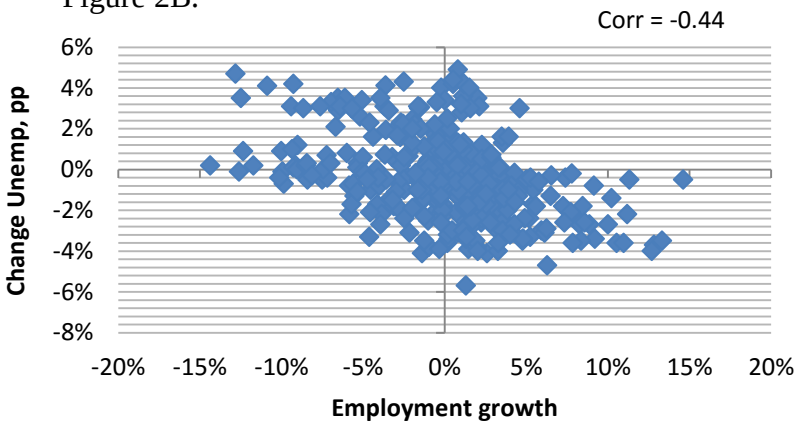


Figure3A

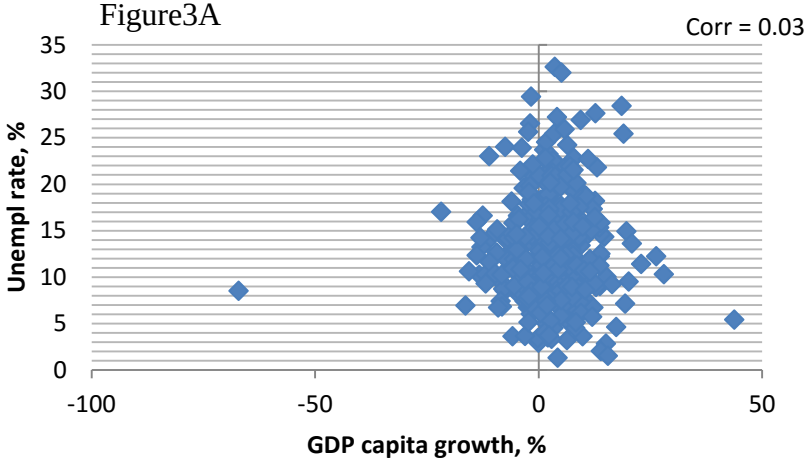


Figure 3B.

5.2 Sub question 2: Determinants regional unemployment rates

The results of the main econometric specification (2) are provided in Table 7. Although the main analyses are based on the Panel data models, FE and RE estimators, pooled OLS specification is provided as a baseline comparison. Nevertheless, FE and RE models that deal with potential unobserved individual heterogeneity are more appropriate. In an attempt to judge the appropriateness of the Panel data estimator, the Hausman test for systematic difference of coefficients is employed. With a p-value of 0.0015, the null hypothesis of equivalence is rejected at the 1% level, and as such the discussion that follows is entirely based on the results obtained under the FE estimator (Columns 1b through 4b in Table 7). It should be noted that all panel data specifications include year fixed effects as well as panel clustered standard errors to account for serial correlation. Additionally, I estimate two major specifications: one including only the MAIN variables based on theoretical considerations[27] (Table 7: 1b) and one FULL model that corresponds to equation (2) (Table 7: 2b). The main idea is to judge the robustness of the main determinants upon introduction of additional within regional variation. Last but not least, I provide two additional specifications: one that estimates the potential effect of generosity of benefits on unemployment levels (3b), and another one that aims at capturing the effect of standard of living, in terms of wage and costs, on unemployment levels[28].

The results for the panel data estimator indicate that overall there is no evidence of consistent demand driven impact on unemployment levels. Employment growth is insignificant in all specifications, while the significance on output gap depends on the specification employed and does not have the expected sign. Theoretically, the larger the output gap the lower the unemployment should be, however in the case of Bulgarian districts, it is the exact opposite. Based on the FULL model (2b), a one percentage point (pp. thereafter) increase in output gap is associated with 0.424pp. higher unemployment rate. There are two potential explanations for this observation. On the one hand, it is possible that the results are due to econometric issues and the

[27] See Discussion on variable selection

[28] As I do not have region specific CPIs so as to directly measure real wages, I have computed a proxy that takes into account the value of housing prices as they are observed to have the highest contribution to costs of living. The measure (the ratio between average wage and housing price) is therefore indicative to a certain extent of the standard of living within certain regions. It is expected to have a positive impact on unemployment as higher ratio is indicative of lower relative costs.

inability of small time series to adequately capture GRP trends and thus cyclical deviations are not correctly measured. One possible way to deal with this problem is the implementation of alternative proxies for aggregate activity. On the other hand, output gaps measure aggregate deviations from potential GRP. Therefore, high positive deviations that are indicative of larger than usual aggregate demand could subsequently lead to an increase in migration and labor force participation. However, if the additional capacity is not enough to absorb all new entrants in the LF, high unemployment would occur alongside high cyclical fluctuations. I explore both possibilities in additional specifications.

Based on the results, structural factors, such as demographics and industrial composition, lead to consistently significant and larger in absolute magnitude coefficients thus signaling that they are relatively more important in explaining persistent unemployment trends differentials within Bulgarian districts. More specifically, the age structure proxies are significant in all specifications, however expected signs are not related to usual theoretical predictions and results in comparable studies. Based on (2b) the larger the share of young people, the lower the observed unemployment rates are, whereas a 1pp. increase in the share of elderly is associated with 0.347pp. higher unemployment levels all else constant. It seems that the experience of old people is not enough to compensate for the lack of flexibility relative to younger workers and the loss of job is likely to result in persistently higher levels of unemployment. It is possible that unemployment that is of long term nature is more likely to occur in areas with proportionately higher share of old workers. Furthermore, there is some indication that young people are more likely to move easily between jobs and thus have lower spells of unemployment, or alternatively able to migrate to regions with better economic opportunities which would rationalize the negative coefficient.

As far as the educational attainment is concerned, the sign of the coefficients is as expected, nevertheless the significance depends on the estimator employed. In fact variation in levels of education does not seem to affect unemployment rate if one considers FE estimator alone which essentially explores within regional variation. Within the RE model which also incorporates across regional variation of explanatory variables, there is some evidence that heterogenous levels of human capital across regions might be able to explain part of the region specific unemployment rates, nevertheless, results are not robust to alternative specification. It is

furthermore interesting to note that the minority proxy[29] is significant in some of the specifications, indicating that the share of minorities could help explain fraction of unemployment level differentials both across and within regions. Nevertheless, it might also be the case that it is a noisy proxy unrelated to share of minority groups but instead more indicative of low opportunities which result in higher birth rates at a younger age.

Based on the results presented in Table 7, the industrial structure is highly significant in all specifications, signaling of its relative importance and major implications on regional labor dynamics. Larger share of manufacturing, construction and trade are being consistently associated with lower levels of unemployment. Additionally, a 1pp. increase in the share of public administration is associated with approx. 0.6 pp. reduction in unemployment rates (2b). On the one hand, it might be a proxy for artificial job creation that keeps unemployment levels at unnaturally low levels. On the other hand it might measure the proliferation of private activity within certain regions that necessitates strong public administrative support.

In addition to the MAIN and FULL model, Table 7 provides estimates on the effect of benefits (3b) and living standards (4b) on regional labor market conditions. Despite the fact that they have the expected sign, based on FE estimator alone, there is no evidence that either benefits (also proxying for intergenerational transfers) or higher real wages are associated with persistently high unemployment levels. The insignificance of the former indicates that adverse incentives are not likely to be of particular importance on individual`s decision to search for employment. Furthermore, it does not seem plausible that higher real wages (and/or lower costs of living) tend to compensate for worse labor market conditions[30] as the amenity model of unemployment differentials would suggest. Nevertheless, it should be noted that the estimated additional specification (3b and 4b) do not include the whole sample period compared to the MAIN and the FULL model. Specifically, data on benefits on a regional level is only available for the year 2000-2006, whereas information on average annual wages is available for 2008-2013

[29] Proxy for the share of minority groups, for instance: roma community. They tend to have very high birth rates at young age therefore I use a measure of birth rates. See Appendix A for a detailed overview of data.

[30] In alternative specifications, I explore separately the effects of average nominal wage and housing prices on unemployment levels. The main results indicate that an increase in the nominal wage rate is negatively associated with unemployment levels. The effect on house prices is insignificant independent on specifications employed. Results available upon request.

Table 7. **Determinants of Regional Unemployment Differentials**

	OLS				Panel data, FE				Panel data, RE			
	1a	**2a**	**3a**	**4a**	**1b**	**2b**	**3b**	**4b**	**1c**	**2c**	**3c**	**4c**
EMP growth	0.051	-0.026	-0.068	-0.025	0.001	-0.024	0.008	-0.023	0.002	-0.018	0.009	-0.005
	(0.045)	(0.032)	(0.050)	(0.052)	(0.035)	(0.034)	(0.041)	(0.033)	(0.035)	(0.036)	(0.041)	(0.038)
GDP gap	-1.055**	0.507	0.392	-0.199	0.403**	0.424**	0.277	0.411**	0.408**	0.423***	0.346	0.403***
	(0.437)	(0.343)	(0.394)	(0.345)	(0.195)	(0.166)	(0.272)	(0.169)	(0.187)	(0.161)	(0.241)	(0.123)
HighEduc	-0.224**	0.126	0.177*	0.067	-0.103	-0.103	-0.151	-0.039	-0.145***	-0.056	-0.063	0.002
	(0.092)	(0.075)	(0.103)	(0.080)	(0.073)	(0.070)	(0.098)	(0.096)	(0.053)	(0.066)	(0.108)	(0.093)
LowEduc	0.219**	0.185**	0.226**	0.079	0.070	0.027	0.076	0.064	0.117**	0.033	0.093*	0.047
	(0.093)	(0.067)	(0.092)	(0.065)	(0.055)	(0.051)	(0.060)	(0.055)	(0.051)	(0.048)	(0.056)	(0.055)
Young	-0.215*	-0.332***	-0.302*	-0.407	-0.762***	-0.618***	-0.969**	-0.631**	-0.734***	-0.552***	-0.665***	-0.397**
	(0.109)	(0.082)	(0.166)	(0.260)	(0.137)	(0.158)	(0.431)	(0.298)	(0.120)	(0.116)	(0.219)	(0.198)
Old	0.127*	0.193***	0.173*	0.418**	0.430***	0.347***	0.550**	0.197	0.413***	0.309***	0.377***	0.367**
	(0.065)	(0.046)	(0.094)	(0.193)	(0.077)	(0.090)	(0.244)	(0.361)	(0.067)	(0.065)	(0.123)	(0.156)
MAN	-0.201***	-0.393***	-0.346***	-0.223***	-0.218*	-0.345***	-0.248	-0.423***	-0.187**	-0.309***	-0.269***	-0.209***
	(0.057)	(0.057)	(0.074)	(0.071)	(0.110)	(0.110)	(0.155)	(0.110)	(0.080)	(0.075)	(0.098)	(0.062)
CONS	-1.243***	-1.064***	-0.765***	-0.578***	-0.350**	-0.520***	-0.506**	-0.278	-0.383**	-0.495***	-0.426**	-0.203
	(0.221)	(0.136)	(0.241)	(0.169)	(0.169)	(0.167)	(0.208)	(0.184)	(0.158)	(0.151)	(0.190)	(0.131)
TRADE		-1.268***	-1.132***	-0.505*		-0.384**	-0.492	-0.219		-0.357**	-0.386	-0.241
		(0.198)	(0.265)	(0.260)		(0.185)	(0.304)	(0.182)		(0.168)	(0.274)	(0.176)
ADM		-0.800**	-0.478	-0.256		-0.642**	-0.432	-1.118***		-0.658***	-0.461**	-0.590**
		(0.351)	(0.405)	(0.403)		(0.281)	(0.344)	(0.305)		(0.221)	(0.227)	(0.284)
Activity Rate		0.003	-0.058	0.201**		-0.120***	-0.058	-0.029		-0.109***	-0.068	0.005
		(0.088)	(0.106)	(0.093)		(0.037)	(0.044)	(0.052)		(0.037)	(0.047)	(0.049)
Urban		-0.103*	-0.085	-0.201***		-0.403***	-0.495**	-0.452*		-0.249***	-0.207***	-0.242***
		(0.051)	(0.062)	(0.064)		(0.130)	(0.206)	(0.241)		(0.047)	(0.065)	(0.052)
Minority		1.196***	1.287**	0.955***		1.090	0.785	0.801*		1.006**	0.912	0.867***
		(0.397)	(0.532)	(0.326)		(0.655)	(0.969)	(0.403)		(0.477)	(0.685)	(0.304)
Benefits			-0.033***				0.023				0.026*	
			(0.011)				(0.015)				(0.014)	
Wage				0.264				0.291				0.199
				(0.205)				(0.194)				(0.239)
Rsqrt.	0.57	0.77	0.81	0.80	0.55	0.74	0.69	0.64	0.64	0.78	0.81	0.78
R between					0.27	0.71	0.65	0.57	0.43	0.69	0.69	0.71
R within					0.88	0.90	0.93	0.92	0.88	0.90	0.93	0.91
Obs.	364	364	224	168	364	364	224	168	364	364	224	168

Note: The table reports OLS and Panel data estimates of the determinants of regional unemployment differentials. Robust standard errors are reported in parenthesis. Panel data estimators include year fixed effects

***1% significance level, ** 5 significance level and * 10% significance level.

and thus conclusions w.r.t those variables might not be as reliable as they include less observations. It should also be noted that the magnitude and significance of some of the right hand side variables change depending on the specification, signaling of potential heterogeneity of coefficients depending on the time period. For instance, the effect of output gap is insignificant during the first period (See T7:3b) while it becomes significant and increases in magnitude during the second period (See T7:4b). Similarly, the relative importance of the industrial composition (MAN and ADM) as well the minority proxy in regressions on unemployment level differentials might be driven by the period in consideration.

The results thus far indicate that the main determinants of within regional unemployment levels are related to demographic factors (age structure), industrial composition and fluctuations in output around long term trend. Three robustness specifications are provided in Table 8 in Appendix C. First of all, I consider the importance of industrial diversity (DIV index) as opposed to simple shares of employment within major sectors as well as concentration of particular industries relative to national average (LC measures) included to control for different impact contingent on the industry specialization[31]. Additionally, I employ GRP per capita growth as an alternative measure of demand driven differences in unemployment rates (T8:3, 4). Thirdly, I examine if the main conclusions are driven by outliers. More specifically, I remove district Sofia as it is consistently associated with extreme values for unemployment levels as well as some of the explanatory variables (T8: 5 & 6). Although overall conclusions do not change substantially, there are several differences that deserve some attention.

With respect to the demographic variables, the age structure has the same impact independent of the alternative specifications or the exclusion of Sofia district. However the impact of educational attainment is no longer insignificant with higher levels of human capital proving to be especially important when considering labor market dynamics outside the main economic region. Furthermore, the relative contribution of the industrial variables changes as diversification proxy is introduced. More specifically, higher relative concentration in the construction sector is the only industrial composition variable that is associated on average with lower unemployment levels. One possible explanation of the observed differences, compared to simple shares presented in Table 7, is the lower within regional variation in LC.

[31] See Section 2.3 for a discussion on the measures.

Concentration of manufacturing and trade shares relative to aggregate values changes very little within a district over the observed time period, therefore across regional differences might be better able to capture impact on unemployment levels. Furthermore, the proxy for industrial diversification is significant but does not have the expected sign. Theoretically, higher index value signifies higher concentration which in turn should lead to an increase in both structural and frictional unemployment all else constant. Nevertheless, in this case it is the opposite. It provides some evidence that unemployment might be industry specific and regions with high concentration in growing industries have relatively lower unemployment rates. In this case employment shares are sufficient to capture within regional variation in unemployment.

Furthermore, output gap might not be the most appropriate measure to estimate the impact of cyclical fluctuations due to the inherent pitfalls in its calculations (Cotis et al., 2004). Therefore I include growth of GDP per capita as an alternative measure. However, the variable is not significant in any of the specifications and coefficients for the rest of the variables do not change.

One possible explanation for the positive coefficient on output gap is the inflow of workers or net migration that leads to an increase in the labor force and high aggregate demand that is not capable of absorbing all new entrants. Based on Table 8: (2b), the impact of deviations from potential output are persistent in nature with one period lagged values showing significantly positive effect. In an attempt to explore whether the observed associations could be explained by migration flows, I look at correlation coefficients between output gap and net migration for the Bulgarian districts for the time period 2000-2013 as well as for two sub-periods (Pre and Post Crisis). Results are reported in Table 9 in Appendix C. Overall the table provides evidence that migration could explain at least part of the observed positive association between output gap and high unemployment rates on a regional level with positive correlation predominantly stronger during the second period. Nevertheless, the relationship seems to be driven by several main districts and it is not consistently strong.

5.3 Spatial effects

Increasingly, the empirical literature on regional labor market dynamics has tried to incorporate spatial effects when modeling variations in local unemployment rates. The underlying reasoning is related to the understanding that regional labor markets are not separate entities but are correlated across space, with outcomes dependent on developments within neighboring regions

(Lottman, 2012). Proponents of spatial econometric models argue that failure to adequately account for spatial interrelations might result in a serious bias and/or inefficiency of coefficient estimates in traditional panel data models even after allowing for individual specific intercepts (Anselin et al., 1998). Clustering of regional outcomes might be indicative of spatial interactions, i.e. labor market dynamics in one region depends on its neighbors. Figure 2 in Appendix C plots a map of regional unemployment levels for the year 2013 which provides evidence on the possibility of spatial interactions. Formal tests are presented in Table 10. It shows Moran`s I indices for 2001 and 2013. The index measures the degree of spatial correlation in selected variables (unemployment, human capital, industry composition etc.) based on data. The global values (T10: (A)) serve as an indicator of the extent of overall similarity within the variables of interest among neighboring districts while local indices show the main districts that contribute to the observed clustering. Corresponding p-values indicate that unemployment rates in one region are positively associated with corresponding rates in its neighbors[32] with ten main districts that contribute to the observed spatial correlation. Additionally, Moran`s Is are calculated for several of the explanatory variables indicating that these spatial effects should also be modeled.

There are three main spatial panel models which differ in their assumptions and the way they model spatial correlations. The general econometric specification takes the following form:

$$Un_{it} = \rho W * Un_t + \beta X_{it} + \theta W * X_t + \mu_i + \delta_t + \varepsilon_{it} + \sigma W * \nu_t \quad (3)$$

With three main spatial models specifications frequently employed in empirical research:

1. Durbin Model (SDM) : $\rho \neq 0$; $\theta \neq 0$ and $\sigma = 0$
2. Lag (autoregressive) Model (SAM): $\rho \neq 0$; $\theta = 0$ and $\sigma = 0$
3. Error model (SEC): $\rho = 0$; $\theta = 0$ and $\sigma \neq 0$

Where W is the spatial weight matrix and X is a vector of explanatory variables, which in the Bulgarian case includes the independent variables in the MAIN as well as FULL model as described above. Due to the strict non-missing value and balanced panel data restrictions required for correct evaluation of the spatial model, the analyses are restricted to the time period

[32] Neighbors are districts that share a common border, i.e. the matrix employed to estimate spatial effect is the binary Queen contiguity matrix.

2004-2013[33]. With respect to spatial models selection, Elhorst (2010) proposes to always start with an estimate of the spatial Durbin model as it captures several channels of spatial correlation in the data. Furthermore, Global Indices presented in Table 10 evince that spatial dependence is also present in the independent variables rendering spatial interactions between one region`s unemployment rates and the characteristics of its neighbors plausible. Therefore I provide the results of the Durbin model in Table 11 below.

As spatial panel data models should be ultimately compared to their non-spatial counterparts, the first column presents the panel FE estimator for the restricted time period, denoted as specification (1). The main conclusions are largely consistent with the ones presented in Table 7 (2b) with unemployment rates being particularly sensitive to changes in output gap, age and industry structure. The rest of the columns present the results for the fixed effects SDM estimator (Specification (2))[34].The first two columns include the results for the coefficients as presented in equation (3). The parameter ρ denotes the degree of spatial dependence in the data. Within a spatial model the interpretation of the rest of the parameter estimates becomes more complicated as a change in the independent variable within region *i* has an impact on both the region itself, as well as on unemployment rates in neighboring regions, introduced by the W*X relationship, which furthermore influences region *i* through the spatial dependence W*Un (Elhorst, 2010). Those complex interrelations are referred to as "feedback loops". The scale of this reinforcing effects naturally depend on the weight matrix that represents the connectivity between regions, the strength of the spatial dependence, rho (ρ), and the estimates β and θ. Nevertheless, the results of the spatial models are more accurately evinced through the estimates for direct, indirect and total impact (LaSage & Pace, 2009). The estimates for the direct impact show the effect of a change in the independent variables in *i* region on unemployment rates in the region itself. The estimates can be compared to the coefficients in the non-spatial models, nevertheless, they might also incorporate the feedback effects, i.e. changes in independent variables in region *i* lead to changes in unemployment in neighboring regions. The indirect effect provides a measure of spillover effects and represents the average impact of a unit change in the independent variables in all neighboring regions (other than *i*) on unemployment in region *i*, while the total effect is the

[33] Due to data limitation on LF activity rates

[34] STATA package "xsmle" has been utilized for the spatial analyses. For further reference see Belotti(2013).

sum of both direct and indirect effect and signals total impact of a unit change in any of the independent variables on unemployment levels in region *i*.

Based on the results presented in Table 11, the direct effect estimates provide a measure of the main variables within a region that have an impact on observed unemployment rates with GDP gap, demographics and industrial composition exerting statistically significant influence. A one pp. increase in output gap within a region is on average associated with a 0.698pp. higher unemployment rate. The value is larger in magnitude than the non-spatial model coefficients, which indicates the presence of reinforcing feedback effects, which go through the impact on unemployment in neighboring regions. As far as the age structure of the population within a spatial model is concerned, an increase in the share of younger people is associated on average with a 0.676pp. lower unemployment rates while the opposite is true for the share of old people. In terms of the summary statistics related to the industry structure, the general conclusions remain similar to the ones obtained under the non-spatial FE estimator, with the sign of the coefficients consistently negative while magnitude slightly differs between the employed econometric specifications. Furthermore, an increase in the activity rate is negatively associated with unemployment in both the spatial and non-spatial model, while the effect of urbanization is larger in absolute value and statistically significant only in the non-spatial model.

Overall, a comparison between the coefficient estimates of the non-spatial and the direct effect measures in the spatial model, provide evidence that feedback effects might be important to consider in econometric specifications that model regional variation of labor market variables in Bulgaria. The coefficients that determine the sensitivity of regional unemployment rates to changes in local characteristics might differ in magnitude as well as significance contingent on the model estimated. Thus the impact of spillover effects might need to be adequately accounted for in regional studies of economies in transition.

The spillover effects due to variation in explanatory variables are captured by the indirect coefficient measures which indicate which characteristics of the neighboring regions contribute the most to the observed spatial dependence. Based on the results, changes in the output gap and activity rate in neighboring regions are associated with statistically significant changes in unemployment in the one region under consideration. In absolute magnitude, the spillover effects

are even larger than the direct effects, with a 1pp. increase in the output gap in neighboring regions leading on average to 1.568pp. increase in unemployment in the one region under consideration with the corresponding value for an increase in the activity rate in neighboring regions equal to -0.281pp. The estimates for the indirect effect lead me to conclude that common output variation within clusters of regions could determine labor market outcome. Alternatively gdp shocks in one region might not only affect unemployment in the region itself but might also have large spillover effects on its neighbors, rendering spatial models the better option to explore within regional variation.

Nevertheless, the results for the spatial model should be treated with caution as they provide only preliminary evidence on spatial interactions while facing some limitations. First of all, the Moran`s Is values that indicate spatial clustering are largely dependent on the weight matrix employed. I have utilized only the binary Queen contiguity matrix, that is entirely based on regions neighbors being determined by a common border. Despite the fact that it is frequently employed in econometric research, LaSage and Pace (2009) suggest the calculation of alternative matrices for robustness of the results. Second of all, spatial panel data models might need a larger dataset to provide more informative results (Elhorst, 2010). Utilization of a limited number of panel-year observations might even introduce additional bias. In the case for Bulgaria, a further segregation beyond district level data is likely to lead to better insights w.r.t. regional variation in unemployment outcomes. Last but not least, within the spatial framework, the channels through which spillover effect of neighbors have an impact on one region`s unemployment rates are not explicitly clear. Therefore, although the presence of spatial interactions seems to be an important consideration in the analyses of regional labor market outcomes, the current level of data availability does not permit a more extensive study of those interactions.

Table 11. **Spatial model Regional Unemployment, 2004 - 2013**

	(1)	(2)				
	FE	Main, β	θ,ρ	Direct	Indirect	Total
EMP growth	-0.062	-0.058***	0.068	-0.052	0.084	0.031
	(0.161)	(0.021)	(0.063)	(0.020)	(0.096)	(0.108)
GDP gap	0.588***	0.580***	0.849***	0.698**	1.568***	2.267***
	(0.161)	(0.128)	(0.252)	(0.143)	(0.325)	(0.394)
HighEduc	-0.046	-0.068	0.221	-0.043	0.309	0.265
	(0.066)	(0.047)	(0.148)	(0.060)	(0.265)	(0.303)
LowEduc	0.054	0.051	0.078	0.063	0.149	0.212
	(0.047)	(0.047)	(0.087)	(0.046)	(0.149)	(0.165)
Young	-0.709***	-0.656***	-0.286	-0.676***	-0.756	-1.432**
	(0.226)	(0.186)	(0.364)	(0.186)	(0.589)	(0.695)
Old	0.376	0.539***	-0.437	0.573***	-0.368	0.205
	(0.266)	(0.197)	(0.425)	(0.192)	(0.640)	(0.679)
MAN	-0.291**	-0.256***	-0.145	-0.283**	-0.362	-0.646
	(0.106)	(0.092)	(0.234)	(0.111)	(0.385)	(0.453)
CONS	-0.300**	-0.351***	0.207	-0.343***	0.112	-0.230
	(0.144)	(0.113)	(0.326)	(0.130)	(0.532)	(0.634)
TRADE	-0.183	-0.163	-0.006	-0.164	-0.106	-0.270
	(0.161)	(0.153)	(0.216)	(0.184)	(0.355)	(0.491)
ADM	-0.760***	-0.608***	-0.164	-0.637**	-0.522	-1.159
	(0.274)	(0.232)	(0.396)	(0.259)	(0.681)	(0.881)
Activity Rate	-0.120***	-0.120**	-0.099	-0.139***	-0.218*	-0.357**
	(0.042)	(0.046)	(0.068)	(0.049)	(0.114)	(0.141)
Urban	-0.315*	-0.328*	0.579	-0.280	0.673	0.392
	(0.183)	(0.185)	(0.425)	(0.214)	(0.723)	(0.857)
Minority	0.308	0.298	0.378	0.035	0.585	0.621
	(0.374)	(0.456)	(0.591)	(0.362)	(0.880)	(1.018)
Rho			0.356***			
			(0.085)			

Note: Panel data FE (1) and SDM estimates (2) of the determinants of regional unemployment, 2004-2013. Robust standard errors reported.Year fixed effects are included.

***1% significance level, ** 5 significance level and * 10% significance level.

5.4 Long term unemployment

The last panel data specification that I consider is the regional determinants of long term unemployment. Based on the results in Section 5.2 regional unemployment is consistently higher for relatively larger proportion of old people in the working age group. I thereby argued that old

people might not be as flexible as younger ones or lack the appropriate skills which in turn would induce longer spells of unemployment. In an attempt to see of this might be the case I look at regional determinants of long term unemployment, whereas long term is defined as people registered in a labor office for more than a year. Unfortunately, the limited public information enables model estimation over a relatively small period of time, namely from 2008-2013. Table 12 in Appendix C provides the results for both non-spatial and spatial panel data models. As the substantially smaller dataset might lead to additional bias of coefficient estimates in spatial regressions[35], I derive the main conclusions based on the panel data FE model. To preserve the coherence of the presented results thus far, I nevertheless report estimates of the spatial Durbin model.

The results do not show that long term unemployment is predominantly driven by the presence of larger share of older people. The coefficients that remain significant at conventional levels are the ones on output gap, share of young people, and from the industry mix variables, those on Manufacturing and Trade. The sensitivity of long term unemployment to the regional characteristics do not change in sing, with larger output gap associated with consistently high levels of long term unemployment all else constant, while a 1pp. increase in the share of younger people, regional levels of manufacturing and trade, is associated with a fall in long turn unemployment of 0.258, 0.321 and 0.468pp respectively. Given the high degree of correlation between overall and long term unemployment[36], the similarity between the results is not surprising. Despite of that, the significant positive association between long term unemployment and output gap is far from intuitive. Positive output shocks are likely to induce larger than usual regional in migration which could explain higher overall unemployment level, nevertheless, it is not apparently clear what is the mechanism through which positive shocks could lead to persistently higher long term unemployment. One possible explanation could be related to structural problems. If targeted regional funds[37] tend to disproportionately focus on underdeveloped regions with relatively worse labor market performance, larger shocks in aggregate demand that lead to positive output gaps could also explain the positive association. In

[35] See Elhorst (2003)
[36] 0.93 for the period 2009-2013.
[37] For instance after becoming part of the European Union.

this case the reverse causality problem could hamper correct evaluation of determinants of labor market outcomes and alternative specifications should be considered.

6. Overview and discussion of results

The analyses on regional labor market outcomes within Bulgaria for the period 2000-2013 has presented several major conclusions with respect to recent developments. Preliminary examination of labor market trends on a local level has revealed that despite the recent financial crisis, unemployment rate observed in its aftermath was on average much lower for the majority of regions compared to values in the pre-crisis period, thereby showing signs of aggregate improvement in labor conditions for the Bulgarian economy. Despite aggregate trends of declining unemployment levels within regions, there is little evidence on convergence towards a common equilibrium. The majority of districts with the highest rates in the beginning of the sample period remained consistently among the worst labor market performers throughout. Nevertheless, the majority of districts witnessed an absolute decline in the number of registered unemployed with neighboring districts shown to follow common trends and two major geographical groups with contrasting developments are observed. Data consistently reveals that the Northern part of the country is lacking behind its Eastern and Southern peers with significant decline in both number of persons employed and working age population for both pre and post crisis period. It signals the presence of regional problems of long lasting nature.

The results for the panel data estimator indicate that overall there is no evidence of consistent demand driven impact on unemployment levels. Employment growth is insignificant in all specifications, while the significance on output gap depends on the specification employed and does not have the expected sign. Trends in migration are considered as a possible explanation of the divergence between theoretical predictions and observed outcomes, nevertheless, results do not seem to indicate persistently strong associations between output gaps and migration flows. In fact consistently positive net migration over the whole period is observed in four main districts[38] with little to no relation to respective output gap trends. Furthermore, the insignificance of the coefficient based on an alternative specification that incorporates output growth directly seems to suggest that there is no relation between aggregate demand factors and unemployment rates on a

[38] Sofia, Plovdiv, Varna and Burgas

regional level. Thus the relationship between output gap and unemployment can be mechanical association rather than a robust economic pattern.

Based on the results, structural factors, such as demographics and industrial composition, are relatively more important in explaining unemployment trends within Bulgarian districts. More specifically, the age structure proxies are significant in all specifications, however expected signs are not related to usual theoretical predictions and results in comparable studies. In fact larger share of young people is associated with lower unemployment rates while the opposite is true for the share of elderly. This might evince negative future trends in Bulgarian districts. Young people are in general more flexible and, based on the econometric results, better able to find jobs or have lower durations of unemployment. Alternatively, they might not be as risk averse and find it easier to move to areas where unemployment is low and more opportunities are present. In that case the direction of the relationship would go the other way around. In fact, a closer look at migration flows in relation to age structure[39] reveals that young people are disproportionately more likely to migrate as they represent a larger share of net migration. This could have important implications for future labor market developments in relatively worse performing regions. Higher share of elderly might signal deteriorating human capital conditions which in turn would hamper private investment and exacerbate local economic development. This could lead to a vicious circle and certain regions could fall in the so called "poverty traps" which would prove a serious problem to utilizing the productive capacity and potential of certain regions.

With respect to the other explanatory variables, it is interesting to note that the minority proxy is significant in some of the specifications, indicating that the share of minorities could help explain fraction of unemployment level differentials both across and within regions. In general minority groups (members of the Roma community) are mostly associated with low levels of human capital and relatively quick drop out of the labor force for females due to marriage and childbearing. The fact that it is significant even after the inclusion of educational proxies and labor force participation rates might indicate that this groups find it relatively more difficult to find jobs overall. Finally, the industrial structure is highly significant in all specifications, signaling of its relative importance and major implications on regional labor dynamics. Larger

[39] Results not shown. Available upon request.

share of manufacturing, construction and trade are being consistently associated with lower levels of unemployment. The fact that the magnitude and significance of some of the right hand side variables change depending on the specification reveals potential heterogeneity of coefficients depending on the time period employed. Furthermore, the exclusion of Sofia region, points that educational attainment consideration might be more relevant when considering labor market dynamics outside the main economic regions.

Recent emphasis on the importance of modeling spatial clustering has urged me to consider spatial panel data models as well. A formal index of spatial correlation evinces that unemployment rates in one region are positively associated with corresponding rates in its neighbors with ten main districts that contribute to the observed spatial correlation. Additionally, Moran`s Is calculated for the set of explanatory variables indicate that spatial clustering of regional characteristics is also present and thus spatial effects should also be best modeled including both independent and dependent variables rationalizing the implementation of the spatial Durbin model. Overall, a comparison between the coefficient estimates of the non-spatial and the direct effect measures in the spatial model, provide evidence that feedback effects might be important to consider in econometric specifications that model regional variation of labor market variables in Bulgaria. The coefficients that determine the sensitivity of regional unemployment rates to changes in local characteristics might differ in magnitude as well as significance contingent on the model estimated. Thus the impact of spillover effects might need to be adequately accounted for in regional studies of economies in transition. Nevertheless, the limited dataset serve as a reminder that conclusions based on the spatial model might introduce additional bias and adequate consideration of spatial clustering could require the inclusion of additional observations.

Policy considerations

With respect to labor market intervention and regulation, there are two opposing views. One of them favors complete liberalization with no state intervention. This was the prevalent political view at the onset of market liberalization in the beginning of 1990s. The belief that markets will quickly eliminate inefficiencies is still being challenged today with the formulation of views that issues related to employment and unemployment should be at the center of political agenda. The results of the analyses point to the conclusion that as regions were disproportionately affected by

the economic shock the divergence in development trends seems difficult to eliminate. Improvement in unemployment rates in worse performing regions relative to pre-crisis levels was not due to better economic conditions but a result of outmigration that reduced the total labor force and the number of both employed and unemployed. This signals that political debate should focus on reconsidering policies that aim at addressing region-specific problems. The lacking regions are characterized by high share of elderly in the labor force as well as lower share of industry and trade related activities. The outmigration of younger individuals further signals that the predominant problem of those regions might in fact be the lack of opportunities. The higher share of elderly people points that the inflexibility of the labor force could be an issue. Elderly workers have specific set of skills that might not be as easily converted to the demands of a changing labor market. Individual investment in training to keep oneself competitive on the labor market could be especially costly for older people. On the other hand, on the job training directly increases the cost of labor and employers might not find it economically beneficial to invest in such activities. Furthermore, relatively limited mobility and higher risk aversion means lower probability of migration towards regions that might provide better employment opportunities. Generally, one way to increase employment for this age group is to introduce government financed training programs in line with business specific needs. Unfortunately, Bulgarian experience so far has proved that they are highly inefficient and in fact do not lead to any improvement in employment prospects for unemployed. Alternatively, the policy makers could consider ways to decrease labor cost for low productivity workers so as to increase their employment value. One way to achieve it is to introduce minimum wages that are region specific. In fact the minimum wage has increased by more than 400% for the period 1999 -2013[40]. While this was introduced with the intention to provide minimum standard of living for the persons employed, failure to take into account of labor productivity issues which might in fact be region specific, would result in higher unemployment. Furthermore, it is more likely to increase the share of the grey economy. Additionally, the introduction of the minimum insurance threshold in 2003 has led to the growing concern that regions lagging behind are more severely affected due to the generally lower productivity of the labor force. Therefore, introduction of region specific minimum wages or insurance thresholds might induce higher investment and improvement in employment prospects for the most vulnerable.

[40] The minimum wage was approx.. 31 EUR at 1999 and close to 182EUR in 2013.

Last but not least, the lack of strong economic association between educational attainment and unemployment rates is largely in contrast to the usually cited political argument that higher education alone would decrease unemployment. What matters is ultimately the successful match between education and demand for specific skills. In fact, recent surveys of employers have shown that there is an increasing gap between specialists demanded in jobs related to engineering and technology and respective graduates. Effectively, this results in an oversupply of specialist in certain areas with limited ability for those to be successfully included in the labor market or the so called "underemployment problem". Policy intended at that direction would prove beneficial in the long run as it would decrease skills mismatch in the aggregate economy. One way to deal with the problem is to introduce incentives in the form of scholarships or larger educational grants to students in respective disciplines in demand.

7. Conclusion

The current thesis focused on labor market trends in the Bulgarian economy observed during the most recent period. Despite the fact that more than twenty years have passed since the introduction of the capitalist system, the large structural changes associated with the rapid liberalization of the economy has introduced large concussions on local labor markets. Regions did not manage to respond uniformly to the altered conditions and as a result unemployment levels started to diverge relative to national average. Despite an overall decline in unemployment rates in the post crisis period, there is little evidence on convergence towards a common equilibrium. The majority of districts with the highest rates in the beginning of the sample period remained consistently among the worst labor market performers. This naturally raises the question of the main determinants of regional labor market outcomes within the country. The theoretical and empirical literature has managed to identify several main factors that could largely explain the persistent divergence. By utilizing a panel district level data that spans from 2000 to 2013, the thesis has tried to determine the relative contribution of supply and demand side forces towards the observed unemployment rates. The main results are largely in line with existing studies with demographics and industrial structure variables exhibiting consistently significant coefficients. Contrary to conventional theory, however, the share of elderly in the labor force is associated with worse labor market outcomes. The relatively limited flexibility and higher training costs offers a potential explanation. Educational attainment seems to be

particularly important once influential observations are removed. The results for the panel data estimator indicate that overall there is no evidence of consistent demand driven impact on unemployment levels. Employment growth is insignificant in all specifications, while the significance on output gap depends on the specification employed and does not have the expected sign. The data also shows the presence of clustering of districts with persistently similar unemployment rates and regional characteristics which necessitates the inclusion of an econometric specification that allows for spatial spillovers. Some preliminary evidence of statistically significant spatial interactions also point at these conclusions.

The thesis faces several considerable shortcomings that need to be addressed in future endeavors on labor developments. First of all, it should be noted that despite increased reliability in statistical datasets, there still might be problems with adequate reporting of employment and unemployment figures, especially at the beginning of the sample period. For instance, employment in the grey economy is still a pervasive factor within the Bulgarian economy and could be relevant in explaining the lack of sensitivity between employment growth and unemployment rate. On the other hand, hidden unemployment could result in measurement error in the dependent variable. As long as it is random it does not directly impact the consistency of the estimated regression coefficients however, this is not the case if certain regions are consistently more likely to be characterized by hidden unemployment. Second of all, the empirical results are essentially a summary of major associations and the direct causality channels cannot be readily assessed. As a result it cannot be argued that government interventions that aim at stimulating, for instance, manufacturing in certain regions would automatically result in lower unemployment rates. The results, nevertheless, show that the districts with relatively older labor force are associated with worse outcomes. The underlying reasons for such outcomes could enable a better policy framework and region specific intervention. Last but not least, the empirical analyses do not include minimum wage considerations which could have a better explanatory power in unemployment equations. More specifically, a regulation in the year 2003 has introduced minimum insurance thresholds for employees depending on industrial occupation and position within the firm. They essentially serve as industry specific minimum wages which determine the labor costs. Therefore regional

labor costs and labor productivity interaction could serve as a key missing predictor of regional divergence.

There are several suggestions for future research that could lead to a better explanation of observed labor market dynamics. The inclusion of broader administrative unit classification that is beyond NUTS-3 level would certainly provide a better explanation of unemployment determinants and would furthermore be more appropriate for assessing spillover effects on a regional level. Presently, data restrictions hamper a more detailed and informative study of such nature with only a limited set of labor market variables publicly available at lower classification clusters. Furthermore, the inclusion of private investment and regional infrastructure considerations in relation to unemployment levels would lead to additional insights given the recent increase in foreign direct investment flows and large projects for infrastructural improvement. In this line of thought, it would be interesting to assess the contribution and effectiveness of EU structural funds on a regional level so that timely measures are introduced in case observed results and targets are not aligned. Future research should also aim at addressing labor cost considerations as suggested above. On a broader scale it might be relevant to include a larger regional dataset on other countries in transition so as to compare relative outcomes or to take into account of possible interactions that extend to regions of neighboring countries so as to assess whether peripheral location could provide insights into unemployment differentials.

Reference list

Adanu, K. (2005). A cross-province comparison of Okun's coefficient for Canada, in *Applied Economics*, 37, pp. 561–570.

Ajilore, T. & Yinusa, O. (2011). An Analysis of Employment Intensity of Sectoral Output Growth in Botswana. Southern African Business Review, Vol.15, No 2

Baker, D. & Schmitt J. (1999). The Macroeconomic Roots of High Unemployment: The Impact of Foreign Growth. Economic Policy Institute.

Beck, N., & Katz, J.N. (1995). What to do (and not to do) with Time-Series Cross-Section Data. The American Political Science Review, Vol.89, No3, pp 634-647

Beck, N. (2001). Time-Series-Cross-Section Data: What Have We Learned in the Last Few Years? Annual Review Political Science, No.4 pp271-293

Blanchard, O. &. Jimeno J. (1995). Structural Unemployment: Spain versus Portugal. The American Economic Review, Vol. 85, No. 2.

Blanchard, O. J. & Katz, L. (1995). Regional Evolutions. Brookings Papers on Economic Activity.

Bratoeva-Manoleva, S. H. (2009). Bulgarian Labor Market in the context of the European Union Membership. VFU "Chernorizets Hrabar"

Boeri, T. & Scarpetta, S. (1996). Regional Mismatch and the Transition to a market economy. Labor Economics No3, pp.233-254

Bornhorst, F., & Commander, S. (2006). Regional Unemployment and its Persistence in Transition Countries. *Economics of Transition*. Vol.14 (2), pp. 269 – 288

Bronars, S.G. & Jansen D. (1987). The Geographical Distribution of Unemployment Rates in the U.S. A Spatial-Time Series Analysis. Journal of Econometrics, Vol.36, pp.251-279

Brown, S., & Sessions, J. (1997). A profile of UK Unemployment: Regional versus demographic Influences. *Regional Studies*, Vol 31.4, pp. 351-366.

Brunello, G., Lupi, C., & Ordine, P. (2001). Widening Differences in Italian Regional Unemployment. Labor Economics, Vol. 8, pp. 103 – 129

Crivelli, E., Furceri, D., & Toujas-Bernate, J. (2012). Can Policies Affect Employment Intensity of Growth? A Cross-Country Analysis. International Monetary Fund. Working Paper Series No. 12/218

Dopke, J. (2001). The Employment Intensity of Growth in Europe. Kiel Institute of World Economics, Kiel Working Paper No. 1021

Ehrenberg, R. G., & Smith, R.S. (2009). Modern Labor Economics: Theory and Public Policy. 11th Edition. Pearson Education, Inc.

Elhorst, J. P. (2003). The Mystery of Regional Unemployment Differentials: Theoretical and Empirical explanations. *Journal of Economic Surveys*, Vol. 17 No5. pp. 709-748.

Erber, G. (1994), Verdoorn's or Okun's Law? *German Institute for Economic Research Discussion Paper* No. 98. Berlin.

Euranet Plus BNR Discussion. (2014) "На първо място сме в ЕС по процент на млади хора, които живеят с родителите си." Available at: http://bnr.bg/euranetplus/post/100291221/na-parvo-masto-sme-v-es-po-procent-na-mladi-hora-koito-jiveat-s-roditelite-si

Evans, P. & McCormick, B. (1994). The New Pattern of Regional Unemployment: Causes and Policy Significance. The Economic Journal, Vol. 104. No.424, pp. 633-647

Flaig, G., & Rottmann H. (2000). Input Demand and the Short- and the Long-Run Employment Thresholds — An Empirical Analysis for the German Manufacturing Sector. *CESifo Working Paper* No. 264.

Garibaldi, P., Makovec, M., & Stoyanova, G. (2001). From Transition to EU Accession. The Bulgarian Labor Market During the 1990s. World Bank Technical Paper No.4

Herwartz, H. & Niebuhr. A. (2007). Determinants of the (un)employment intensity of growth: Evidence from a cross section of EU regions.

Hristoskov, Y. (1994). Unemployment and Labor Market in Bulgaria. Center for the Study of Democracy.

Islam, I. & Nazara, S. (2000), Estimating employment elasticity for the Indonesian economy, *ILO Technical Note*, Jakarta

Islam, R. (2010) "Pattern of Economic Growth and its Implication for Employment" in C.Banerjee, A. Dasgupta and R. Islam (eds.), *Development, Equity and Poverty: Essays inHonour of Azicur Rahman khan*, Delhi and New York: Macmillan India and UNDP

Izraeli, O., & Murphy K. (2003). The effect of industrial diversity on state unemployment rate and per capita income', *Annals of Regional Science*, vol. 37, iss. 1, pp. 1-14.

Kangasharju, A. & Pehkonen, J. (2001). Employment-output link in Finland: evidence from regional data, in *Finnish Economic Papers*, 1, pp. 41-50.

Kapsos, S. (2005). The Employment Intensity of Growth: Trends and Macroeconomic Determinants. International Labor Office. Employment Strategy Papers.

Kessler, E., E. Mack, & T. Grubesic. (2007). Indices of Industrial Diversity and Regional Economic Composition. *Growth and Change* Vol. 38 No. 3, pp. 474–509

Kieselbach, T. (2003). Long term Unemployment among young people: The risk of social Exclusion. Americal Journal of Community Psychology, Vol. 32, No.1/2

Korobilis, D. & Gilmartin, M. (2010). On regional unemployment: an empirical examination of the determinants of geographical differentials in the UK . The Rimini Centre for Economic Analyses

Kosfeld, R. & Dreger, C. (2006), Thresholds for Employment and Unemployment. A Spatial Analysis of German Regional Labour Markets 1992-2000, *Papers in Regional Science*, 85, 523-542.

Landmann, O. (2004). Employment, productivity and output growth, *ILO Employment Strategy Papers*, No 7.

Lopez-Bazo, E., del Barrio, T., & Artıs, M. (2000) The geographical distribution of unemployment. Paper presented at the XXV European Meeting of the Regional Science Association, Barcelona.

Lopez-Bazo, E., del Barrio, T., Artıs M. (2002). The Regional Distribution of Spanish Unemployment: Spatial Analysis. Papers in Regional Science. Vol. 81, pp. 365–389

Maddah, M. (2013). An Empirical Analysis of the Relationship between Unemployment and Theft Crimes. International Journal of Economics and Financial Issues Vol.3, No.1, p. 50 – 53

Martin, R. (1997) Regional Unemployment Disparities and their Dynamics. Regional Studies, Vol.31, No.3, p.237-252

Marston, S.T. (1985) Two Views of the Geographical Distribution of Unemployment. The Quarterly Journal of Economics, Vol.100, No.1, pp. 57-79

Malizia E.E., Ke S. (1993) The Influence of Economic Diversity on Unemployment and Stability. *Journal of Regional Studies* 33, pp. 221-235.

McPherson, S.H. & Flores, O. (2012). Provincial Disparities and Structural Unemployment in Canada. *Regional and Sectoral Economic Studies*, Vol. 12, Issue 1, pages 25-42

Misra, S. & Suresh, A.K. (2014). Estimating Employment Elasticity of Growth for the Indian Economy. RBI Working Paper Series, Department of Economic and Policy Research

Mizutani, F., Mizuno, K. & Nakayama, N. (2003). Industrial Diversity and Metropolitan Unemployment Rate. *Annals of Regional Science* 40(1): 157–172.

Moon, S., Sekwat, A. & Stanley, R.E. (2004) Testing Assumptions of Pooled Time Series Cross Sectional Regression Analysis through ARIMA and interrupted Time Series Modeling. Institute of Government, Tennessee State University

Nesporova, A. (2002). Why Unemployment remains so high in Central and Eastern Europe. International Labor Office Geneva. Employment Paper N. 43

Nickell, S. (1997). Unemployment and Labor Market Rigidities: Europe versus North America. The Journal of Economic Perspectives, Vol. 11, No. 3, pp. 55 -74

Oelgemoller, J. (2013). Economic and Employment Growth in Germany – The Sectoral Elements of Verdoorn`s Law with regional Data. CAWM – Discussion Paper No. 63

Orlandi, F. (2012). Structural Unemployment and its Determinants in the EU Countries. European Commission. Economic Papers 455

Padalino, S. & Vivarelli, M. (1997). The Employment Intensity of Growth in the G-7 Countries. *International Labor Review* 136: 199–213;

Pehkonen, J. (2000) Employment, unemployment and output growth in booms and recessions:time series evidence from Finland, 1970-1996, Applied Economics, 32:7, 885-900

Perugini, C. (2008). Employment Intensity of Growth in Italy. A Note Using Regional Data. *Regionla and Sectoral Economic Studies*, Vol. 9(1)

Petersen, M. (2005) Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. NBER

Pfaff, T. (2013). Clustered Correlation and Few Clusters: Some Practical Advice for Adjusting Standard Errors and Inference. Center for Interdisciplinary Economics

Phelps, E. (1998). Structural Slumps: The Modern Equilibrium Theory of Unemployment, Interest, and Assets.

Poterba, J.M & Summers, L.H. (1995). Unemployment Benefits and Labor Market Transitions: A Multinomial Logit Model with Errors in Classification. *Review of Economics and Statistics* 77,

pp. 207–216.

Saget, C. (2000). Can the Level of Employment be Explained by GDP Growth in Transition Countries? (Theory versus Quality of Data). LABOUR, Vol.14 (4) pp. 623 -644.

Saridakis, G., & Spengler, H. (2009). Crime, deterrence and unemployment in Greece: A panel data approach. DIW Berlin.German. Institute for Economic Research

Seyfried, W. (2007). A Cross – Country Analysis of the Employment Intensity of Economic Growth. *Southwest Business and Economics Journal*.

Tang, C.F. &. Lean, H.H. (2009). New evidence from the misery index in the crime function. Economics Letters, 102, 112-115.

Taylor, J. & S. Bradley. (1997). Unemployment in Europe: A Comperative Analyses of Regional Disparities in Germany, Italy and the UK. KYKLOS, Vol. 50, pp. 221 – 245

Vega, S.H. & Elhorst, J.P. (2012). Modeling Regional Labor Market Dynamics in Space and Time. Papers in Regional Science, Vol. 93, No. 4

Veracierto, M. (2008). On the Cyclical Behavior of Employment, Unemployment and Labor Force Participation. Journal of Monetary Economics 55, p.1143 – 1157

Appendix A. Data

Variable	Definition	Source
Employed	Persons aged 15 and over who work for the production of goods and services for at least an hour during the period and get paid for their work.	Statistical Yearbooks, NSI Regional Statistics database
Unemployed	Persons aged 15-74 who do not work but are actively seeking employment during the observed period	Statistical Yearbooks, NSI Regional Statistics database
Labor Force	Economically active population. Persons aged 15 and over who invest or offer labor to produce goods and services. LF includes both employed and unemployed.	Statistical Yearbooks, NSI Regional Statistics database

Female % LF	Share female in Labor Force to Labor Force	NSI Regional Statistics Database
Activity Rate	The ratio of Labor Force to population in the same age category	Statistical Yearbooks
AR Female	Female activity rate	NSI Regional Statistics Database
Employment Rate	The ratio of number of employees to population in the same age category.	Statistical Yearbooks
WA pop	Working Age population. Population that falls within the age category 15-64	Statistical Yearbooks
Female % WA	Share female to total working age population	Own calculations
Participation Rate	Share of economically active population over total WA pop. Excludes those who are part of WA but not in labor force – students, early retirees, homemakers)	Own calculations
Unemp LO	Regional unemployment rate equal to number of unemployed over total labor force. The number unemployed is based on registration at regional labor offices.	Statistical Yearbooks, NSI Regional Statistics database
Unemp LF	Regional unemployment rate based on annual labor market surveys	Statistical Yearbooks, NSI Regional Statistics database
Young	Share of population aged 15-24 to WA population	Own calculations
Old	Share of population aged 50-64 to WA population	Own calculations
High Educ	Share of population aged 25-64 years old having completed tertiary education	Official Inquiry NSI Regional Office - Sofia
Low Educ	Share of population aged 25-64 with primary or lower education	Official Inquiry NSI Regional Office - Sofia
Agriculture	Share of employed in Agriculture	Own calculations
Industry & Cons	Share of employed in Industry and Construction	Own calculations
Services	Share of employed in Services	Own calculations
Private	Share of employed in private sector	Own calculations
Public	Share of employed in public sector	Own calculations
Public Adm	Share of employed in public administration.	Own calculations
DIV index	Herfindahl Index of industrial diversification.	Own calculations

Urban	Share of urban population to total population within a region.	Statistical Yearbooks
Density	Population per sqr. km.	Statistical Yearbooks
Minorities 1	Share of births with mother under 20 years of age over total number of births.	Own calculations
Minorities 2	Share of births with mother under 20 years of age over population aged 15 -19.	Own calculations
CPI	Harmonized Index of Consumer Prices, 2005 =100.	NSI Economic Statistics, Inflation and CPIs
GVA	Gross Value Added. Available aggregates for Agriculture, Industry and Services	Official inquiry, NSI Regional Statistics database
GRP	Gross Regional Product	Official inquiry, NSI Regional Statistics database
Output gap	Percentage of deviation of real GDP from logarithmic time trend. Calculated based on Hodrick -Prescott filter.	Own calculations

Appendix B. Variables selection procedure

The number of regional labor variables that have been collected in an attempt to analyze the most important determinants of variation in unemployment rates is 25. As the number of district-year observations is relatively limited, I decided to employ a model specification methodology introduced by Lottman et al. (2012). They divide the variables in three groups based on their relative contribution in similar regional labor market studies. The first group includes the factors that are most frequently employed as explanatory variables, namely measures of labor demand, such as employment and output growth, as well as factors summarizing labor supply: demographic structure and educational attainment. Furthermore, employment shares in manufacturing and construction are also included as they are frequently found to predict regional variations. The second group of variables includes additional explanatory variables such as population density, participation rate, share of private sector as well as agriculture and trade. The third group includes all other measures related to industrial classification as well as a measure of urbanization and the minority proxy. The final selection of the variables is achieved through

minimization of AICC and BIC measure as follows. Variables of Group 1 are included in all specifications. Model 1 (M1) in the table below provides AICC and BIC for the MAIN specification. M2 includes all variables from both Group 1 and 2 with the additional variables from the second group that minimize AICC and BIC presented in bold. M3 includes all variables from Group 1 and 3 as well as the selected variables in the second step. The final specification determined based on the procedure can be found in the last column.

Table 3. Variable selection procedure

Groups of variables analyzed				**Variables selection**		
Group1	**Group 2**	**Group 3**	**M1**	**M2**	**M3**	**Final Model**
EMP growth	Density	Urban	EMP growth	Density	**Activity Rate**	EMP growth
GDP gap	Activity Rate	Minority	GDP gap	**Activity Rate**	**ADM**	GDP gap
MAN	ADM	EDUC	MAN	**ADM**	**TRADE**	MAN
CONS	TRADE	SOCIAL	CONS	**TRADE**	**Urban**	CONS
LowEduc	AGRI	EGW	LowEduc	AGRI	**Minority**	LowEduc
HighEduc	PRIVATE	TRANS	HighEduc	PRIVATE	EDUC	HighEduc
Young	Female	FIN	Young	Female	SOCIAL	Young
Old		RE	Old		EGW	Old
		HOTELS			TRANS	Activity Rate
					FIN	ADM
					RE	TRADE
					HOTELS	Urban
						Minority
		AICC	2370	**2288**	**2280**	2280
		BIC	1480	**1470**	**1471**	1471

Appendix C. Additional Tables and Graphs

Table 2. **Mean Values according to Unemployment Quartiles**

	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Unemp LO	8.27	12.33	15.42	19.49
Female	47.72	49.39	46.97	46.97
Young	20.33	21.67	20.97	20.51
Old	21.39	23.48	22.47	23.42
High Educ	22.81	17.70	16.98	15.84
Low Educ	22.53	26.71	29.83	32.97
Agriculture	2.42	4.27	6.08	6.28
IND & Cons	38.84	41.18	37.87	33.78
Services	37.10	30.68	29.22	28.81
Private	70.92	68.35	66.71	60.95
ADM	4.59	5.37	5.95	7.16
DIV index	0.16	0.16	0.16	0.14
Empl gr	1.14	-0.23	-0.05	-0.85
GDP gr	2.60	1.84	1.16	0.27
GDP cap gr	3.48	2.98	2.70	2.37
Activity Rate	52.09	49.27	48.55	47.41
Urban	76.20	65.82	61.45	55.38
Density	196.19	56.71	53.28	48.03

Note: The table provides mean values for several variables of interest according to different unemployment quartiles. All varialbes, except for DIV index, are in percentage points. Detailed definitions are available in ApendixA.

Table 4. Unit-Root Test for Stationarity

	no trend	trend		no trend	trend		no trend	trend
Unemp LO	0.00	0.07	MAN	0.00	0.00	Empl g	0.00	0.00
Female	0.00	0.00	CONS	0.00	0.00	GDP g	0.00	0.00
Young	1.00	1.00	ADM	0.01	0.00	GDP cap g	0.00	0.00
Old	1.00	0.00	TRADE	0.00	0.00	Urban	1.00	0.83
High Educ	0.76	0.00	Act Rate	0.00	0.00	Minority	0.00	0.00
Low Educ	0.11	0.00	DIV index	0.00	0.00			

Note: The table reports p-values for Levin-Lin-Chu unit-root test for panel data (xtunitroot llc). *Ho* : Panels contain unit roots; *Ha* : Panels are stationary.

Table 5. Basic Correlations Main Variables

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	**Unempl**	**1.00**																
2	**Emp growth**	0.01	**1.00**															
3	**Output gap**	-0.22	0.19	**1.00**														
4	**LowEduc**	0.52	0.13	-0.01	**1.00**													
5	**Young**	0.10	0.15	0.07	0.26	**1.00**												
6	**Old**	0.04	-0.06	0.04	-0.16	0.71	**1.00**											
7	**Female**	-0.06	0.03	0.07	-0.02	0.88	0.92	**1.00**										
8	**Act rate**	-0.49	0.09	0.09	-0.26	-0.11	-0.16	0.00	**1.00**									
9	**MAN**	0.04	0.02	0.05	0.14	0.09	0.04	-0.02	-0.12	**1.00**								
10	**CONS**	-0.51	-0.03	0.22	-0.13	-0.06	-0.12	0.03	0.46	-0.40	**1.00**							
11	**TRADE**	-0.58	-0.07	0.01	-0.43	-0.21	-0.05	0.03	0.41	-0.54	0.38	**1.00**						
12	**ADM**	0.16	-0.20	0.15	0.07	-0.06	0.11	-0.01	-0.26	-0.24	-0.13	-0.11	**1.00**					
13	**Urban**	-0.56	0.11	0.00	-0.71	-0.05	0.03	0.07	0.35	-0.21	0.23	0.52	-0.40	**1.00**				
14	**Minority**	0.26	-0.08	-0.02	0.18	-0.03	-0.04	-0.09	-0.22	-0.03	-0.11	0.04	0.03	-0.16	**1.00**			
15	**Benefits**	-0.65	-0.14	0.54	-0.35	-0.18	0.02	0.01	0.35	-0.07	0.51	0.48	0.37	0.08	-0.04	**1.00**		
16	**DIV index**	-0.11	0.05	0.05	0.03	0.05	0.00	-0.03	0.00	0.93	-0.34	-0.40	-0.28	-0.04	-0.13	-0.03	**1.00**	
17	**GDP cap gr**	0.03	0.25	0.49	0.04	0.15	0.10	0.12	0.07	0.06	-0.09	-0.05	-0.08	0.03	-0.04	-0.04	0.07	**1.00**

Note: The table reports pairwise correlations of regional variables over the period 2000-2013.

***Table 6*. Labor Market, pre- & post- crisis developments**

	North West									
	Vidin		**Vratsa**		**Montana**		**Pleven**		**Lovech**	
	t=1	t=2	t=1	t=2	t=1	t=2	t=1	t=2	t=1	t=2
Employed	34.2	33.9	67.4	67.9	55.7	53.8	103.6	103.0	56.0	54.5
Unemployed	9.8	6.3	14.7	7.3	10.0	7.7	14.0	11.4	9.3	4.8
Pop WA	66.6	57.2	125.4	112.5	94.9	85.3	178.8	161.6	91.7	82.4
Activity Rate	40%	44%	43%	45%	44%	46%	44%	47%	46%	47%
Unemp LO	22%	17%	20%	16%	22%	17%	16%	13%	15%	14%
	North Central									
	Veliko Tarnovo		**Gabrovo**		**Ruse**		**Razgrad**		**Silistra**	
	t=1	t=2	t=1	t=2	t=1	t=2	t=1	t=2	t=1	t=2
Employed	99.8	99.2	56.2	52.3	93.2	99.6	42.7	43.7	45.1	43.5
Unemployed	16.4	13.4	5.6	3.9	19.8	10.5	12.4	10.7	10.5	6.7
Pop WA	174.1	163.8	81.4	72.5	159.2	149.0	87.1	79.4	83.1	74.5
Activity Rate	46%	48%	50%	50%	50%	51%	45%	49%	47%	47%
Unemp LO	15%	10%	8%	6%	14%	9%	22%	16%	18%	16%
	North East									
	Varna		**Dobrich**		**Targovishte**		**Shumen**			
	t=1	t=2	t=1	t=2	t=1	t=2	t=1	t=2		
Employed	185.6	195.2	67.3	76.8	41.8	46.5	65.5	69.7		
Unemployed	29.2	20.5	17.6	13.2	11.4	6.7	18.6	22.8		
Pop WA	288.3	299.5	129.3	120.0	81.4	75.8	123.0	116.6		
Activity Rate	55%	54%	47%	54%	46%	49%	49%	57%		
Unemp LO	11%	7%	17%	11%	26%	17%	20%	16%		
	South West									
	Blagoevgrad		**Kyustendil**		**Pernik**		**Sofia**		**Sofia cap**	
	t=1	t=2	t=1	t=2	t=1	t=2	t=1	t=2	t=1	t=2
Employed	144.5	150.7	56.5	56.0	53.9	55.4	100.1	101.3	540.9	634.9
Unemployed	9.9	11.7	9.4	7.3	7.1	4.8	11.0	7.2	52.6	38.7
Pop WA	211.0	209.5	91.5	82.5	76.4	80.4	152.2	148.4	795.3	844.4
Activity Rate	55%	58%	48%	49%	48%	51%	49%	50%	56%	61%
Unemp LO	10%	12%	12%	12%	11%	9%	13%	11%	3%	3%
	Soth Central									
	Kardzhali		**Pazardzhik**		**Plovdiv**		**Smolyan**		**Haskovo**	
	t=1	t=2	t=1	t=2	t=1	t=2	t=1	t=2	t=1	t=2
Employed	62.4	55.0	102.4	112.6	264.0	286.5	50.5	50.8	96.0	98.5
Unemployed	4.9	3.0	17.5	16.8	25.7	26.6	15.5	11.4	15.8	13.5
Pop WA	102.8	98.1	184.3	175.2	437.5	430.8	85.8	78.6	159.9	151.6
Activity Rate	49%	44%	47%	53%	47%	52%	57%	57%	48%	51%
Unemp LO	14%	13%	18%	14%	11%	8%	19%	17%	13%	10%
	South East									
	Burgas		**Sliven**		**Yambol**		**Stara Zagora**			
	t=1	t=2	t=1	t=2	t=1	t=2	t=1	t=2		
Employed	153.8	166.5	66.0	72.2	54.7	54.1	136.0	139.4		
Unemployed	23.2	16.4	17.0	13.7	9.0	7.7	13.2	9.6		
Pop WA	260.9	264.7	127.1	119.7	87.2	78.8	218.8	207.1		
Activity Rate	50%	52%	47%	52%	49%	53%	48%	50%		
Unemp LO	11%	8%	16%	14%	17%	13%	12%	8%		

Note: Average values for selected labor market variables for 28 administrative regions; t=1 denotes 2000-2007 period and t=2 :2008-2013. Employed, Unemployed and Pop WA are in thousands, Activity Rate and Unemploument are in percentage points.

Table 8. **Alternative specifications**

	Panel data, FE				w/o SOFIA	
	1	**2**	**3**	**4**	**5**	**6**
EMP growth	-0.001	-0.007	0.009	0.013	-0.029	-0.026
	(0.031)	(0.031)	(0.034)	(0.035)	(0.028)	(0.031)
GDP gap	0.464**	0.457***			0.419***	
	(0.190)	(0.153)			(0.142)	
GDP gap (t-1)		0.444***			0.416***	
		(0.139)			(0.136)	
GDP cap gr			-0.011	-0.004		-0.004
			(0.007)	(0.005)		(0.005)
HighEduc	-0.081	-0.113	-0.087	-0.107	-0.148**	-0.144**
	(0.080)	(0.071)	(0.081)	(0.072)	(0.063)	(0.064)
LowEduc	0.103**	0.063	0.104**	0.067	0.047	0.049
	(0.046)	(0.049)	(0.047)	(0.052)	(0.044)	(0.047)
Young	-0.682***	-0.545***	-0.695***	-0.585***	-0.502***	-0.527***
	(0.131)	(0.150)	(0.137)	(0.158)	(0.169)	(0.173)
Old	0.382***	0.306***	0.393***	0.329***	0.283***	0.297***
	(0.073)	(0.085)	(0.076)	(0.089)	(0.096)	(0.098)
DIV Index	-0.476*	-0.619***	-0.456*	-0.594***	-0.675***	-0.654***
	(0.237)	(0.180)	(0.237)	(0.189)	(0.174)	(0.183)
LC MAN	0.004	-0.001	0.003	-0.006	0.021	0.018
	(0.024)	(0.019)	(0.024)	(0.019)	(0.013)	(0.014)
LC CONS	-0.033**	-0.042***	-0.031**	-0.040***	-0.033***	-0.031***
	(0.013)	(0.011)	(0.012)	(0.012)	(0.009)	(0.010)
LC TRADE		-0.047		-0.049	-0.019	-0.019
		(0.028)		(0.029)	(0.027)	(0.028)
LC ADM		-0.026		-0.022	-0.018	-0.015
		(0.018)		(0.019)	(0.018)	(0.019)
Activity Rate		-0.093**		-0.096**	-0.104***	-0.109***
		(0.034)		(0.037)	(0.033)	(0.036)
Urban		-0.415***		-0.391***	-0.335**	-0.313**
		(0.127)		(0.132)	(0.146)	(0.151)
Minority		0.973		1.095	0.917	1.023
		(0.673)		(0.665)	(0.659)	(0.646)
Rsqrt.	0.64	0.73	0.65	0.76	0.744	0.748
R between	0.44	0.72	0.44	0.73	0.638	0.641
R within	0.88	0.90	0.88	0.89	0.910	0.906
Obs.	364	364	364	336	351	352

Note: The table reports FE estimates of the determinants of regional unemployment differentials. Robust standard errors in parenthesis. Year fixed effects included.

***1% significance level, ** 5 significance level and * 10% significance level.

Table 9. Output gap & Net Migration flows

	Total	t=1	t=2		Total	t=1	t=2
BLG	0.527	0.506	0.931	RUS	0.147	0.065	0.196
BUR	0.301	0.341	0.318	SHU	0.28	0.268	0.265
DOB	0.418	0.537	-0.008	SLS	-0.35	-0.51	-0.26
GAB	0.446	0.371	0.516	SLV	0.219	-0.28	0.662
HAS	0.308	-0.05	0.559	SML	0.034	-0.08	0.09
KDZ	-0.288	-0.535	0.302	SOF	0.148	0.19	0.712
KUS	-0.116	-0.368	-0.378	CAP	-0.455	-0.882	-0.401
LOV	0.533	0.513	0.558	STZ	0.109	0.264	0.072
MON	-0.456	-0.606	-0.311	TGV	0.085	-0.285	0.31
PDZ	0.457	0.492	0.418	VAR	0.483	0.475	0.838
PRN	0.06	0.003	0.496	VTV	0.154	-0.079	0.113
PVN	0.419	0.609	0.262	VID	-0.189	0.1908	-0.556
PVD	-0.001	-0.455	-0.225	VTZ	0.311	0.724	-0.486
RAZ	0.376	-0.619	0.688	YAM	0.28	-0.128	0.603

Note: Pairwise correlation of net total migration flows (relative to WA) and output

Figure 2. Regional Unemployment Levels, 2013

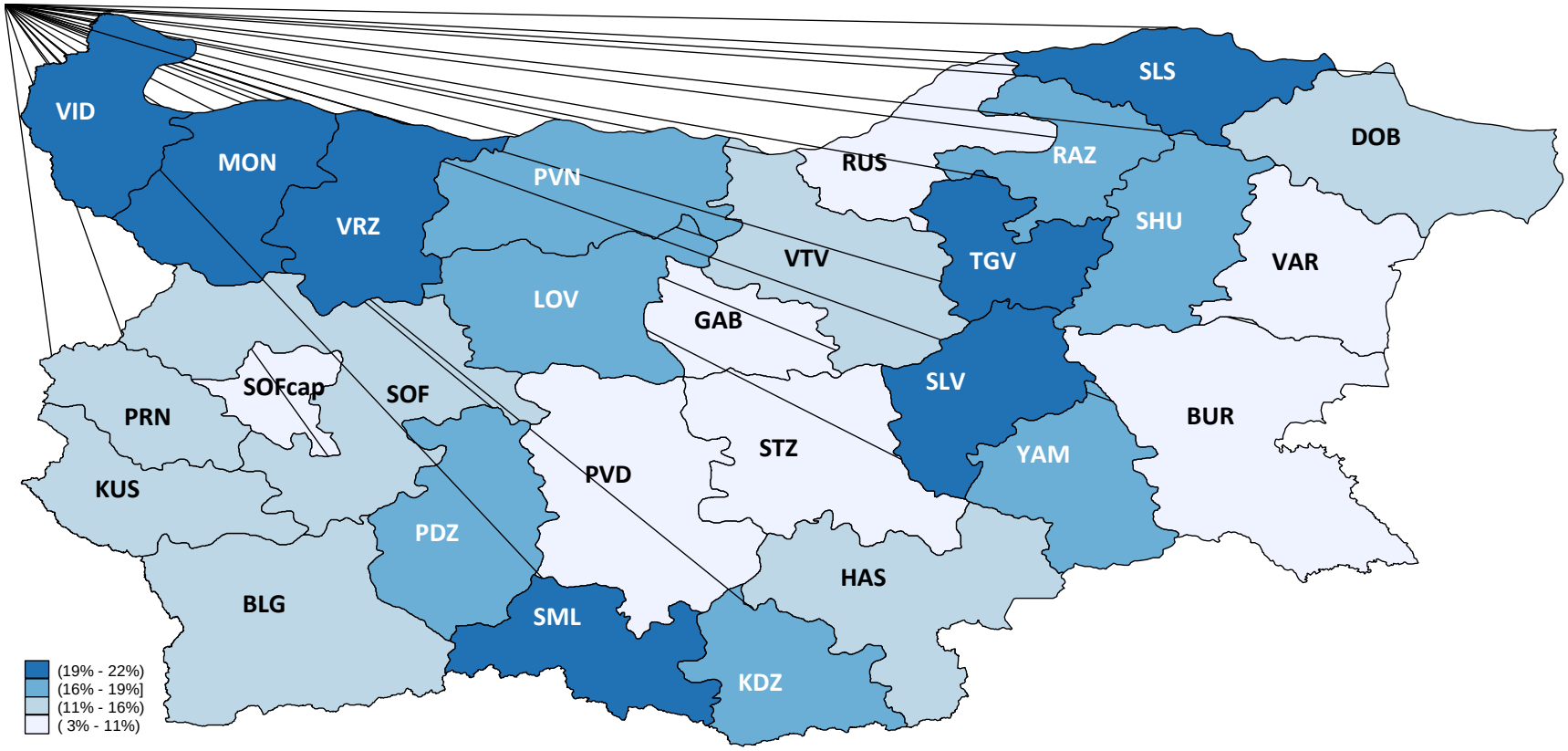


Table 12. **Long Term Unemployment rates, 2009-2013**

	Non spatial	**Spatial regression, SDM**				
	FE	Main, β	θ,ρ	Direct	Indirect	Total
EMP growth	-0.035	-0.033	0.057	-0.035	0.048	0.012
	(0.044)	(0.041)	(0.067)	(0.034)	(0.066)	(0.078)
GDP gap	0.513***	0.485***	-0.249	0.505***	-0.269	0.235
	(0.177)	(0.15)	(0.408)	(0.175)	(0.399)	(0.394)
HighEduc	-0.034	-0.002	0.153**	-0.002	0.137*	0.135*
	(0.032)	(0.035)	(0.065)	(0.037)	(0.089)	(0.079)
LowEduc	0.041	0.052*	0.133**	0.049*	0.123**	0.173***
	(0.031)	(0.029)	(0.058)	(0.028)	(0.053)	(0.067)
Young	-0.258***	-0.295***	-0.109	-0.274***	-0.115	-0.390
	(0.087)	(0.078)	(0.290)	(0.076)	(0.261)	(0.249)
Old	0.190	0.187*	-0.173	0.218**	-0.191	0.027
	(0.160)	(0.113)	(0.266)	(0.107)	(0.266)	(0.277)
MAN	-0.321***	-0.324***	-0.132	-0.318***	-0.107	-0.425*
	(0.070)	(0.084)	(0.280)	(0.089)	(0.266)	(0.225)
CONS	0.025	-0.014	-0.173	-0.028	-0.114	-0.142
	(0.108)	(0.102)	(0.456)	(0.098)	(0.503)	(0.515)
TRADE	-0.468***	-0.539***	-0.453	-0.537***	-0.357	-0.894***
	(0.141)	(0.141)	(0.306)	(0.155)	(0.265)	(0.320)
ADM	0.438	0.288	0.154	0.251	0.165	0.416
	(0.321)	(0.343)	(0.457)	(0.289)	(0.414)	(0.430)
Activity Rate	0.016	0.001	-0.039	0.001	-0.028	-0.028
	(0.022)	(0.022)	(0.046)	(0.018)	(0.042)	(0.046)
Urban	-0.188	-0.200*	0.261	-0.209*	0.269	0.059
	(0.114)	(0.117)	(0.257)	(0.108)	(0.275)	(0.287)
Minority	0.285	0.122	-0.051	0.145	-0.106	0.038
	(0.242)	(0.247)	(0.520)	(0.279)	(0.522)	(0.719)
Rho			-0.096			
			(0.121)			

Note: Panel data FE and SDM estimates of the determinants of LT regional unemployment. Robust standard errors reported.Year fixed effects are included. ***1% significance level, ** 5 significance level and * 10% significance level.

Appendix D. Index of spatial correlation

Moran`s I is the most frequently employed measure of spatial correlation in the econometric literature. It is often preferred to alternatives as it is relatively straightforward and easy to interpret. The index (I) is measured through the following formula:

$$I = \frac{n}{S_o} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} z_i z_j}{\sum_{i=1}^{n} z_i^{\,2}}$$

Where *n* equals the number of observations, $w_{i,j}$ is an element in the weight matrix that describes the spatial relation between regions i and j; z is the deviation of the variable to be tested from the overall sample mean and S_0 is the sum of all elements in the weight matrix equal to $\sum_i \sum_j w_{i,j}$. The index alongside the p-value indicate whether the observed pattern is clustered, random or dispersed with positive significant values pointing at high—high values and/or low-low values spatial distribution.

Table 10. (A) Global Index

	2001		2013	
	Index	**p-value**	**Index**	**p-value**
Unemp	0.349	0.001	0.252	0.012
Output gap	0.032	0.293	0.237	0.012
Young	0.265	0.009	0.078	0.175
Old	0.42	0.00	0.022	0.317
Low Educ	0.316	0.002	0.358	0.001
MAN	0.187	0.034	0.221	0.019
CONS	0.092	0.154	0.206	0.024
TRADE	0.029	0.472	0.175	0.046
Urban	0.162	0.054	0.134	0.086
Minority	0.184	0.032	0.175	0.038
Activity rate	0.063	0.211	0.186	0.039

(B)Local Indices, Unemployment

	2001		2013	
	Index	**p-value**	**Index**	**p-value**
RAZ	1.273	0.002	0.385	0.179
PRN	1.164	0.012	0.809	0.059
VID	1.572	0.047	1.75	0.033
CAP	1.898	0.002	1.389	0.017
SHU	0.555	0.033	0.151	0.281
TRG	1.25	0.001	0.194	0.283
MON	0.597	0.117	0.982	0.03
GAB	0.219	0.316	0.68	0.092
STZ	0.328	0.179	0.507	0.088
VRZ	0.13	0.356	0.633	0.072

Note:Moran`s I of spatial correlation. Null hypothesis: No spatial correlation exists