Correlates of Regional Unemployment in Regions of East-Central Europe

Jeffrey Parker
George Hay Professor of Economics
Reed College
3203 S.E. Woodstock Blvd.
Portland, Ore. 97202
(503) 517-7308
parker@reed.edu

Prepared for presentation at the annual meeting of the
Western Economic Association, International
Vancouver, British Columbia
June 2018

This research was initiated at the Economická Univerzita v Bratislave during a 2015 Fulbright grant. I am grateful to the Slovak Fulbright Commission and its executive director Nora Hložeková, and to the Fulbright Scholar program in general for making this opportunity possible. I wish to thank the many Slovak colleagues who helped me learn about the region's economy and labor markets, in particular Daniel Dujava, Martín Kahanec, Zuzana Polačková, Miroslav Štefánik, and Marek Radvanský. Finally, three extraordinary Reed student collaborators contributed greatly to this research: Helena Pedrotti, Samuel Olson, and Ryan Keithahn. Responsibility for remaining shortcomings rests with the author.

Correlates of Regional Unemployment in Regions of East-Central Europe

1. Introduction

Since emerging from Communism in the 1990s, the countries of East-Central Europe have followed differing political and economic paths. This paper focuses on differences in labor-market performance in the four countries known as the "Visegrád" (V-4) group after the historic Hungarian town in which leaders of Poland, Czechoslovakia, and Hungary met in 1991 facilitating a cooperative path toward EU membership.

Figure 1 shows the varying paths of the national unemployment rates of the four countries since the divorce of the Czech and Slovak Republics in 1993. Over the decade since the Great Recession, all four countries have seen a large increase in unemployment followed by a consistent recovery after 2013. However, the levels of unemployment, patterns of change, and especially their paths prior to 2008 differ widely.

Moreover, while there are important differences in the paths of the national unemployment rates, these aggregates mask large differences across regions within the countries. The map in Figure 2 shows 2008 unemployment rates in the regions of the V-4 countries. There is considerable within-country variation in all four countries, but perhaps most strikingly in Slovakia, where the Bratislava region (marked with a single black dot in Figure 2) had an unemployment rate of 3.6 percent while, just 100 miles to the east, the mountainous Banská Bystrica region of central Slovakia (marked with two dots) suffered with 18.2 percent unemployment.

This paper aims to explore the extent to which these differences in unemployment rates are associated with differences in variables measuring policies, demographics, labor-market characteristics, industrial composition, and distance from Western markets. The variables included in my models are able to explain over 70 percent of the variation in unemployment rates over time and across regions.

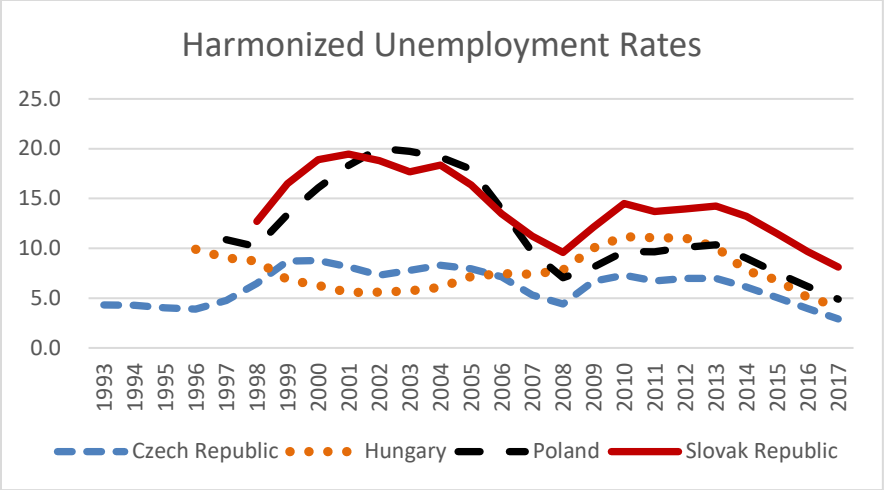

Figure 1. Visegrád harmonized national unemployment rates, 1993–2017

Source: OECD online iLibrary

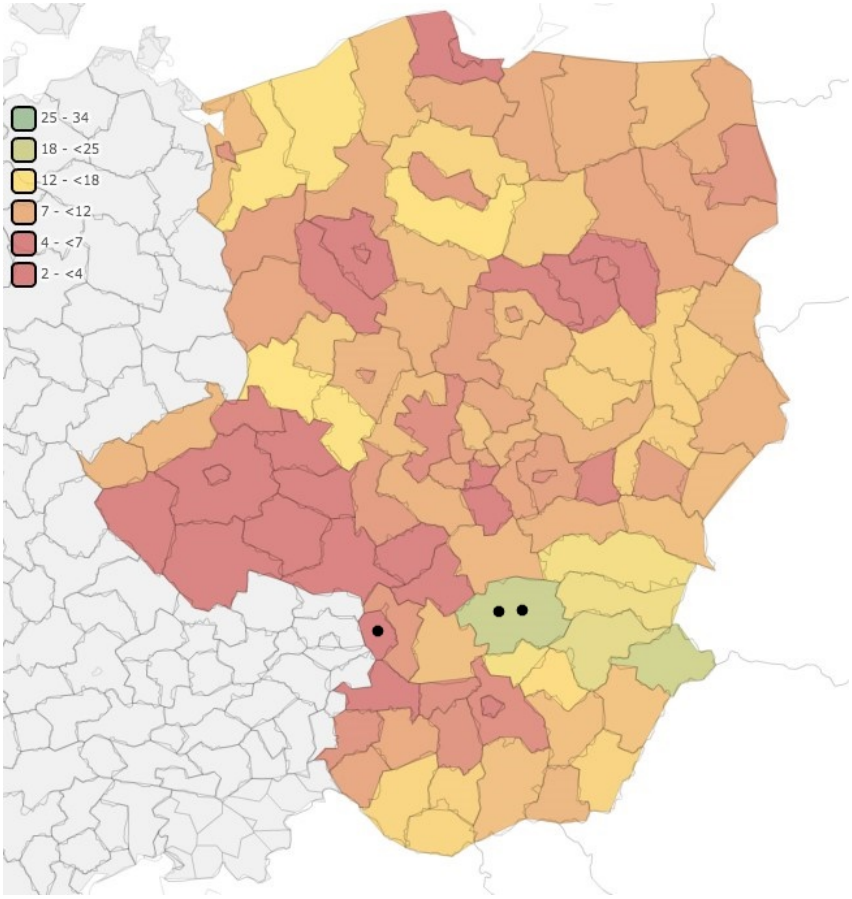

Figure 2. Unemployment in Visegrád regions, 2008

Source: OECD iLibrary

2. Brief Sketch of the Countries

The V-4 countries emerged from Communism following the collapse of the Soviet Bloc in 1989. As in other Communist countries, they retained a legacy of inefficient and loss-making state industries. Resource allocation had been based on decisions of central planners rather than markets. International trade was coordinated by COMECON, which was established by the Soviet Union to organize trade and industry across Communist Eastern Europe and a handful of other Communist countries. Trade with the market economies of Western Europe was severely limited, insulating these countries from effective competition with their western neighbors.

Privatization of industry proceeded rapidly after 1989 with private-sector employment rising from low levels (except in Poland) to roughly 60 percent of the economy by 1995. (See Borish and Noël (1996).) Some industries (agriculture, retail trade, and many services) had immediate local markets, but the transition was much more difficult for manufacturing enterprises that were suddenly deprived of their captive COMECON export markets and forced to compete with more efficient Western firms. Defense industries that had armed the Warsaw Pact were left without buyers at the end of the Cold War.

The Czech and Slovak Republics were created at the beginning of 1993 out of the former Czechoslovakia. Figure 1 shows the large and persistent difference in unemployment rates across the two newly separated states, the former of which nestles between Germany and Austria while the latter's longest borders are with Poland and Hungary.

The V-4 all joined the European Union (EU) in January 2004. They all use different currencies with only Slovakia having adopted the euro (in January 2009). All have been members of the Schengen Area since 2007, eliminating border checks between the countries and their western neighbors. All have distinct languages, although the Czech and Slovak languages are so similar that speakers can communicate without difficulty.[1]

In the nearly 25 years they have been EU members, all four economies have grown strongly. Figure 3 plots total real GDP and reflects a tripling of the Polish economy (from a

[1] As one local told me, all Slovaks speak Czech fluently due to the dominance of Czech television in Slovakia, but most Czechs speak very imperfect, but understandable, Slovak.

relatively depressed level) and increases of between 50 and 100 percent in the other three countries.

Neoclassical growth theory predicts that countries with similar characteristics should converge over time to similar levels of per-capita income. Table 1 shows that in terms of per-capita real GDP, Czechia remains the richest of the four countries, but two of the other three have converged strongly. Slovakia has nearly caught up, Poland has narrowed the gap from 41 percent to 18 percent, but Hungary has barely converged at all, falling below Slovakia and Poland in per-capita income.

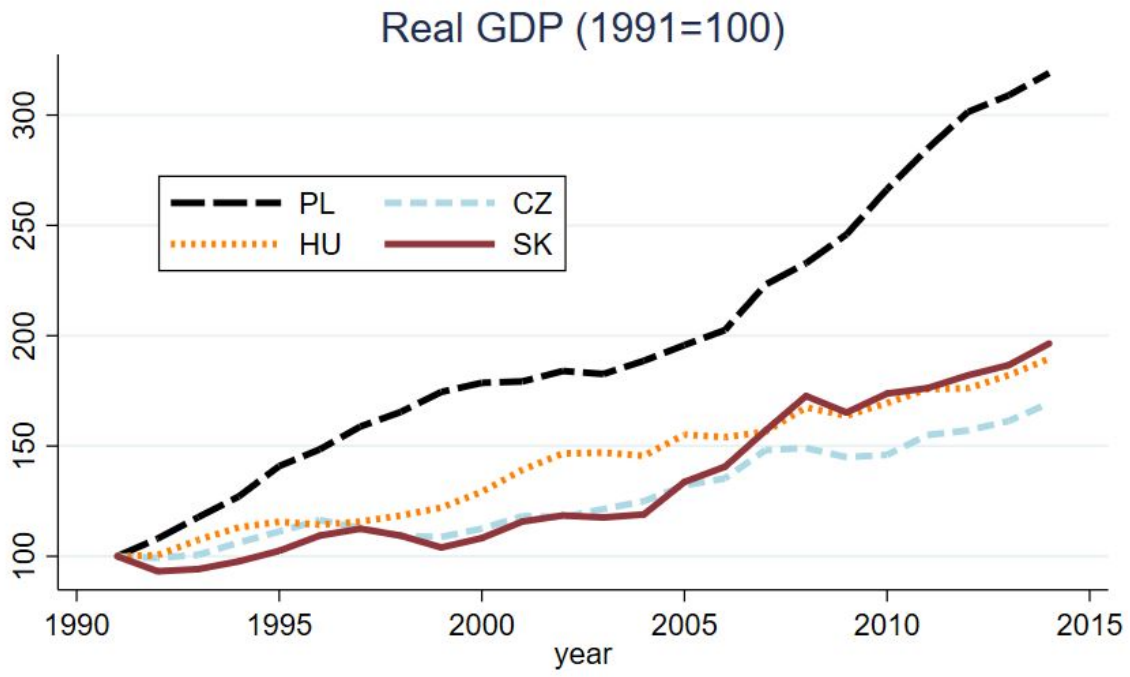

Figure 3. Real GDP growth since 1991 in Visegrád countries

Source: OECD iLibrary

Table 1. Per-capita real GDP, 1995 and 2017

	Czechia		Hungary		Poland		Slovakia	
	GDP per capita	Share of Czechia	GDP per capita	Share of Czechia	GDP per capita	Share of Czechia	GDP per capita	Share of Czechia
1995	18,835	---	14,753	78.3%	11,027	58.5%	13,102	69.6%
2017	31,858	---	25,756	80.8%	26,085	81.9%	29,957	94.0%

Note: GDP in 2010 U.S. dollars at purchasing-power parity
Source: OECD iLibrary

Each of the four Visegrád countries has a strong national identity, which has flourished after release from the forced alliances of the Warsaw Pact. Ethnic considerations contributed to the separation of Czechia and Slovakia and to the recent rise of populist governments hostile to immigration and refugees. Marginalized Roma ("Gypsy") minorities are spread unevenly across the region and are poorly integrated into the formal economy. Empirically, the estimated regional Roma population is strongly correlated with unemployment.

Figure 1 and Figure 3 both show that these countries have experienced strong business cycles since the 1990s. However, at shown by Table 2, the cycles have not been strongly synchronized. Poland and Slovakia experienced a strong growth slowdown around 2000, then rebounded strongly, but growth in Czechia and Hungary accelerated during this period. All four countries suffered slower growth following the Great Recession in 2008, but reductions in growth were larger in Czechia and Slovakia than in Hungary and Poland. Since 2013, all four countries have had moderate growth.

Table 2. Real GDP growth in 5-year intervals

	Czechia	Hungary	Poland	Slovakia
1995 – 1998	1.07	2.48	5.55	5.45
1998 – 2003	2.73	3.85	3.13	2.79
2003 – 2008	5.17	2.85	5.08	7.10
2008 – 2013	– 0.44	– 0.82	2.84	1.05
2013 – 2017	2.96	2.33	2.81	2.57

Note: Data before 1995 are considered less reliable by OECD, so the initial period is truncated to three years.
Source: OECD iLibrary

3. Modeling the determinants of unemployment

Models of unemployment typically separate so-called natural factors from cyclical movements. Natural unemployment variations are those, either over time or across geographic regions, that are due to microeconomic or structural characteristics of labor markets or policies. Cyclical movements are associated with short-run fluctuations in economic activity,

usually proxied by real GDP growth. In this section, I describe variables that were examined as potential correlates of regional unemployment.

Among the thousands of studies of unemployment rates, the ones most directly relevant to this study include Nickell, Nunziata, and Ochel (2005), Orlandi (2012), and Partridge and Rickman (1997).

Regional sample

The sample consists of sub-national regions of the V-4 countries. Data are drawn from the OECD iLibrary database, which includes data at two different levels of aggregation. "TL2" regions are large, with an average population of under 2 million and 35 total regions. "TL3" regions are a finer level of disaggregation, sometimes coinciding with the TL2 regions and sometimes breaking them down into smaller sub-regions. There are 108 TL3 regions with an average 2010 population of just under 700,000.[2]

Most of the key variables in the model are observed at the finer, TL3 level, but some (education, for example) are only available at the TL2 level. Some variables—in particular policy variables—do not vary at all across regions within a country.

Cyclical unemployment

I model the cyclical component of the unemployment rate as depending on the estimated GDP gap: the percentage difference between actual and potential output. Because unemployment is strongly countercyclical, the coefficient on the gap is expected to be negative. There are two key questions to be decided in determining a suitable business-cycle measure: (1) the level of aggregation at which to measure cycles and (2) the method to be used to decompose GDP into trend vs. cycle in calculating the gap. Because unemployment and output are jointly determined, there is a significant risk of endogeneity in including a measure of output in an equation predicting unemployment. However, unemployment generally lags the business cycle; the regressions reported below use a one-year lag of the business cycle indicator, which should reduce the likelihood of bias due to endogeneity.

Measuring the cycle at a low level of aggregation (country or even sub-national region) allows a more precise measure of fluctuations that apply to the individual region. However,

[2] These regions correspond to the NUTS-2 and NUTS-3 regions in the EU's Eurostat database. Several Polish TL2 regions lacked consistent data at the TL3 level; these regions were aggregated to the TL2 level, leaving a sample of 88 usable regional observations in the dataset.

it increases the risk of endogeneity because local shocks are likely to affect both unemployment and output strongly. Choosing a more aggregated measure dilutes the effects of local shocks on the cycle measure, but also makes it a less accurate measure of the macroeconomic cycles that apply to any particular region.

If business cycles are largely coordinated across a large area then a more aggregated measure can provide an accurate representation of macroeconomic cycles in the constituent regions. For example, it is common to use a U.S. national business cycle measure in modeling unemployment in individual states. However, using an analogous EU-wide measure seems inappropriate for the V-4 countries because of their unique characteristics as they privatized their economies and gradually integrated with Western Europe.

Two levels of aggregation are considered here. The first is the aggregate of the V-4 countries. However, because of the considerable heterogeneity in national cycles within the V-4, modeling at the individual-country level might be preferable. When using the V-4 gap, interaction terms between country dummies and the GDP gap variable are highly significant, principally due to the strong negative correlation between Hungary's cycle and those of the other three countries in the 2000s.

The second question, trend-cycle decomposition, can be addressed in two principal ways: structural estimates and purely statistical methods. The OECD estimates a "potential output" series and a corresponding series for the "output gap" (the percentage gap between actual and potential) for each country. These measures are based on simple growth-accounting methods that rely on national estimates of productivity and the utilization of labor and capital. Because a major contributor to the calculation of labor utilization (and therefore the GDP gap) is the unemployment rate, there are endogeneity concerns with this variable. If the level of aggregation is sufficient that a given region's unemployment rate has a negligible effect on the aggregate gap measure, then endogeneity may not be a serious problem.

An alternative is to use a purely statistical method for calculating a trend in log-output and computing the output gap as the deviation of log-output from its trend. Although a simple linear trend could be used, a more "flexible" trend such as the Hodrick-Prescott (H-P) filter seems more appropriate for these countries. [Hodrick and Prescott (1997)]

As one might expect, the two gap measures are positively correlated, as shown for the V-4 aggregate in Figure 4. Because the H-P filter fails to capture the recession in the early 2000s, the OECD gap measure is used in the analysis described below.

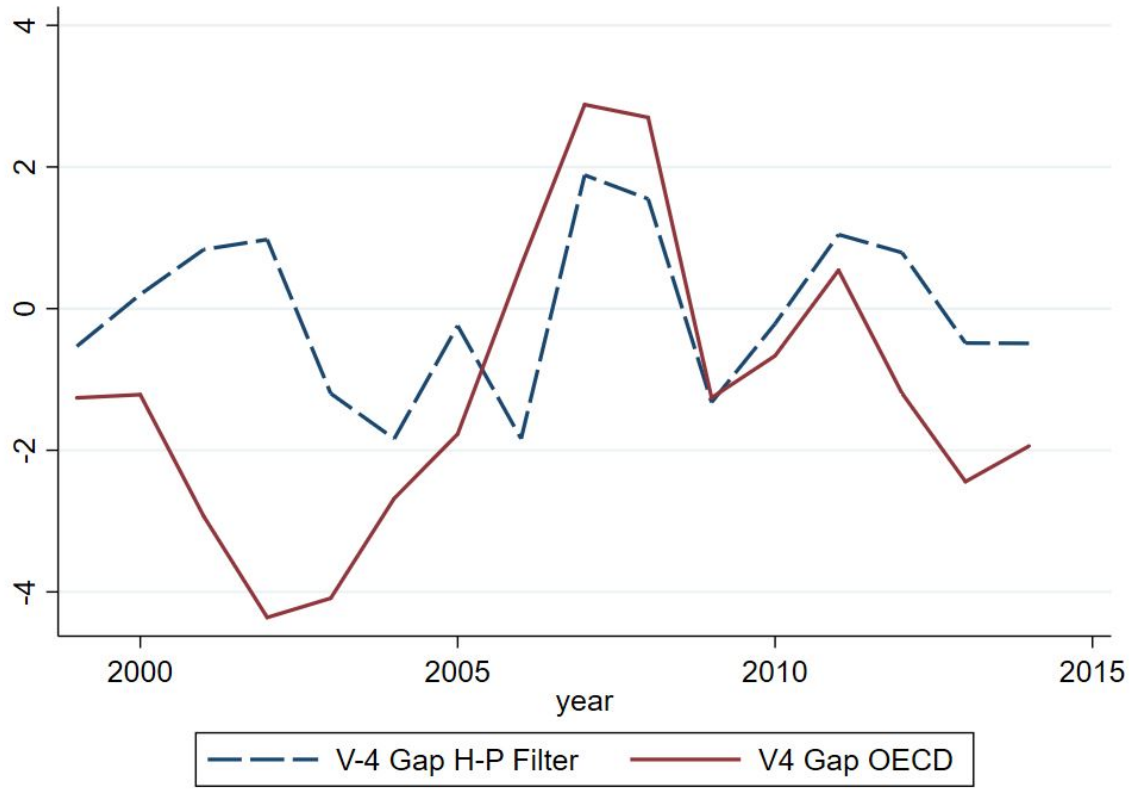

Figure 4. Comparison of V-4 GPD gap series

Industry composition

In any given year, the industry composition of a region may favor or disfavor its labor market. If the region's most prominent industries are booming, the local labor market is likely to be strong; if they are slack then the local labor market will probably suffer. I model the effect of industry composition by calculating a predicted employment growth rate for each region as a weighted average of the growth rates of V-4 employment levels in 11 ISIC4 industries, with weights given by the previous year's employment shares within the region. More precisely, the industry-based predicted employment growth rate for region r in year t is $g^{P}_{r,t} := \sum_{i} \frac{E_{i,r,t-1}}{E_{r,t-1}} g_{i,t}$, where $E_{i,r,t}$ is employment in industry i in region r in period t, $E_{r,t}$ is total period t employment in region r, and $g_{i,t}$ is employment growth in industry i in period t over the entire V-4. A measure similar to this is used in Partridge and Rickman (1997) to model unemployment rates in the U.S. states.

If the important local industries are growing rapidly averaged over the V-4, this should lower unemployment in a region, so I expect a negative sign for the industry-predicted employment growth variable. This variable seems to have greater impact in the model when entered with a one-year lag, which also helps limit any potential endogeneity.

Agriculture is known for low incidence of measured unemployment because of its high degree of self-employment and relative cyclical stability. To capture this effect, I include the percentage of a region's workers who work in agriculture as a separate variable with an expected positive sign.

Education and demographics

A higher level of education often allows a worker to qualify for a wider range of jobs. Educated jobs are often more stable than those requiring less education. I measure education with the percentage of workers with tertiary educational attainment and the percentage whose highest attainment is secondary education.[3] (These data are available in the OECD's regional innovation database at a coarse regional breakdown.)

I expect that regions with more-educated workers should have lower unemployment rates, other things equal. Thus, secondary-education attainment should have a negative sign; tertiary-education attainment should have a negative sign that is larger in absolute value than that of secondary education.

The age composition of the region may also affect unemployment rates, though the directions are not always clear. Higher youth population is likely to raise unemployment rates.

Policy variables

Labor-market policies in these countries are enacted on a national level, so there is typically no variation at all across regions. Moreover, they tend to be very stable over time, so these variables are somewhat collinear with country dummies when the latter are included in the model.

---

[3] "Tertiary education" here is ISCED codes 5–8, "short-cycle tertiary education" through "doctoral or equivalent level." "Secondary education" is ISCED codes 3–4, "upper secondary education" and "post-secondary non-tertiary education." The remaining portion of the population has "primary education" attainment, which is ISCED codes 0–2, "early childhood education" through "lower secondary education."

Taxes

The size of the tax wedge in labor markets may affect the incentives for work in ways that affect equilibrium unemployment. I expect that higher marginal taxes on labor income would raise unemployment.

I use the marginal personal income and social-security tax rate for the average wage earner (from OECD) as an estimate of the tax wedge. Within countries, there are no available regional data on tax rates or on the average wage.

Unemployment insurance

The terms and rates that apply to unemployment insurance programs affect the incentives for the unemployed to accept jobs and for employed workers to avoid unemployment. Most unemployment-insurance programs have two basic dimensions: duration of benefits and the "replacement rate," the benefit as a share of previous wages. There are also myriad other program-specific details such as strictness with which search activity is monitored, maximum monthly benefit amount, and sometimes the replacement rate varies with duration (declining as the length of the unemployment spell increases) and with family size.

The OECD data on "social protection and well-being" include series for replacement rates for various durations and demographics. Most of the variation in unemployment programs in these countries is in replacement rates, so I use the initial (at time of job loss) unemployment insurance replacement rate for a married individual with two children as a single representative measure of the generosity of unemployment benefits. As with tax rates, there is no regional variation in this variable.

Minimum wage in relation to market wage

Many studies find that unemployment is higher when the minimum wage is higher relative to the median or mean of market wages. Minimum wage laws in Europe are national, but they may have different effects across regions depending on the regional wage. They may have a larger effect on unemployment in low-wage regions than where wages are higher.

A good measure would use median (or even better, a lower-percentile) rather than mean regional wage to eliminate effects of tail-end high earners, but no data were found on mean or median regional wages. The measure I use divides the national minimum wage by the regional average labor compensation, both reported by Eurostat. All compensation data are

available only at the coarser TL2 level of regional disaggregation, so cross-regional variation in this variable is limited to differences across larger regions.

Other regional characteristics

Union coverage

One of the most important labor-market institutions is the extent to which bargaining is done collectively through unions or at an individual level. Wage bargaining is a complex process that can occur at national, industry, or individual-firm levels. Unionization rate is an imperfect measure of the degree and nature of collective bargaining. Greater union presence could increase unemployment if it increases inter-industry wage differentials and leads to queuing for highly-paid union jobs. Or it could lower unemployment rates if unions allow the labor market to overcome information or mobility imperfections.

The OECD measure of union coverage is national in scope and varies relatively little over time so, as with the policy variables, it has limited variation independent of country dummies.

Urban vs. rural regions

Two measures were considered for the urban/rural nature of regions. Population density has more variation than the OECD's three-level classification scheme of "predominantly urban," "predominantly rural," and "intermediate."[4] Population density is used in the regressions reported below.

Proximity to Western markets

Per-capita incomes in the bordering Western countries were much higher than those in the V-4 at the end of Soviet domination and they remain so today. To the extent that the development of capitalist industries in the V-4 has been driven by exports to Western countries (particularly those in the EU), we would expect, other things equal, regions closer to Western markets to see greater increases in employment (and lower unemployment) than in more remote regions.

[4] Earlier OECD data included a fourth category of "predominantly rural, close to a city." There are no regions within the V-4 countries with this classification in the current OECD database, so I believe that the category has been eliminated.

To measure land distance, I computed the distance from the largest city in each region to the nearest large city in Germany or Austria. The relevant cities are Berlin for northern Poland, Dresden for southern Poland and northern Czechia, and Vienna for southern Czechia and all of Slovakia and Hungary. (Given the more rapid transition of East Germany relative to the V-4, I consider Dresden to be a "Western" market despite its location in the formerly Communist part of Germany.)

For some goods and regions, shipment by sea may be an alternative to land shipment. Three of the V-4 countries are land-locked, but Gdansk is a major port on the Baltic Sea that may facilitate exports from nearby regions of Poland.

The measure that I use combines both land and sea access: it is the lesser of the distance to the nearest Western city and the distance to Gdansk. I expect this variable to have a positive sign: regions farther from Western markets are likely to have higher unemployment.

Roma population

The countries of East-Central Europe are home to relatively large sub-populations of Roma ("gypsies"). The Roma within these countries are often geographically segregated, poorly educated, and weakly integrated into the broader national labor market.[5] They also tend to be under-sampled in the population surveys on which unemployment data are based.

It is difficult to assess the impact on measured unemployment we should expect from a large Roma population. On one hand, they could simply be "invisible" to the data, with Roma showing up as neither employed nor unemployed and with their low education levels excluded from regional averages due to under-sampling. On the other hand, to the extent that they are included in surveys, Roma without formal jobs may be counted as unemployed and lead to a higher regional unemployment rate.

Table 3, taken from Figure 3 of Kahanec (2014), shows estimates of 2011 differences in formal employment rates of Roma and non-Roma in three sample countries. (The share of Roma in Poland is small, so it is not included in many studies of Roma in the region.) Given these large differences in labor-market outcomes and the uneven spread of Roma among the regions in the sample, a measure of the Roma share of the population would likely be a useful explanatory variable for regional unemployment and employment rates.

[5] For more detailed information about the labor-market status of Roma in the sample countries, see Vašečka, Jurásková, and Nicholson (2003), Messing (2014), and Brozovicova et al. (2013).

Table 3. Estimated Roma and non-Roma employment rates

	Employment Rates (%)			
	Males		Females	
Country	Roma	Non-Roma	Roma	Non-Roma
Czechia	43	76	19	64
Hungary	34	45	13	37
Slovakia	20	46	9	30

However, most experts believe that formal census data severely undercount the number of Roma. Undercounting occurs for several reasons, including the concentration of some Roma groups in secluded rural colonies, language differences, and the reluctance of some respondents (especially those of mixed ethnicity) to identify as Roma. Kahanec (2014) provides (Figure 1) the estimates shown in Table 4 of the degree of underestimation of the Roma population in official census sources.

Table 4. Estimated undercounting of Roma in official figures

	Roma population	
	Official	Experts' estimates
Czechia	12,444	179,778
Hungary	205,720	520,000 – 650,000
Slovakia	98,170	320,000 – 380,000

To compute the estimated share of Roma in each region and each year, I have either used expert estimates (when available at a regional level) or inflated official census estimates for each region using the national ratio of the midpoint of the experts' range of estimates to the official estimates. Detailed information on the Roma variable is presented below for each country in the sample. Slovakia is the only country for which temporal variation in Roma share is measured; for the other three countries the share of Roma is held constant over time at the 2002 (2001 for Hungary) level.

Hungary

Regional estimates of Roma population in 2001 are taken from Table 9 of Hablicsek (2008). These are divided by OECD data on regional population in 2001 to get a share. Lacking any information on changes over time, the 2001 share is used throughout the sample.

Slovakia

Annual regional estimates of Roma population share are available from the Slovak Statistical Office. I inflate the Census numbers for each year by a factor of 3.565 based on experts' estimates of undercounting in Table 4.

Czechia

Table 3 of Horáková (1997 (updated 2002)) provides estimated shares of Roma in each region according to the 2002 census. These numbers are inflated by a factor of 14.447 to correct for undercounting as in Table 4. The 2002 estimate is used for all years in the sample due to lack of time-series information.

Poland

Table 1 of Topińska (2011) gives "authority estimates" of the number of Roma in each larger region of Poland in 2002. These numbers are used for all years in the sample. The shares are assumed to be the same for each smaller region within larger regions.

Analysis and Results

The 1990s were a period of "transition" from Communism. The nascent capitalist economies of the V-4 countries were just beginning to emerge and both industrial structure and labor-market institutions were not yet well-established. Data are available for the variables outlined above beginning in 2001–14, which is a reasonable choice for the beginning of the sample for economic reasons as well. The actual sample size varies depending on the presence of lagged variables and some missing observations in early years for some of the 88 regions.

Numerous alternative estimation procedures may be applied to panel data. Basic OLS is informative but neither takes advantage of the panel nature of the sample nor corrects for the correlations it likely introduces. Including standard fixed effects would be an obvious econometric approach, but it creates problems estimating the effect of the many variables in the model that are very stable over time. A compromise approach is to include country-level fixed effects but not effects at the level of individual regions.

A random-effects model would avoid this limitation of fixed effects, but suffers from the usual potential for bias-inducing correlation between the included regressors and the random

effects themselves. Time-series correlation of the error can be accommodated through allowing each region's error term to follow a first-order autoregressive process. All of these estimators are presented in the results tables below.

Summary statistics

Table 5 shows sample statistics for the main variables in the model over the most inclusive estimation sample. The bivariate correlations are in Table 6.

Only two pairwise correlation coefficients between regressors reach 0.7 (between the V-4 GDP gap and industry-predicted employment growth and between minimum wage ratio to average wage and share of agricultural employment). This suggests that, on a bivariate level, collinearity is not severe. Variance inflation factors for the regressions suggest that multicollinearity is possible only among the country dummies and the region-invariant policy variables (taxes and unemployment-insurance replacement rate).

Table 5. Summary statistics for estimating sample

Variable	Number of Obs.	Mean	Standard Deviation	Minimum	Maximum
Unemployment rate (%)	1,048	11.9	6.23	1.9	34.5
Industry predicted emp. growth (%)	655	0.33	1.86	–4.49	4.78
Secondary education only (%)	1,048	71.5	6.2	53.5	83.1
Tertiary education (%)	1,048	16.7	5.5	6.9	40.6
Union membership (%)	1,038	17.3	3.7	10.7	32.3
Own-country GDP gap (%)	1,048	–1.22	2.81	–5.53	5.67
V-4 GDP gap (%)	1,048	–1.14	2.27	–4.36	2.88
Young adult population (% 15–24)	925	14.7	1.83	9.53	18.35
Minimum wage / average compensation	888	0.40	0.11	0.18	0.78
Roma population (%)	1,048	2.40	3.77	0.018	19.1
Agricultural employment (%)	1,048	11.9	10.0	0.075	56.0
Distance to Western markets (miles)	1,048	149.1	76.5	0	361.2
Initial unemp. ins. replacement rate (%)	965	50.1	11. 5	35.1	71.1
Marginal tax rate at average wage (%)	1,048	42.3	6.21	34.1	55.8
Population density (1,000/per sq. km.)	1,048	0.298	0.622	0.046	3.47

Table 6. Bivariate correlation coefficients

	Unemp	Ind. gr.	Sec. ed.	Tert. ed.	Union	Own gap	V-4 gap	Young	Min wage	Roma	Ag. emp.	Dist. mkts.	UI rep. rate	Tax rate
Unemployment rate	1.00													
Ind. pred .emp. growth	–0.30	1.00												
Secondary education only	0.14	0.06	1.00											
Tertiary education	–0.29	0.21	–0.67	1.00										
Union membership	0.06	0.04	0.42	–0.45	1.00									
Own GDP gap	–0.58	0.51	–0.13	0.14	–0.10	1.00								
V-4 GDP gap	–0.35	0.69	–0.10	0.31	–0.31	0.67	1.00							
Young adult pop. share	0.54	0.01	0.12	–0.26	0.42	–0.36	–0.13	1.00						
Min. wage / ave. comp.	0.14	–0.23	–0.40	0.04	–0.26	–0.11	–0.05	0.40	1.00					
Roma pop. share	0.01	0.01	–0.03	–0.12	0.07	0.18	0.02	–0.19	–0.27	1.00				
Ag. emp. share	0.20	–0.34	–0.41	–0.05	–0.11	–0.15	–0.11	0.41	0.70	–0.18	1.00			
Dist. to Western markets	0.23	–0.14	–0.24	0.02	–0.16	–0.09	–0.01	0.32	0.44	0.12	0.46	1.00		
Initial unemp. ins. rep. rate	–0.34	0.05	0.23	–0.16	0.09	0.30	0.12	–0.65	–0.60	0.47	–0.44	–0.35	1.00	
Marg. tax rate at ave. wage	–0.36	–0.01	–0.26	–0.09	0.06	0.34	–0.03	–0.47	–0.45	0.52	–0.13	–0.10	0.64	1.00
Population density	–0.07	0.15	–0.14	0.31	–0.03	–0.02	–0.00	–0.16	–0.15	–0.14	–0.35	–0.01	–0.09	–0.05

Measuring business cycles

As discussed above, the appropriate level of geographical aggregation at which to measure business cycles is not clear, given the differences in the cycles of the four countries. Table 7 shows OLS estimates of three models using the OECD measures for the GDP gap. The "Own gap" measure is the GDP gap at the individual country level. The "V-4 gap" is the gap aggregated across all four countries. The final column of the table shows regressions with the V-4 gap plus interaction terms with the country dummies. All gap measures are lagged one year. Each equation includes the indicated set of other variables.

Included without interactions, each GDP gap measure has the expected negative sign and is strongly statistically significant. The magnitudes of the coefficients suggest that a one-percentage-point drop in GDP below potential output leads to a rise in unemployment of ¾ to 1 percentage point. This value is quite large compared to the conventional "Okun's Law" values for the United States, which are typically assumed to be 0.3 to 0.5. It is surprising that the V-4 gap actually yields a slightly better fit than allowing each country's business cycle to

vary. The best fit is the final column of Table 7 with the lagged V-4 gap interacted with country dummy variables. All countries have negative cycle effects, but the effects are not statistically significant for Czechia and Hungary.

The estimated effects for the other variables of the model will be discussed in more detail below for two business-cycle specifications: the own-country GDP gap and the V-4 gap with interactions. Note from Table 7 that two variables—young adult population share and distance to Western markets—are economically and statistically insignificant in all specifications. These variables are dropped from the subsequent analysis.

Estimates using V-4 business cycle

Table 8 shows the results of alternative estimation methods using the model with the V-4 GDP gap and country interaction terms. Column (1) reports the estimates from using OLS with country effects and cluster-robust standard errors; column (2) uses the Prais-Winsten procedure to correct for first-order autocorrelation. Column (3) shows results from a standard fixed-effects model (at the regional level) and column (4) reports the corresponding random-effects model. Columns (5) and (6) repeat fixed-effects and random-effects estimation allowing the error to follow a first-order autoregressive process. The final column (7) is the "between regions" estimator using as a single observation the regional mean (over time).

The country-effect coefficients are consistent in sign and order across models and suggest that unemployment was, other things equal, highest in Poland, followed in order by Hungary, Slovakia, and (the omitted country) Czechia. As shown by the asterisk notations in Table 8, the coefficients are individually significant in nearly all cases and are collectively significant in all models. The pair-wise differences between the coefficients of Hungary, Poland, and Slovakia vary in statistically significance.

Table 7. OLS comparison of business-cycle variables

	Own gap	V-4 gap	V-4 & interactions
GDP gap variables			
Own GDP gap (lagged)	-0.743		
	(0.076)**		
V-4 GDP gap (lagged)		-1.042	-0.179
		(0.103)**	(0.143)
Country effects			
Hungary	-0.708	0.300	2.281
	(1.113)	(1.125)	(1.128)*
Poland	16.931	15.174	6.160
	(1.389)**	(1.356)**	(2.029)**
Slovakia	5.835	6.710	3.470
	(0.634)**	(0.628)**	(0.805)**
Interactions			
Hungary * V-4 gap (lagged)			-0.102
			(0.339)
Poland * V-4 gap (lagged)			-1.506
			(0.156)**
Slovakia * V-4 gap (lagged)			-0.845
			(0.185)**
Other variables			
Industry-predicted employment growth (lagged)	-0.163	-0.044	-0.116
	(0.117)	(0.123)	(0.119)
Secondary education only	-0.303	-0.275	-0.160
	(0.073)**	(0.073)**	(0.070)*
Tertiary education	-0.487	-0.460	-0.290
	(0.074)**	(0.074)**	(0.072)**
Young adult share	0.051	0.073	0.046
	(0.177)	(0.176)	(0.166)
Roma share	0.572	0.567	0.605
	(0.051)**	(0.050)**	(0.048)**
Population density	-0.397	-0.374	-0.660
	(0.275)	(0.274)	(0.260)*
Union coverage	-0.010	-0.239	-0.074
	(0.073)	(0.077)**	(0.085)
Agriculture emp. share	-0.101	-0.093	-0.119
	(0.025)**	(0.025)**	(0.024)**
Distance to Western market	-0.001	-0.001	-0.002
	(0.002)	(0.002)	(0.002)
Tax rate at average wage	0.253	0.217	-0.053
	(0.094)**	(0.094)*	(0.125)
UI replacement rate	0.334	0.217	-0.048
	(0.062)**	(0.058)**	(0.069)
Min. wage / comp.	-4.683	-4.209	1.088
	(2.418)	(2.414)	(2.361)
Constant	6.728	15.604	29.010
	(9.945)	(9.744)	(10.998)**
R^2	0.70	0.71	0.74

*p value < 0.05, **p value < 0.01, standard errors in parentheses below coefficients

The strongly significant coefficients (both individually and collectively) on the interaction terms with the GDP gap reflect the large differences in the business cycle shown in Figure 3. Unemployment in regions of the omitted country (Czechia) is countercyclical to lagged V-4 output fluctuations, with an "Okun's Law" coefficient in the 0.3–0.5 range that is typical of estimates for the United States. The cycle appears to affect unemployment much more strongly in Slovakia and especially Poland, which is consistent with the visual evidence of Figure 3. The Hungary interaction term is consistently positive and in some specifications large enough to make the overall business-cycle effect for Hungary (the sum of its interaction term and the gap coefficient) positive rather than negative. Hungary seems to have a very different business cycle than the other three countries, which is noticeable in Figure 3 and reflected in the estimates of Table 8.

The (lagged) industry-based prediction of regional employment growth should have a negative effect on unemployment: regions with high concentrations of rapidly-growing industries should have lower unemployment as a result. However, the estimates of Table 8 are more often positive than negative and, except in one model, not statistically significant.

The estimated effects of secondary and tertiary education conform to expectations. Both coefficients are consistently negative (though not always significantly so) and the effect of tertiary education is larger than the effect of secondary education. The difference between these coefficients is statistically significant except in the "between" model and the fixed-effects model with AR(1) error.

The strong and negative coefficient (except in the between regression) on union density is surprising. This variable is measured at the national rather than the regional level, which means that its estimated effect is based exclusively on the relatively small changes that occur over time, which may not provide a sufficient basis for estimating the true effect of union bargaining on unemployment in these countries. The coefficients of population density and agriculture are inconsistent in sign across estimation methods and statistically significant only in OLS.

Table 8. Alternative estimation methods with V-4 GDP gap and interactions

	(1) OLS	(2) AR(1)	(3) Fixed effects	(4) Random effects	(5) FE with AR(1)	(6) RE with AR(1)	(7) Between
Country effects							
Hungary	3.519	3.025		4.740		3.845	6.793
	(1.394)*	(1.164)*		(1.258)**		(1.218)**	(4.829)
Poland	4.781	6.411		4.923		6.001	14.090
	(1.039)**	(0.921)**		(1.085)**		(1.041)**	(2.716)**
Slovakia	2.687	4.690		3.090		4.044	1.732
	(1.110)*	(1.138)		(1.174)**		(1.107)**	(2.188)
Gap and interactions							
V-4 GDP gap_{-1}	-0.393	-0.361	-0.443	-0.417	-0.500	-0.357	
	(0.114)**	(0.081)**	(0.096)**	(0.096)**	(0.101)**	(0.104)**	
Hungary * V-4 gap_{-1}	1.019	0.730	0.590	1.011	0.387	0.835	
	(0.140)**	(0.093)**	(0.096)**	(0.117)**	(0.091)**	(0.132)**	
Poland * V-4 gap_{-1}	-1.615	-1.315	-1.978	-1.584	-1.796	-1.433	
	(0.132)**	(0.109)**	(0.087)**	(0.097)**	(0.093)**	(0.109)**	
Slovakia * V-4 gap_{-1}	-1.088	-0.564	-1.472	-1.112	-0.949	-0.788	
	(0.175)**	(0.098)**	(0.131)**	(0.126)**	(0.127)**	(0.145)**	
Industry effects							
Ind.-pred. emp. $growth_{-1}$	-0.014	0.086	0.023	0.010	0.217	0.047	0.362
	(0.103)	(0.077)	(0.077)	(0.077)	(0.076)**	(0.075)	(1.667)
Demographic effects							
Secondary educ. only	-0.157	-0.118	-0.036	-0.081	-0.087	-0.103	-0.184
	(0.103)	(0.065)	(0.103)	(0.080)	(0.105)	(0.076)	(0.136)
Tertiary education	-0.322	-0.345	-0.266	-0.310	-0.064	-0.355	-0.276
	(0.085)**	(0.068)**	(0.082)**	(0.066)**	(0.109)	(0.063)**	(0.130)*
Union density	-0.245	-0.256	-0.291	-0.276	-0.145	-0.265	2.695
	(0.058)**	(0.047)**	(0.048)**	(0.047)**	(0.073)*	(0.044)**	(1.630)
Roma share	0.556	0.511	-0.571	0.474	-0.570	0.486	0.575
	(0.081)**	(0.084)**	(0.351)	(0.096)**	(0.522)	(0.092)**	(0.102)**
Population density	-0.620	-0.433	3.066	-0.478	14.797	-0.364	-1.192
	(0.433)	(0.471)	(6.963)	(0.471)	(11.390)	(0.449)	(0.561)*
Agriculture emp. share	-0.105	-0.049	0.031	-0.072	0.147	-0.053	-0.112
	(0.038)**	(0.035)	(0.069)	(0.035)*	(0.079)	(0.034)	(0.109)
Policy effects							
Marginal tax rate	-0.407	-0.319	-0.451	-0.441	-0.350	-0.370	†
	(0.084)**	(0.069)**	(0.060)**	(0.059)**	(0.058)**	(0.057)**	
IU replacement rate	-0.016	0.020	-0.000	-0.005	-0.001	0.019	†
	(0.019)	(0.014)	(0.023)	(0.023)	(0.022)	(0.021)	
Constant	46.031	37.183	41.672	41.169	30.084	38.679	-23.510
	(9.810)**	(6.825)**	(10.159)**	(7.824)**	(5.278)**	(7.463)**	(33.128)
Autocorr. coefficient		0.74			0.51	0.51	
R^2	0.75	0.71	0.76				0.74

*p value < 0.05, **p value < 0.01, standard errors in parentheses below coefficients, cluster-robust standard errors in (1) and (2) †collinear with country effects

The Roma share variable is very strongly correlated with regional unemployment rates and has the expected positive sign, except in the fixed-effect models where its effect is undermined by the regional effects themselves.[6] The *t* statistics on this variable (in the models other than fixed effects) range exceed 5, suggesting that Roma presence is an important correlate of unemployment.

The two policy variables show surprising results. The marginal tax rate has a large and negative effect that is statistically significant in all models. As with union density no data are available at the regional level, so most of the variation in this variable is across countries and correlated with the country dummies. Unemployment-insurance replacement rates show no effect on the level of unemployment. Given the limitations of the data, this is not a surprising result because, again, there is little time variation and no intra-national regional variation in this variable.

I have presented the estimates from seven different estimation methods in Table 8 so that the reader can see the alternative results. Given that the interesting variation in unemployment that motivated this study is across regions, the fixed-effect model that "dummies out" this variation seems less interesting. However, random-effects models may suffer from correlation between the (random) region effects and the other variables of the model, leading to inconsistent estimates. A Hausman test may be used to compare the fixed-effects and random-effects estimates, implicitly testing for consistency of the random-effects estimator. This test for Table 8 returns a test statistic of $\chi^2_{12} = 15.7,$ with a *p* value of 0.21. Thus, we cannot reject the consistency of the random-effects estimator for these regressions.

The "between" estimator ignores all variation over time, which makes it impossible to estimate coefficients on business-cycle variables or on variables that are observed only at a national level (which become collinear with the country dummies). It is included here for comparison purposes. The three estimators with the autoregressive error model show a serial-correlation coefficient larger than one-half, so these may be preferred to the estimators that assume no serial correlation. These considerations point to column (6) as potentially the best model.

[6] Recall that the Roma share in my data varies over time only for the eight regions of Slovakia. No time-series variation could be estimated for the other three countries.

Estimates using country-specific business cycle

Table 9 presents an alternative set of corresponding estimates using the own-country-specific GDP gap measures. Some estimates are quite similar, but there are some variables for which the conclusions differ across the two business-cycle measures.

As expected, the business-cycle measures are negatively associated with unemployment. Allowing aberrant Hungary to have its own cycle improves the coherence of this result. For all but the fixed-effects models (which suffer from the same limitations discussed with respect to Table 8) the industry-predicted employment-growth variable has the expected negative sign and a high level of statistical significance, in contrast with the insignificant and positive effects in Table 8.

The results for the demographic variables in Table 9 are quite similar to those in Table 8, though there are variations across estimation methods. Region/years with higher education levels have lower unemployment. Union density has a similar negative effect in most models. The share of Roma population is strongly significant and positive (except in the fixed-effects models). Population density and agricultural-employment share do not have strong and consistent effects.

A strong contrast between Table 8 and Table 9 is in the estimated policy effects. The coefficients on the marginal income-tax rate and the unemployment-insurance replacement rate are significantly positive in Table 9. These effects confirm the usual implications of economic theory about how these policies should affect unemployment.

The interpretation of the estimates in Table 9 is made more difficult by the strong rejection of the validity of the random-effects model by the Hausman test. The test statistic is $\chi^2_{10} = 39.1,$ which has a p value less than 0.001. Thus, while Table 9 tells a reasonably coherent story about unemployment, it is difficult to decide which estimator should be used. Fixed-effects emasculate variables that vary most across regions; random-effects seem to violate the independence assumption according to the Hausman test; OLS is inefficient because the coefficient estimates do not account for within-region correlation in the error (though the robust standard errors do). Given the rejection of the random-effects model and the limitations of the fixed-effects model in failing to estimate cross-regional variation, the Prais-Winsten estimator of column (2) may be the most reliable.

Table 9. Alternative regression methods with own-country GDP gap

	(1) OLS	(2) AR(1)	(3) Fixed effects	(4) Random effects	(5) FE with AR(1)	(6) RE with AR(1)	(7) Between
Country effects							
Hungary	-1.667	-1.249		-0.986		-1.267	6.793
	(1.380)	(1.123)		(1.306)		(1.247)	(4.829)
Poland	16.493	12.583		15.647		13.767	14.090
	(1.013)**	(1.017)**		(1.000)**		(1.010)**	(2.716)**
Slovakia	6.187	6.338		6.760		6.575	1.732
	(1.170)**	(1.165)**		(1.168)**		(1.099)**	(2.188)
GDP gap							
Own GDP gap$_{-1}$	-0.808	-0.389	-0.767	-0.790	-0.492	-0.497	
	(0.054)**	(0.044)**	(0.047)**	(0.047)**	(0.045)**	(0.048)**	
Industry effects							
Ind.-pred. emp. growth$_{-1}$	-0.293	-0.201	-0.143	-0.205	0.039	-0.233	0.362
	(0.094)**	(0.066)**	(0.074)	(0.074)**	(0.076)	(0.075)**	(1.667)
Demographic effects							
Secondary educ. only	-0.237	-0.221	-0.340	-0.328	-0.218	-0.247	-0.184
	(0.102)*	(0.079)**	(0.118)**	(0.088)**	(0.115)	(0.084)**	(0.136)
Tertiary education	-0.449	-0.489	-0.691	-0.678	-0.057	-0.567	-0.276
	(0.087)**	(0.076)**	(0.092)**	(0.070)**	(0.121)	(0.069)**	(0.130)*
Union density	-0.134	-0.088	-0.266	-0.231	0.300	-0.172	2.695
	(0.052)*	(0.040)*	(0.045)**	(0.043)**	(0.076)**	(0.048)**	(1.630)
Roma share	0.532	0.489	-0.669	0.425	-0.493	0.469	0.575
	(0.081)**	(0.081)**	(0.392)	(0.098)**	(0.618)	(0.093)**	(0.102)**
Population density	-0.381	-0.137	11.480	0.128	29.636	0.006	-1.192
	(0.4457)	(0.502)	(7.835)	(0.478)	(13.741)*	(0.462)	(0.561)*
Agriculture emp. share	-0.125	-0.060	0.096	-0.095	0.335	-0.074	-0.112
	(0.039)**	(0.037)	(0.077)	(0.037)*	(0.088)**	(0.036)*	(0.109)
Policy effects							
Marginal tax rate	0.295	0.155	0.117	0.187	0.228	0.184	†
	(0.069)**	(0.058)**	(0.063)	(0.062)**	(0.053)**	(0.056)**	
UI replacement rate	0.307	0.148	0.327	0.318	0.084	0.200	†
	(0.020)**	(0.016)**	(0.020)**	(0.020)**	(0.024)**	(0.022)**	
Constant	2.364	16.564	27.163	18.359	-3.002	16.770	-23.510
	(9.173)	(7.400)*	(11.582)*	(8.454)*	(5.220)	(8.138)*	(33.128)
Autocorr. coefficient		0.76			0.56	0.56	
R^2	0.71	0.62	0.68				0.74

*p value < 0.05, **p value < 0.01, standard errors in parentheses below coefficients, cluster-robust errors in (1) and (2), †collinear with country effects

Discussion

Given the obvious differences between Hungary's business cycle and those in the other V-4 countries, separate measurement of cycles at the national level (as in Table 9) makes sense. However, the equations using the V-4 cycle with dummy interactions for each country in Table 8 have slightly better fit in terms of the R^2 criterion.

It is, of course, risky to choose one's estimation model based on their conformity with prior expectations, but the models of Table 9, are also "better" in terms of consistency with economic theory. Both sets are presented here to allow the reader to assess the alternatives and the presence (or lack) of robustness across specifications.

Several effects are strongly robust. A high share of Roma population is associated with high unemployment in all models (except the fixed-effect models where regional variation is eliminated). Given the presumption of omitted variables, one must be very careful about causality, both here and throughout the models. There are several reasons why a large Roma population might contribute to high regional unemployment. Most obviously, as shown in Table 3, Roma have very low employment levels. Those who are not employed may be either unemployed or out of the labor force, so low employment does not necessary lead to high unemployment. To the extent that Roma respondents to the population surveys indicate that they are seeking formal-sector work, they may count as unemployed and contribute directly to high unemployment. However, there may be other unmeasured characteristics of Roma-intense regions that cause high unemployment, so one must be very cautious in attributing causation directly to the Roma presence.

Other variables that are robust across specifications are education, business cycles (more or less), and union density. Unemployment moves countercyclically with respect to cycles in output, except in Hungary where the business cycle seems to be out of phase with the overall V-4. The Okun's Law coefficient of –0.39 (from column 2 of Table 9, with confidence interval –0.30 to –0.48) is consistent with estimates from the United States and elsewhere. Higher levels of high-school and college education attainment are generally associated with lower unemployment rates as expected. College education has a consistently larger effect than high-school and the difference is generally statistically significant. Union density usually has a negative association, which is perhaps surprising.

Some effects are less robust across cycle specifications. Region/years with a strong prevalence of fast-growing industries should have lower unemployment. The effect of industry-

predicted employment growth is negative, as expected, in Table 9, with the country-specific cycle measure, but not in Table 8, where the V-4 cycle is used. The coefficient of –0.2 (in column 2 of Table 9) indicates that a one-percentage-point increase in last year's V-4 employment growth rate weighted by the region's industry mix would be associated with a 0.2 percentage-point reduction in the unemployment rate (confidence interval 0.07 to 0.33).

The estimated policy effects also differ between the tables. Table 9 shows a statistically significant increase in unemployment as a result of higher taxes, consistent with the theoretical effects of a greater tax wedge in the labor market, but the opposite result holds in all specifications of Table 8. Using the column (2) model of Table 9, an increase of one percentage point in the marginal tax rate of the average wage-earner would raise unemployment by 0.16 points (confidence interval 0.04 to 0.27), a relatively small but plausible effect.

The replacement rate for unemployment insurance has no effect in the V-4-cycle model of Table 8, but has the expected positive and strongly significant effect on unemployment in Table 9. An increase of 10 percentage points in the replacement rate would be associated with a 1.5 percentage point (confidence interval 1.2 to 1.8) increase in unemployment.

While the lack of robustness in some effects across business-cycle measures may lower confidence in the results, the analysis reported in Table 9 confirms the predictions of economic theory about the unemployment effects of business cycles, industry composition, some demographic variables, and two major economic policy instruments. In particular, the strong estimated association of Roma population with unemployment suggests that more consistent time-series and cross-section data on this marginalized sub-population might lead to a better understanding of local labor markets.

Conclusions

The strongest conclusion of my analysis is the association between high unemployment and the concentration of Roma population. This result holds despite controlling for regional differences in population density, education, and share of employment in agriculture; it is not simply a result of "backward" rural regions having high unemployment. Scholars of the status of Roma will not be surprised by this result and their works (some cited here) provide important detailed information that helps us understand its underlying causes. Better data on the size, distribution, and demographic characteristics of the Roma minority would allow a more careful quantitative analysis.

Due to data limitations, my sample ends in 2013. Unemployment has fallen throughout the region in the last four years, so adding these years to the sample should provide considerable additional information once the required data are available.

It is also possible that additional countries from the region could eventually be added to the analysis, though most of these countries are small compared to the V-4 and have few regions. Slovenia and Estonia joined the OECD in 2010 and Latvia in 2016, so some comparable data are available for recent years. Lithuania is in the process of accession. Croatia and other states in the former Yugoslavia are not yet affiliated with the OECD. As time passes data from these countries may be sufficient to add them to the sample.

References

Borish, Michael S., and Michel Noël. 1996. Private Sector Development in the Visegrad Countries. *Finance and Development* 33 (4):45-48.

Brozovicova, Klara, Brian Fabo, Martin Kahanec, and Vera Messing. 2013. Overview of the Labour Market Situation of Low-Educated and Roma Population and Regulations Affecting their Employment. Bratislava: Central European Labour Studies Institute.

Hablicsek, László. 2008. The Development and the Spatial Characteristics of the Roma Population in Hungary: Experimental Population Projections till 2021. *Demográfia* 51 (5 (English Edition)):85-123.

Hodrick, Robert J., and Edward C. Prescott. 1997. Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking* 29 (1):1-16.

Horáková, Milada. 1997 (updated 2002). Romani in the Czech and Slovak Republic. Prague: Research Institute for Labour and Social Affairs.

Kahanec, Martin. 2014. Roma Integration in European Labor Markets. *IZA World of Labor* (39).

Messing, Vera. 2014. Patterns of Roma Employment in Europe. In *NEUJOBS Policy Brief*. Budapest: Center of Policy Studies, Central European University.

Nickell, Stephen, Luca Nunziata, and Wolfgang Ochel. 2005. Unemployment in the OECD since the 1960s: What Do We Know? *Economic Journal* 115 (500):1-27.

Orlandi, Fabrice. 2012. Structural Unemployment and Its Determinants in the EU Countries. Brussels: European Union. Economic Papers, No. 455.

Partridge, Mark D., and Dan S. Rickman. 1997. The Dispersion of US State Unemployment Rates: The Role of Market and Non-market Equilibrium Factors. *Regional Studies* 31 (6):593-606.

Topińska, Irena. 2011. Poland: Promoting Social Inclusion of Roma. Brussels: European Commissions DG of Employment, Social Affairs and Inclusion.

Vašečka, Michal, Martina Jurásková, and Tom Nicholson, eds. 2003. *Čačipen Pal o Roma: A Global Report on Roma in Slovakia*. Bratislava: Institute for Public Affairs.